Evaluating the causal relevance of diverse risk markers: horizontal systematic review

Hannah Kuper, senior lecturer,1 Amanda Nicholson, clinical research fellow,2 Mika Kivimaki, professor of social epidemiology,2 Amina Aitsi-Selmi, academic clinical fellow in public health medicine,2 Gianpiero Cavalleri, biomedical research lecturer,3 John E Deanfield, professor of cardiology,4 Peter Heuschmann, professor,5 Xavier Jouven, professor of cardiology,6 Sofia Malystina, professor,7 Bongani M Mayosi, professor of medicine,8 Susanna Sans, director of monitoring and research programme on chronic diseases,9 Troels Thomsen, consultant in preventive cardiology,10 Jacqueline C M Witteman, professor of cardiovascular epidemiology,11 Aron D Hingorani, professor of genetic epidemiology,2 Debbie A Lawlor, professor of epidemiology,12 Harry Hemingway, professor of clinical epidemiology2

ABSTRACT
Objectives To develop a new methodology to systematically compare evidence across diverse risk markers for coronary heart disease and to compare this with guideline recommendations.
Design “Horizontal” systematic review incorporating different sources of evidence.
Data sources Electronic search of Medline and hand search of guidelines.
Study selection Two reviewers independently determined eligibility of studies across three sources of evidence (observational studies, genetic association studies, and randomised controlled trials) related to four risk markers: depression, exercise, C reactive protein, and type 2 diabetes.
Data extraction For each risk marker, the largest meta-analyses of observational studies and genetic association studies, and meta-analyses or individual randomised controlled trials were analysed.
Results Meta-analyses of observational studies reported adjusted relative risks of coronary heart disease for depression of 1.9 (95% confidence interval 1.5 to 2.4), for top compared with bottom fourths of exercise 0.7 (0.5 to 1.0), for top compared with bottom thirds of C reactive protein 1.6 (1.5 to 1.7), and for diabetes in women 3.0 (2.4 to 3.7) and in men 2.0 (1.8 to 2.3). Prespecified study limitations were more common for depression and exercise. Meta-analyses of studies that allowed formal Mendelian randomisation were identified for C reactive protein (and did not support a causal effect), and were lacking for exercise, diabetes, and depression. Randomised controlled trials were not available for depression, exercise, or C reactive protein in relation to incidence of coronary heart disease, but trials in patients with diabetes showed some preventive effect of glucose control on risk of coronary heart disease. None of the four randomised controlled trials of treating depression in patients with coronary heart disease reduced the risk of further coronary events. Comparisons of this horizontal evidence review with two guidelines published in 2007 showed inconsistencies, with depression prioritised more in the guidelines than in our review.
Conclusions This horizontal systematic review pinpoints deficiencies and strengths in the evidence for depression, exercise, C reactive protein, and diabetes as unconfounded and unbiased causes of coronary heart disease. This new method could be used to develop a field synopsis and prioritise future development of guidelines and research.

INTRODUCTION
Clinical guidelines for the prevention of coronary heart disease are important not only because they influence practice but also because they present a highly cited collation of evidence for a multitude of risk markers.12 For example, the European primary prevention guidelines published in 2003 mentioned more than 40 risk markers and have been cited more than 800 times.

A fundamental problem in developing rational clinical guidelines has been the lack of explicit, systematic comparisons of the strength of causal evidence across the diverse range of risk markers, which compete for clinical attention. Traditional vertical systematic reviews, which focus on one risk marker or a relatively homogeneous group of related risk markers, are an important influence on the development of clinical guidelines. However, individual risk markers may be championed by different experts, with few attempts at harmonising, displaying, and comparing the evidence across different markers. This may contribute to an ad hoc selection not based on strength of causal evidence, of which risk markers beyond smoking, blood pressure, and cholesterol concentration are included in guidelines. The European guidelines,1 for example, did not consider atrial fibrillation, unlike the
We purposively selected one risk marker—diabetes—widely accepted as having an established causal role, as well as three markers where the causal role is not universally accepted. We hypothesised that concordance of research evidence from differing research designs each with different sources of error provides the strongest evidence on the causal relevance of a putative risk factor. Specifically, we sought evidence from three major study designs that offer different approaches to tackling confounding and reverse causation: traditional prospective observational studies with multivariate adjustments, studies that use genetic variants as instruments to tackle confounding (so called Mendelian randomisation), and randomised controlled trials where exposure to the risk marker is experimentally manipulated. Finally, in the light of the horizontal comparison we compared the recommendations made for these four risk markers in the most recent guidelines on prevention of coronary heart disease.

