

Bias in Associations of Emerging Biomarkers With Cardiovascular Disease

Ioanna Tzoulaki, PhD; Konstantinos C. Siontis, MD; Evangelos Evangelou, PhD; John P. A. Ioannidis, MD, DSc

Importance: Numerous cardiovascular biomarkers are proposed as potential predictors of cardiovascular risk.

Objective: To evaluate whether there is evidence for biases favoring statistically significant results and inflating associations in this literature.

Design and Setting: PubMed search for meta-analyses of cardiovascular biomarkers that are not part of the Framingham Risk Score.

Main Outcome Measures: We estimated summary effects and between-study heterogeneity (considered “very large” for $I^2 > 75\%$). We evaluated whether large studies had significantly more conservative results than smaller studies (small-study effects) and whether there were too many studies with statistically significant results compared with what would be expected on the basis of the findings of the largest study in each meta-analysis.

Results: Of 56 eligible meta-analyses, 49 had statistically significant results. Very large heterogeneity and small-study effects were seen in 9 and 13 meta-analyses,

respectively. In 29 meta-analyses (52%), there was a significant excess of studies with statistically significant results. Only 13 of the statistically significant meta-analyses had more than 1000 cases and no hints of large heterogeneity, small-study effects, or excess significance. These included the associations of glomerular filtration rate and albumin to creatinine ratio in general and high-risk populations with cardiovascular disease mortality and of non-high-density lipoprotein cholesterol, serum albumin, *Chlamydia pneumoniae* IgG, glycosylated hemoglobin, nonfasting insulin, apolipoprotein B/AI ratio, erythrocyte sedimentation rate, and lipoprotein-associated phospholipase mass or activity with coronary heart disease.

Conclusions and Relevance: Selective reporting biases may be common in the evidence on emerging cardiovascular biomarkers. Most of the proposed associations of these biomarkers may be inflated.

JAMA Intern Med. 2013;173(8):664-671.

Published online March 25, 2013.

doi:10.1001/jamainternmed.2013.3018

Author Affiliations: Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece (Drs Tzoulaki, Siontis, and Evangelou); Department of Epidemiology and Biostatistics, Imperial College London, London, United Kingdom (Dr Tzoulaki); Mayo School of Graduate Medical Education, College of Medicine, Mayo Clinic, Rochester, Minnesota (Dr Siontis); and Stanford Prevention Research Center, Departments of Medicine and Health Research and Policy, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California (Dr Ioannidis).

NUMEROUS CARDIOVASCULAR biomarkers are proposed as potential predictors of cardiovascular risk.¹⁻³ Despite intensive research efforts, to date, few emerging cardiovascular biomarkers have shown clear improvements in predictive discrimination, reclassification, and/or calibration, and their clinical utility remains equivocal.⁴⁻⁶ Methodologic limitations in this field, including poor reporting, lack of validation, and unjustified claims of improved prediction over well-established risk scores, cast doubts on the validity and magnitude of reported effect sizes.^{4,7-10}

Numerous meta-analyses have been performed to date aiming to summarize the predictive value of individual cardiovascular biomarkers. However, to our knowledge, there has been no recent effort to summarize the evidence from these meta-analyses on their associated limitations such as publication, selective analysis, and outcome reporting biases. Here we assembled for the first time a comprehensive

systematic sample of meta-analyses that examine associations of emerging biomarkers with cardiovascular outcomes.

*See Invited Commentary
at end of article and
see also page 611*

We sought to evaluate whether there is evidence for biases in this literature favoring statistically significant and/or inflated results and how many and which of these cardiovascular biomarker effects have no hints of bias.

METHODS

IDENTIFICATION OF ELIGIBLE STUDIES

We assembled meta-analyses that examined any emerging biomarker, defined as any biological parameter¹¹ other than those included in the Framingham Risk Score,¹² in relation to cardiovascular disease (CVD), coronary heart dis-

ease (CHD), or cardiovascular mortality. Meta-analyses were selected as a tool to provide us with systematic summary evidence on each examined biomarker. All types of biomarkers were eligible (blood, urine, tissue, imaging, or physical measurement). We excluded meta-analyses of single common genetic variants because these markers have limited prognostic ability when examined in isolation, but multigene scores were eligible.

We used 3 different approaches to collect a comprehensive sample of biomarker meta-analyses indexed in MEDLINE (with no year restriction and last update as of January 30, 2012). First, we used the algorithm “(‘Biological marker’ [MeSH Terms]) AND (cardiovascular OR coronary [Title/Abstract])” limited to meta-analyses, English language, and human studies. Second, we performed targeted MEDLINE searches for meta-analyses of 71 additional specific emerging biomarkers included in recent comprehensive reviews¹³⁻¹⁵ using the same algorithm, but instead of applying the generic “Biological marker” term, we tracked the name of each biomarker (see eTable for full list of biomarkers searched; <http://www.jamainternalmed.com>). We perused the title and abstract of each of these citations, and potentially eligible articles were then retrieved in full text. Finally, we identified meta-analyses of individual participant data published by major consortia operating in the field (Emerging Risk Factor Collaboration, Fibrinogen Studies Collaboration, Ankle Brachial Index Collaboration, Homocysteine Studies Collaboration, and Chronic Kidney Disease Prognosis Consortium).

We included studies regardless of the baseline characteristics (clinical setting) of the examined populations. Whenever an article presented separate meta-analyses on more than 1 eligible biomarker, outcome, or type of clinical settings, those were kept separately. Meta-analyses were eligible regardless of whether the included studies used adjustment for some covariates or score (eg, the Framingham Risk Score) or unadjusted analyses. We excluded meta-analyses of randomized controlled trials assessing the change of a biomarker in relation to an outcome. When more than 1 meta-analysis examining the same biomarker and same outcome on the same clinical setting were identified, only the most recent one with eligible data was kept.

DATA EXTRACTION

Data extraction was performed independently by 2 investigators (K.C.S. and I.T.) and, in the case of discrepancies, consensus was reached. From each eligible meta-analysis, we recorded the first author, journal, year of publication, and number of studies in the meta-analysis, and we noted the biomarker, risk factors or score used for covariate adjustment and the outcome examined. The study-specific relative risk (RR) estimates (risk ratio, odds ratio, hazard ratio, or incident risk ratio, as reported by the meta-analysis authors) along with the corresponding 95% CIs and the number of cases in each study were extracted for each biomarker and outcome. The number of control participants in addition to the number of cases for each study were extracted when needed for the power calculation (see the Statistical Analysis subsection). When data on number of cases (or number of cases and controls when needed) were missing, meta-analyses were excluded and substituted with a previously published meta-analysis on the same biomarker, whenever available. Meta-analyses of cardiovascular outcomes typically assume that study-specific RRs (odds ratio, hazard ratio, and risk ratio) are similar and combine different RRs under this assumption. This assumption is a fair approximation whenever the disease incidence is low or modest. We, therefore, extracted the metric of the largest study (study with smallest variance) and assumed that the remaining RRs correspond to the same metric. Whenever data were provided with differ-

ent adjustments, we preferred estimates that adjusted for the Framingham Risk Score or the model with Framingham Risk Score variables; if neither of these options was available, we preferred the model with the larger number of adjusting factors. When subgroups were presented, we extracted the data for each subgroup separately unless the study combined the results across all subgroups.

STATISTICAL ANALYSIS

For each meta-analysis, we estimated the summary effect size and its 95% CI using random effects models¹⁶ and calculated the I^2 metric for heterogeneity. The I^2 metric ranges from 0% to 100% and is the ratio of between-study variance to the sum of the within- and between-study variances. Values exceeding 50% or 75% are considered to represent “large” or “very large” heterogeneity, respectively. The 95% CIs of I^2 estimates can be wide when there are few studies.¹⁷ Furthermore, we used the regression asymmetry test proposed by Egger et al¹⁸ and examined by Sterne et al.¹⁹ $P < .10$ with more conservative effect in larger studies was considered evidence for small-study effects—that is, that the results of small studies differed from those of larger studies. Various biases or genuine heterogeneity may cause small-study effects.¹⁹