METHODS

The horizontal systematic review assesses causal relevance across a range of risk markers and study designs. We set out a priori eligibility criteria for studies, systematically obtained the studies, and extracted and displayed the data. Firstly, we separated information on risk markers for first coronary heart disease events in people initially free from clinical disease and prognostic factors in patients with existing coronary heart disease because of the clinical importance of distinguishing between primary and secondary prevention. Secondly, where more than one systematic review was identified we displayed the largest meta-analysis or study (which tended to be better quality according to the MOOSE,7 and QUORUM8 statements) in the tables of main results and included details of the others in the web extra appendix. Thirdly, we agreed a priori that if we could find no systematic review for any randomised controlled trials of the risk marker then we would review the largest individual study. Fourthly, we stratified data extraction and synthesis by study design (observational studies, genetic association studies, and randomised controlled trials), but analysed horizontally.

Observational studies

In January 2008 we searched Medline to identify meta-analyses of observational studies in healthy populations (aetiologic) and among patients with existing coronary disease (prognostic), and contacted experts. Existing coronary disease included patients with myocardial infarction or those undergoing coronary revascularisation or coronary angiography. Search terms included the expanded medical subject heading (MESH) of cardiovascular disease, meta-analysis as a MESH topic or publication type, and then the four individual risk markers as either the MESH term or text words. English and non-English language publications were eligible. Eligible outcomes were fatal coronary heart disease and non-fatal myocardial infarction (aetiologic and prognostic studies) and, for prognostic studies only, all cause mortality. Meta-analyses were only eligible for inclusion if they reported summary estimates based on longitudinal studies.

Two reviewers (HK and AAS) extracted data, with recourse to a third reviewer in the event of disagreement. We extracted summary data on prespecified items: age adjusted (or unadjusted) relative risks with 95% confidence intervals; adjusted relative risks with 95% confidence intervals; number of studies adjusting for smoking, blood pressure, and cholesterol concentration (aetiologic) or disease severity (prognostic); attenuation between age and multivariate adjusted relative risks; the prevalence of exposure in individual studies; the methods used to measure exposure; the number and type of outcomes; the level of meta-analysis (literature only, or pooled analysis of individual participant data across the studies); measure of heterogeneity; whether separate estimates were reported among people aged over 75 years, women, or non-Western populations; evidence of the presence of a dose-response relation—that is, extending beyond dichotomous comparisons; the extent to which the duration of follow-up influenced the strength of the estimates (if effects are stronger with shorter periods of follow-up this is consistent with reverse causality); and evidence of publication bias.

Genetic studies

We searched for meta-analyses of the associations between genotype and each of the four individual risk markers through Medline using gene as the MESH term or text words or Mendelian in any field; meta-analysis as a MESH topic or publication type; and then the four individual risk markers as either the MESH term or text words. To identify meta-analyses of the association between genotype and coronary heart disease outcome we searched individual single nucleotide polymorphisms in all fields identified from recent systematic reviews for depression (SLC6A4, MTHFR, APOE*4), the CRP gene, and type 2 diabetes (TCF7L2, FTO, CDKN2A/
CDKN2B, PPARG, ICF2BP2, KCNJ11, HHEXIDE, CDKAL1, SLC30A8<sup>15</sup>) together with expanded MESH headings of cardiovascular disease and meta-analysis as a MESH topic or publication type. For each variant two independent reviewers (HH and AN) extracted information on prespecified items: whether the single nucleotide polymorphism was identified from genome wide scans, the number of outcome events, the number of studies in the meta-analysis, unadjusted relative risks (95% confidence intervals), and whether there was a formal test on the use of the genetic variant as an instrumental variable.<sup>6</sup>

Randomised trials
Randomised controlled trials and meta-analyses of these were identified through searches of Medline. Search terms included the MESH heading of coronary heart disease or CHD or myocardial infarction or MI; trial as a MESH heading; the four individual risk markers as either the MESH term or text words; and meta-analysis or systematic review as a MESH heading or review as a publication type. Given the importance of evidence from randomised controlled trials to inform guidelines, we accepted individual trials where no meta-analyses were available. We also searched through the reference list from the guideline publications to identify relevant randomised controlled trials. Only randomised controlled trials that reported coronary heart disease event outcomes were eligible (aetiological and prognostic studies) or death in the setting of patients with coronary heart disease (prognostic studies). Two independent reviewers (HK and AAS) extracted details on the nature of the intervention, the number of studies in the meta-analysis, the number and type of end points, the relative risk (95% confidence interval) of coronary heart disease or death, and whether the intervention had an effect on the risk marker.

Selection of guidelines
To make contemporaneous comparisons with our evidence review we included guidelines only published in 2007 as this was the most recent information available to us. We identified two guidelines, which were developed through independent processes, from the Scottish Intercollegiate Guideline Network<sup>9</sup> and the fourth Joint European Societies<sup>10,11</sup> (coordinated by the European Society of Cardiology, and representing nine professional organisations). Across each guideline and risk marker we compared the evidence cited, description of the causal relevance of the marker, recommendations on measurement in healthy population settings, inclusion in risk scores, recommendations for specific interventions, and target levels or goals for risk marker levels.

RESULTS
The figure shows the results of the meta-analyses of observational studies, genetic variants, and randomised controlled trials for depression, exercise, C-reactive protein, and diabetes in relation to risk of coronary heart disease. These risk markers differed noticeably in the type and amount of evidence identified.

Observational studies
Table 1 summarises the largest meta-analyses found for each of the risk markers. Meta-analyses of observational studies reported adjusted aetiological relative risks of coronary heart disease for depression of 1.9 (95% confidence interval 1.5 to 2.4; 1262 events in 11 studies), for top compared with bottom fourth of exercise of 0.7 (0.5 to 1.0; 500 events in three studies), for top compared with bottom third of C reactive protein of 1.6 (1.5 to 1.7; 7068 events in 22 studies), and for diabetes in women of 3.0 (2.4 to 3.7) and in men of 2.0 (1.8 to 2.3; >4964 events in 29 studies). Meta-analyses of the association in patients with coronary heart disease (prognostic studies) reported an adjusted relative risk of further coronary heart disease events or death for depression of 1.6 (1.3 to 1.9; >525 events in 11 studies) and an adjusted relative risk of cardiac death for diabetes of 1.5 (1.1 to 2.0; 240 events in four studies). No prognostic meta-analyses were identified for the effect of exercise or C reactive protein on outcome among patients with coronary heart disease. A dose-response effect was reported for intensity of physical activity but not consistently for major depression compared with minor depression, concentration of C reactive protein, or glucose control in people with diabetes. None of the meta-analyses reported effects separately among those older than 75 years (or other age groups of older people), and only for diabetes was there evidence of consistent effects in non-Western populations. Effect estimates for C reactive protein and diabetes were presented separately for men and women, with observational evidence for diabetes showing a stronger association with coronary heart disease among women than among men. No sex differences were found for the association of C reactive protein with coronary heart disease. The exercise meta-analysis was restricted to women and the depression meta-analysis did not report effects separately in women and men.

Confounding—Adjustments for a priori confounders of smoking, blood pressure, and cholesterol concentration were found in four of 22 aetiological studies for depression and coronary heart disease and one of three studies for physical activity. Adjustments in the C reactive protein studies (20/22) and diabetes studies (26/37) were generally more consistent and complete, although beyond smoking it was unclear which variables were included in multivariate analyses. For aetiological meta-analyses of the four risk markers, the effect on coronary heart disease was apparent after multivariable adjustments. Reporting of unadjusted or age adjusted and multivariate adjusted results was inconsistent.

Biases—Statistical heterogeneity, present in all the meta-analyses, was partly attributable to differences in measurement of exposure for depression and physical activity, and year of publication for C reactive protein. Depression was defined by 12 different methods,
but relatively standardised methods were used for measurement of C reactive protein concentration and diabetes. For depression, but not for C reactive protein or diabetes, stronger effects were observed with shorter follow-up. Adjustment for severity of coronary heart disease in prognostic studies reduced the relative risk for depression by 45%. Evidence of small study bias (often indicative of publication bias or a strong association between methodological weakness and a non-null association in the expected direction in smaller studies compared with larger studies) was present for depression but absent for physical activity, diabetes, and C reactive protein.

**Other meta-analyses of observational studies**—The findings of the meta-analyses of observational studies included in our analysis were consistent with results of other eligible meta-analyses on this subject (see web extra appendix). For depression, adjusted relative risks ranged between 1.5 (1.4 to 1.6) and 1.9 (1.5 to 2.4). Similarly, a consistently strong association was found between diabetes and the incidence and mortality after coronary heart disease, with stronger relations identified for women than for men. Exercise was consistently protective for coronary heart disease, ranging from 0.5 (0.3 to 0.6) to 0.8 (0.6 to 1.2). Other meta-analyses confirmed the increased risk for coronary heart disease for people in the highest third of C reactive protein concentrations.

**Genetic evidence**

Table 2 summarises the results of our search for meta-analyses of genetic variants, indexing differences in the risk markers of interest. Meta-analyses for two genetic variants were identified, which have been investigated in relation to both depression and coronary heart disease: MTHFR and APOE. The MTHFR variant was positively associated with both depression and coronary heart disease, whereas the ε2 APOE genotype was linked to a reduced risk of depression and of coronary heart disease. No replicated genetic variant for physical activity was identified that could be assessed in relation to coronary heart disease. We identified two syntheses of a genetic variant in the CRP gene in which the expected relation for coronary heart disease events under a causal model, in the light of its effect on C reactive protein levels, was not observed. Despite the large number of meta-analyses showing genes associated with diabetes, none examined the association of these variants with coronary heart disease.