We applied the excess significance test to evaluate whether the observed number of studies (O) with statistically significant results (“positive” studies) in each meta-analysis is larger than their expected number (E).²⁰⁻²² We also summed the O and E across all meta-analyses. For each meta-analysis, E is the sum of the power estimates of the studies it includes. The estimated power depends on the plausible effect size. The true effect size for any meta-analysis is unknown. Herein, we assumed that the most plausible effect is given by the largest study. Different equations were used to estimate the power when the largest study reported a hazard ratio²³ or odds ratio.²⁴ Excess significance for single meta-analyses was claimed at $P < .10$ (1-sided $P < .05$ with $O > E$ as previously proposed²²).

Predefined subgroup analyses applied the excess significance test in subgroups of meta-analyses with or without large between-study heterogeneity, small-study effects, individual-level data, or nominally significant summary effects and by primary (general populations) vs secondary (high risk or populations with CVD) prevention and biomarker (biological fluid measurement or physical measurement/ imaging).

Stata, version 10.1 (StataCorp), was used for statistical analyses. P values were 2-tailed.

RESULTS

ELIGIBLE STUDIES AND META-ANALYSES

Overall, 582 articles were searched and 35 articles corresponding to 56 meta-analyses were deemed eligible²⁵⁻⁵⁹ (eFigure). Examining 42 unique biomarkers for the 56 meta-analyses, the median (range) number of studies was 12 (3-68) and number of events was 2459 (34-12 785). Meta-analyses pertained to a range of biomarkers and populations examined; most data corresponded to primary prevention (general populations, 42 meta-analyses) and biomarkers measured in biological fluids (37 meta-analyses). The outcome examined was CHD in 28 meta-analyses, CVD in 21, and CVD mortality or cardiac death in 7 (**Table 1**). Fourteen meta-analyses were of individual participants and 42 analyzed published literature. All but 4 meta-analyses reported RR estimates