**Randomised controlled trials**

We identified four randomised controlled trials of different interventions for the treatment of depression among patients with coronary heart disease (table 3). None of these trials showed a beneficial effect on death or cardiovascular events. No randomised controlled trials were identified on the effect of treating depression in healthy populations (or on the effect of prevention of depression) in relation to risk of coronary heart disease. No randomised controlled trials...
<table>
<thead>
<tr>
<th>Variables</th>
<th>Depression</th>
<th>Exercise</th>
<th>C reactive protein</th>
<th>Type 2 diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcomes</td>
<td>Non-fatal myocardial infarction or fatal coronary heart disease</td>
<td>Death</td>
<td>Coronary heart disease</td>
<td>Fatal coronary heart disease</td>
</tr>
<tr>
<td>Comparison</td>
<td>Depression v no depression</td>
<td>Depression v no depression</td>
<td>Top fourth v bottom fourth</td>
<td>Top third v bottom third</td>
</tr>
<tr>
<td>Largest meta-analysis</td>
<td>Nicholson 2006 \cite{10}</td>
<td>Nicholson 2006 \cite{10}</td>
<td>Oguma 2004 \cite{11}</td>
<td>Danesh 2004 \cite{12}</td>
</tr>
<tr>
<td>No of outcome events</td>
<td>4016</td>
<td>1867</td>
<td>500</td>
<td>7068</td>
</tr>
<tr>
<td>No of studies in meta-analysis</td>
<td>22</td>
<td>34</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Measurement of exposure</td>
<td>Scale=16*; antidepressant=4; doctor diagnosed=2</td>
<td>Scale=34†</td>
<td>Leisure time physical activity=2; total physical activity=1</td>
<td>NR</td>
</tr>
<tr>
<td>Prevalence of exposure (%)</td>
<td>2-50</td>
<td>2-51</td>
<td>25</td>
<td>33</td>
</tr>
<tr>
<td>No of studies adjusted for smoking/cholesterol</td>
<td>4/22</td>
<td>13/14</td>
<td>20/22 reported adjustment for smoking and &quot;some other established risk markers&quot;</td>
<td>20/37</td>
</tr>
<tr>
<td>Heterogeneity (%) (any markers that explain)</td>
<td>52 (Measurement instrument)</td>
<td>54</td>
<td>NR</td>
<td>54 (year of publication)</td>
</tr>
<tr>
<td>Older people (≥75)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Women</td>
<td>NR</td>
<td>NR</td>
<td>Only women</td>
<td>Effect stronger in women</td>
</tr>
<tr>
<td>Non-Western populations</td>
<td>Studies lacking</td>
<td>Studies lacking</td>
<td>Studies lacking</td>
<td>Yes</td>
</tr>
<tr>
<td>Effect changes with duration of follow-up</td>
<td>Yes; stronger effects with shorter follow up</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Length of follow-up (years)</td>
<td>3-37</td>
<td>0.25-15</td>
<td>3-20</td>
<td>4-36</td>
</tr>
<tr>
<td>Publication bias</td>
<td>Present (Egger's test: P=0.08)</td>
<td>Present (Egger's test P=0.05)</td>
<td>No evidence (Egger's test P=0.45)</td>
<td>No evidence (Egger's test)</td>
</tr>
<tr>
<td>Do response</td>
<td>Evidence for dose response: stronger association for clinical depression (2.3, 1.8 to 3.1) than depressive symptom scale (1.7, 1.4 to 2.0)</td>
<td>No evidence for dose response: weaker association for clinical depression (1.4, 0.8 to 2.5) than depressive symptom scale (1.9, 1.6 to 2.3)</td>
<td>NS (P=0.20)</td>
<td>NR</td>
</tr>
</tbody>
</table>

NR not reported.

*General wellbeing schedule, \cite{1}.
†Center for Epidemiological Studies depression scale, \cite{5}.
‡Diagnostic interview schedule, \cite{2}.
\S Zung self rating depression scale, \cite{1}.
\| General health questionnaire/present state examination, \cite{1}.
\| Hamilton rating scale, \cite{2}.
\| Minnesota multiphasic personality inventory—depression, \cite{1}.
\| Diagnostic interview schedule, \cite{2}.
\| Center for Epidemiological Studies depression scale, \cite{6}.
\| Becks depression inventory—anorexia, \cite{7}.
\| Becks depression inventory combined, \cite{3}.
\| Depression anxiety stress scales, \cite{1}.

Center for Epidemiological Studies depression scale and Becks depression inventory combined were not formally investigated using Egger's test.
Table 2 | Genetic variants (single nucleotide polymorphisms, SNPs) associated with risk marker (depression, exercise, C reactive protein, and diabetes) and coronary heart disease in healthy populations

<table>
<thead>
<tr>
<th>Variable</th>
<th>MTHFR C677T (TT v CC) rs1801133</th>
<th>APOE carriers (ε2 v ε3/ε3)</th>
<th>No SNPs</th>
<th>CRP T1444C rs1130864 (TT v any C)</th>
<th>8 SNPs identified in recent review*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcomes</td>
<td>Depression</td>
<td>Coronary heart disease</td>
<td>Depression ε2 v ε3 allele</td>
<td>Coronary heart disease ε2 carriers ε3/ε3</td>
<td>Exercise</td>
</tr>
<tr>
<td>SNP identified from genome wide scans</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>—</td>
</tr>
<tr>
<td>No of outcome events</td>
<td>1280</td>
<td>26 000</td>
<td>827</td>
<td>21 331</td>
<td>—</td>
</tr>
<tr>
<td>No of studies in meta-analysis</td>
<td>10</td>
<td>80</td>
<td>7</td>
<td>17</td>
<td>—</td>
</tr>
<tr>
<td>Unadjusted relative risk (95% CI)</td>
<td>1.36 (1.11 to 1.67)</td>
<td>1.14 (1.05 to 1.24)</td>
<td>0.51 (0.39 to 0.68)</td>
<td>0.80 (0.70 to 0.90)</td>
<td>1.21 (1.09 to 1.43)</td>
</tr>
<tr>
<td>Instrumental variable test*</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Null finding (underpowered)</td>
</tr>
</tbody>
</table>

NA = not available.

*Testing whether risk marker is associated with coronary heart disease to extent it is associated with genetic variant—that is, exploiting part of phenotypic variation which is not related to potential confounding markers. A positive finding supports causality whereas a null finding suggests that observed association between risk marker and coronary heart disease may be confounded or due to reverse causality.

†Adjusted estimate.

were found of interventions that specifically increased physical activity [in isolation from, for example, improvements in diet, or compliance with drugs] in healthy populations, whereas exercise based rehabilitation reduced the risk of mortality among people with coronary heart disease (0.7, 0.6 to 1.0).9 Currently there are no interventions that specifically lower C reactive protein levels, and hence no randomised controlled trials that could test the causal importance of C reactive protein in coronary heart disease. Randomised controlled trials [four meta-analyses and 70 individual trials] of different hypoglycaemic agents in patients with diabetes without or with manifest coronary heart disease provided some, but not consistent, support that lowering glucose concentrations reduced the rate of coronary heart disease events.17-21

Guidelines

Neither the Scottish Intercollegiate Guideline Network nor the European guidelines adopted an explicit method for displaying and comparing evidence across risk markers [table 4]. The Scottish but not European guidelines reported a level of evidence for some statements. Both guidelines made clear, albeit differing, statements on the importance of depression in the onset and progression of coronary heart disease. Neither the only randomised trial with statistical power to detect differences in event rates (ENRICHD)22 nor the Cochrane meta-analysis of psychological interventions17 was cited in the executive summary of the European guidelines. Post hoc subgroup analyses were cited.16 The Scottish guidelines cited neither trials nor meta-analyses but did cite a previous position statement, which itself cites only an older narrative systematic review. Neither guideline suggested that C reactive protein was an important risk marker: for the Scottish guidelines it was neither mentioned nor any rationale given as to why not. Observational studies and Mendelian randomisation studies were cited for the European guidelines meta-analyses, and the association was stated as “often seriously confounded.”

DISCUSSION

The horizontal systematic review is a new method to compare the evidence on diverse risk markers in a unified explicit framework of the largest available syntheses of the most important forms of evidence. This approach highlighted differences and deficiencies in the evidence of causal relevance across the four selected risk markers: psychosocial, behavioural, biomarker, and metabolic disease. The evidence that depression, low physical activity, or C reactive protein concentration causes coronary heart disease seems less strong than that for diabetes. Randomised trials of specific interventions are lacking for C reactive protein and null for depression, and although they support the role of exercise in the secondary prevention of coronary heart disease they are not available to test the causal hypothesis. Neither the European nor Scottish guidelines gave explicit criteria for assessing evidence to enable prioritisation of the impact of individual risk markers. However the emphasis given to the causal
and clinical relevance of depression in the guidelines was inconsistent with the available evidence.

Closing the translational gap

Discordance exists between the large number of markers that are associated with coronary heart disease and the small number of targets for intervention. This highlights the need for our approach, which aims to prioritise targets. Our approach is horizontal in two senses: a comparison was made across diverse markers and a comparison was made across different forms of evidence. We selected four risk markers as examples; the approach is scaleable to all risk markers.

Three complementary designs for obtaining causal evidence

Associations between putative risk markers and coronary heart disease are easy to show in observational studies but may be confounded, as has been shown by negative trials of hormone replacement therapy and vitamins on coronary heart disease.19 We marshalled three approaches, which aimed, with varying limitations, to mimic the ideal, unconfounded experiment; prospective observational studies (multivariate adjustment for confounders), genetic studies (which utilise genetic variants that influence the modifiable exposure and that are assigned at random and can therefore be used as an instrument for the unconfounded and unbiased association of the genetic variant with the outcome of interest), and randomised trials (where the investigator influences exposure). Low density lipoprotein cholesterol provides an example with converging evidence from all these approaches: robust associations between high concentrations of low density lipoprotein cholesterol and coronary heart disease shown in observational cohort studies20; genetic variants that relate to lower concentrations of low density lipoprotein cholesterol (for example, in PCSK9,21 the APOE,22 and LDL receptor gene)23 also found to be associated with a reduced risk of coronary heart disease; and trials on low density lipoprotein cholesterol lowering confirmed the protection against coronary heart disease.24 We excluded study designs that may be associated with lower validity, such as individual observational studies that have not (yet) been synthesised, non-randomised trials, and studies on biological mechanisms. Such studies have been the basis for guideline recommendations.25 We included the forms of large scale evidence, which aims to evaluate the causal hypothesis, with a low tolerance for false positive findings. Our approach could be extended to incorporate small scale experimental studies in humans (as part of a “teleoanalysis” approach)26 and, with due caveats,27 experimental studies in animals. Such extensions to other forms of evidence should acknowledge that studies in the discovery phase have a higher tolerance for false positive findings, as the aim is not to abandon a potentially important risk marker prematurely.28

Observational evidence

Risk markers were associated with relative risks from 1.5 (C reactive protein) to over 3 (diabetes), but adjustment for established risk factors of smoking, blood pressure, and cholesterol concentration was least common among studies of depression and exercise. These psychosocial (depression) and behavioural (exercise) factors were also more prone to information biases, with multiple instruments used to determine exposure. There was evidence of reverse causality and publication bias for depression. We found no meta-analyses in a prognostic setting of C reactive protein or exercise. Given that the guidelines make recommendations in secondary prevention and since aetiological markers may not necessarily be prognostic, this lack of synthesised evidence is important. Thus, for example, meta-analyses of body mass index in the prognosis of patients with coronary disease suggest no adverse effect for obesity,29 whereas those for aetiological associations show an increased risk.30

Genetic evidence

Genetic studies using Mendelian randomisation have been more frequently applied to assessing C reactive protein than for our other three risk factors. Despite being relatively underpowered, the emerging evidence does not suggest an important role for C reactive protein in causing coronary heart disease. A new and large collaboration should provide a more definitive answer in the near future.24 For depression, exercise, and diabetes, evidence from Mendelian randomisation for their causal role in risk of coronary heart disease was limited. The robust positive associations of MTHFR with both depression and coronary heart disease could indicate a causal effect of depression on coronary heart disease but is more likely to reflect folate intake and metabolism as a causal marker for both outcomes. The emergence of whole genome association studies and complete genome sequencing is improving our understanding of the genomic architecture underlying complex traits. Mendelian randomisation may offer a powerful tool to understand causality, particularly for risk marker traits that are controlled by a limited number of genetic variants of relatively strong effect.

Randomised controlled trials

Successful treatment of depression in patients with established coronary heart disease in randomised controlled trials does not show benefits in subsequent death or rates of coronary heart disease events. This provides no support for the causal hypothesis that avoiding depression is important in the secondary prevention of coronary heart disease, but it is a matter of debate whether it provides evidence against. For example, it might be argued that it is the intervention rather than the hypothesis that is wrong. Trials were lacking for the effect of C reactive protein concentration or physical activity among healthy populations, whereas there was evidence that glucose control may provide some reduction in coronary heart disease events. Interpretation was difficult in the situation of trials with
### Table 3: Randomised trials of interventions that alter depression, exercise, C reactive protein, and glycaemic control and their effects on coronary heart disease events

<table>
<thead>
<tr>
<th>Population</th>
<th>Depressive disorder</th>
<th>Exercise</th>
<th>C reactive protein</th>
<th>Glycaemic control</th>
<th>Outcomes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy Coronary heart disease</td>
<td>Depression only</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Cardiac event</td>
<td>Jolliffe 2005</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>CBT or IPT</td>
<td>Exercise</td>
<td>Pioglitazone</td>
<td>Pioglitazone</td>
<td>Death</td>
<td>Huang 2001, Lincoff 2007</td>
</tr>
<tr>
<td>Coronary heart disease plus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-fatal event</td>
<td></td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coronary heart disease</td>
<td></td>
</tr>
<tr>
<td>Cardiac event</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mortality</td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Survival in (0.6 to 1.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower in intervention compared with control</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.6 to 1.0) to 2.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Impact of IPT on reducing mortality on death</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant but modest impact of IPT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>on mortality</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Significant but modest impact of IPT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>on mortality</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** IPT = interpersonal psychotherapy; CBT = cognitive behavioral therapy; NR = not reported.

---

### Clinical implications and consistency of the guidelines

By using horizontal systematic reviews, clinicians, guideline developers, funders of research, public health policy makers, and journal editors and their peer reviewers might be aided in making more consistent and less biased decisions. The graphical summary of evidence may serve a practical purpose in guideline groups, facilitating more explicit debate of the importance of risk markers across the multiple fields of interest of contributors. The two most recent guidelines on primary prevention of cardiovascular disease refer to a wide “penumbra” of risk markers beyond smoking, blood pressure, and cholesterol concentration; over 50 markers in the most recent European guidelines. The guidelines cite more than 1100 references (joint societies) and 315 references (Scottish Intercollegiate Guideline Network) but do not provide a systematic comparison of the quality or strength of evidence across the risk markers that were included. There was inconsistency in the conclusions reached by the two guidelines across the four risk markers that we evaluated. For instance, C reactive protein was not considered by the Scottish Intercollegiate Guideline Network. Depression was accorded higher prominence in the European guidelines than in the Scottish guidelines and within the European guidelines was accorded higher prominence than C reactive protein, which is not consistent with the evidence in the horizontal systematic review. Depression is worth treating in its own right, irrespective of any causal relation with coronary heart disease; but the same is true for other conditions, such as chronic obstructive airways disease, which are not mentioned in the guidelines.

### Limitations of this horizontal systematic review

There are important limitations in this initial illustration of a horizontal systematic review. Firstly, the method depends on the availability and quality of pleiotropic effects; for instance, the evidence that exercise in combination with other aspects of rehabilitation reduced the risk of death among patients with coronary heart disease.

---

### Method

By using horizontal systematic reviews, clinicians, guideline developers, funders of research, public health policy makers, and journal editors and their peer reviewers might be aided in making more consistent and less biased decisions. The graphical summary of evidence may serve a practical purpose in guideline groups, facilitating more explicit debate of the importance of risk markers across the multiple fields of interest of contributors. The two most recent guidelines on primary prevention of cardiovascular disease refer to a wide “penumbra” of risk markers beyond smoking, blood pressure, and cholesterol concentration; over 50 markers in the most recent European guidelines. The guidelines cite more than 1100 references (joint societies) and 315 references (Scottish Intercollegiate Guideline Network) but do not provide a systematic comparison of the quality or strength of evidence across the risk markers that were included. There was inconsistency in the conclusions reached by the two guidelines across the four risk markers that we evaluated. For instance, C reactive protein was not considered by the Scottish Intercollegiate Guideline Network. Depression was accorded higher prominence in the European guidelines than in the Scottish guidelines and within the European guidelines was accorded higher prominence than C reactive protein, which is not consistent with the evidence in the horizontal systematic review. Depression is worth treating in its own right, irrespective of any causal relation with coronary heart disease; but the same is true for other conditions, such as chronic obstructive airways disease, which are not mentioned in the guidelines.

---

### Limitations of this horizontal systematic review

There are important limitations in this initial illustration of a horizontal systematic review. Firstly, the method depends on the availability and quality of...
### Table 4: Guideline recommendations for primary prevention of coronary heart disease in relation to depression, exercise, C reactive protein, and diabetes in fourth Joint European Societies\(^{10,11}\) and Scottish Intercollegiate Network (SIGN) guidelines\(^{9}\) (both published in 2007)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Depression</th>
<th>Exercise</th>
<th>C reactive protein</th>
<th>Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4th Joint European Societies</td>
<td>SIGN</td>
<td>4th Joint European Societies</td>
<td>SIGN</td>
</tr>
<tr>
<td>Description of causal relevance</td>
<td><em>Increasing scientific evidence that psychosocial factors (including depression) contribute independently to the risk of CHD (coronary heart disease)</em> (E26)</td>
<td><em>strong and consistent evidence . . . independent risk factor</em> (p43); evidence level 2++</td>
<td><em>A lack of regular physical activity may contribute to the early onset and progression of CVD (cardiovascular disease)</em> (E19)</td>
<td><em>Independent risk factor</em> (p16); level of evidence 2++</td>
</tr>
<tr>
<td>Measurement recommended in opportunistic assessment of healthy populations</td>
<td>Yes:<em>Assess all patients for psychosocial risk factors</em> (E27)</td>
<td>No clear statement: <em>Depression...should be taken into account when assessing individual risk</em> (p44)</td>
<td>Yes: <em>Assessment . . . core task[s] for physicians and other health workers</em> (E20)</td>
<td>No</td>
</tr>
<tr>
<td>Method of measurement recommended</td>
<td>Yes:<em>clinical interview or standardised questionnaires</em>; two screening questions are proposed (E27)</td>
<td>No</td>
<td>Yes: <em>Brief interview concerning . . . physical activity at work and leisure</em>; no specific standardised instrument recommended (S38)</td>
<td>No</td>
</tr>
<tr>
<td>Use in risk prediction scores*</td>
<td>No</td>
<td>No (but social deprivation is included)</td>
<td>No (but a <em>key element of risk evaluation</em> (E19)</td>
<td>No</td>
</tr>
<tr>
<td>Interventions</td>
<td>Yes:<em>Prescribe multimodal, behavioural intervention, integrating individual or group counselling for psychosocial risk markers and coping with stress and illness. Refer to a specialist in case of clinically significant emotional distress</em> (E27)</td>
<td>No: <em>No clear evidence that treating depression is effective</em> (p43); <em>Smokers with CHD and comorbid clinical depression should have their depression treated both for alleviation of depressive symptoms and to increase the likelihood of stopping smoking</em> (p21)</td>
<td>Yes: <em>Professional advice about the intensity, duration and frequency of exercise</em> (S40)</td>
<td>Not described (national guidance cited)</td>
</tr>
<tr>
<td>Goals for risk marker change</td>
<td>None</td>
<td>None</td>
<td>Yes: <em>30 minutes of moderately vigorous exercise on most days of the week will reduce risk</em> (E19)</td>
<td>Yes: <em>All adults should accumulate 30 minutes of moderate intensity activity on most days of the week</em> (p17); level of evidence 4</td>
</tr>
</tbody>
</table>

**UKPDS**=United Kingdom Prospective Diabetes Study; **EASD**=European Society for Diabetes.

*Assessing cardiovascular risk using SIGN guidelines to assign preventive treatment (ASSIGN score) in SIGN and HeartScore in fourth Joint European Societies.
large scale syntheses of evidence. These are more commonly available for blood based markers than for behavioural or psychosocial markers; horizontal systematic reviews may stimulate research groups to raise or defend the profile of research in their subdisciplines. A range of measures of effect were included in the reviews, and where the confidence intervals for the effect estimates were wide this precluded reaching firm conclusions. Increasing use of horizontal systematic reviews may provide an impetus to improving the number and quality of meta-analyses, particularly those using individual participant data. Secondly, the horizontal systematic review is narrative, without novel methods for data analysis, offering no explicit ranking of causal relevance nor attempting to posit a decision threshold above which a marker might be considered causal.

Research implications and need for unbiased field synopses

Further research is required to develop the method of horizontal systematic review. Firstly, methods could be developed to derive relative weights of evidence building on the judgments of groups of experts,\(^ {37}\) Bayesian methods could be used for the synthesis of evidence, or models could be developed to combine features from different studies to derive quantitative estimates.\(^ {38}\) Secondly, extension is required to the whole range of risk markers that are included in guidelines, thus providing a systematic synopsis of the specialties. Thirdly, extension to other chronic diseases should be explored—for example, in the specialty of cancer, horizontal systematic reviews could build on the assessment of causality used by the International Agency for Research on Cancer. Fourthly, horizontal systematic reviews should be regularly updated as evidence changes to minimise the lag time between the generation of evidence and the development of guidelines and could take advantage of continually updated databases of genetic studies in this process. The conclusions of our review are not altered if the publication year of evidence is truncated two years before the publication of guidelines—that is, in 2005. Fifthly, there are important considerations beyond causal relevance when developing guidelines, such as economic considerations and the additional deleterious effects of the risk markers (for example, the impact of depression on quality of life), and the framework could be extended to encompass these considerations. Sixthly, it should be noted that non-causal markers can be used in risk prediction (for example, socioeconomic position) and this requires distinct consideration in the observational evidence.

Conclusion

Horizontal systematic review in which the causal relevance of diverse risk markers is compared in an explicit framework helps clarify the relative standing of each risk marker. Field synopses, expanded to include the whole range of risk markers considered of potential clinical or public health relevance, should be developed to prioritise research efforts and to focus recommendations on those markers most likely to be causal.

This work was developed during meetings of the Nucleus of Epidemiology and Public Health, European Association of Cardiovascular Prevention and Rehabilitation, European Society of Cardiology in Paris 2007 and Sofia Antipolis 2008 (HH chair, members GC, JED, PH, XJ, SM, BMM, SS, TT, JCMW, ADH, and DAL).

Contributors: The writing group included HK, AN, MK, and HH. HK, AN, and MK contributed equally to the study. HK, HH, AN, and AAS carried out the searches and extracted the data. All authors took part in the discussion group where the new methodology was developed, commented on early drafts of the manuscript through to revisions, approved the final draft, had full access to all the data, and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: MK is supported by the Academy of Finland. HK is supported by a grant from the Wellcome Trust. DAL is supported by a UK Department of Health Career Scientist Award and works in a centre that receives support from the UK Medical Research Council. ADH is supported by a British Heart Foundation senior research fellowship (FS 05/125).

Competing interests: ADH is a member of the editorial board of Drug and Therapeutics Bulletin and has acted as an adviser to GlaxoSmithKline and London Genetics. He has received honorariums for speaking at educational meetings sponsored by the pharmaceutical industry and has donated all or most of these to charity.

Ethical approval: Not required.

Data sharing: No additional data are available.


Accepted: 12 July 2009