Table 1. Description of 56 Eligible Meta-analyses

Source	Biomarker	Outcome	Comparison	Events	Events (Largest Study)	Summary RR (CI)	I ²	RR (Largest Study) (CI)
Primary Prevention (General Populations)								
25	Abdominal obesity	CVD	Per 1 SD higher	5209	826	1.32 (1.24-1.40)	65	1.33 (1.25-1.42)
32	ACR	CHD	10 vs 5 mg/g	3823	2119	1.27 (1.04-1.29) ^a	15	1.16 (1.04-1.29)
48	Adiponectin	CHD	Top vs bottom third	1313	589	0.86 (0.67-1.09)	30	0.89 (0.67-1.18)
36	Aortic PWV	CVD	High vs low	1681	616	1.14 (1.09-1.20)	82	1.03 (1.01-1.06) ^b
33	Apo(a) isoform	CHD	Smaller vs larger	7382	1013	2.08 (1.67-2.60)	85	1.58 (1.33-1.93)
49	Apolipoprotein AI	CHD	Bottom vs top third	6320	899	1.61 (1.42-1.82)	51	1.16 (0.93-1.47)
49	Apolipoprotein B	CHD	Top vs bottom third	6333	899	1.98 (1.65-2.38)	80	1.32 (1.09-1.60) ^b
49	Apolipoprotein B/AI	CHD	Top vs bottom third	3730	1743	1.85 (1.54-2.21)	75	1.84 (1.59-2.13)
37	BNP	CVD	Top vs bottom third	10 625	4618	2.82 (2.40-3.33)	83	1.77 (1.62-1.94) ^b
42	BNP/NT-proBNP	MACE	Positive vs negative	862	96	19.75 (13.18-29.6) ^c	30	26.52 (15.23-46.17) ^c
25	Body mass index	CVD	Per 1 SD higher	5209	826	1.29 (1.22-1.37)	66	1.24 (1.17-1.32)
35	C-reactive protein	CHD	Per 1 SD higher	5373	633	1.38 (1.27-1.49)	26	1.27 (1.11-1.44)
43	CAC	CVD	CAC>0 vs CAC = 0	1903	510	6.28 (4.63-8.53) ^c	59	5.78 (4.25-7.88) ^{b,c}
44	Cardiorespiratory fitness	CVD	Per 1 MET higher	4419	1512	0.85 (0.82-0.88)	73	0.91 (0.89-0.94) ^b
56	Chlamydia IgG	CHD	Top vs bottom third	3169	647	1.16 (1.02-1.31)	0	1.28 (0.88-1.84)
28	Daytime BP	CVD	Per 10 mm Hg higher	1128	317	1.21 (1.16-1.26)	0	1.14 (1.05-1.24) ^b
32	eGFR	CHD	15 vs 95 mL/min/1.73m ²	3823	562	2.66 (2.04-3.46) ^a	0	2.63 (1.50-4.62)
58	ESR	CHD	Top vs bottom third	1703	1462	1.53 (1.01-2.33)	0	1.25 (1.04-1.52)
45	Fasting insulin	CHD	Top vs bottom third	1649	677	1.20 (0.97-1.50)	72	1.03 (0.79-1.37)
51	Fibrinogen	CHD	Per 1 g/L higher	7118	1474	2.64 (2.27-3.07)	76	1.58 (1.54-2.09) ^b
26	HbA1c	CHD	Per 1 SD higher	1639	529	1.19 (1.10-1.30)	39	1.24 (1.17-1.32)
26	HbA1c	CVD	Top vs bottom quartile	2142	18	2.14 (1.47-3.10)	98	3.18 (3.12-3.25)
54	Homocysteine	CHD	Per 25% lower	5073	527	0.72 (0.66-0.78)	73	0.81 (0.71-0.92)
46	IMT	MI	Per 1 SD higher	1428	626	1.11 (1.08-1.14)	28	1.07 (1.03-1.11) ^b
57	Leukocyte count	CHD	Top vs bottom third	7229	1835	1.45 (1.34-1.57)	36	1.35 (1.15-1.60) ^b
39	Lipoprotein(a)	CHD	Per 1 SD higher	9321	1842	1.13 (1.09-1.18)	49	1.14 (1.09-1.20)
34	Lp-PLA2 activity	CHD	Per 1 SD higher	3963	1547	1.10 (1.05-1.16)	21	1.12 (1.06-1.19)
34	Lp-PLA2 mass	CHD	Per 1 SD higher	3963	1547	1.11 (1.07-1.16)	28	1.13 (1.07-1.19)
28	Night-to-day BP ratio	CVD	Per 10 mm Hg higher	1128	317	1.14 (1.03-1.27)	40	1.07 (0.94-1.22)
28	Nighttime BP	CVD	Per 10 mm Hg higher	1128	317	1.24 (1.15-1.33)	61	1.11 (1.04-1.18)
45	Nonfasting insulin	CHD	Top vs bottom third	1980	621	1.44 (1.06-1.94)	59	1.26 (0.93-1.68)
38	Non-HDL	CHD	Per 1 SD higher	12 785	871	1.50 (1.39-1.61)	73	1.33 (1.20-1.48)
28	Nondipping BP	CVD	Yes vs no	1128	317	1.22 (0.94-1.57)	59	1.56 (1.21-2.02)
31	Postload glucose	CHD	Per 1 SD higher	12 652	3723	1.03 (1.02-1.04)	67	1.02 (0.88-1.15)
45	Pro-insulin	CHD	Top vs bottom third	423	219	2.22 (1.66-2.98)	0	2.24 (1.60-3.09)
40	RAC	CHD	Per 20 μm lower	2219	565	1.07 (0.98-1.17)	52	1.09 (0.96-1.23)
40	RVC	CHD	Per 20 μm higher	2219	565	1.07 (0.98-1.16)	42	1.01 (0.89-1.14)
47	Selenium	CHD	Highest to lowest	3586	301	0.58 (0.47-0.72)	83	0.77 (0.59-0.99) ^b
57	Serum albumin	CHD	Bottom vs top third	3770	1865	1.55 (1.38-1.75)	0	1.50 (1.22-1.86)
52	Serum uric acid	CHD	Top vs bottom third	9458	1869	1.19 (1.05-1.34)	96	1.04 (0.90-1.20)
38	Triglycerides	CHD	Per 1 SD higher	12 785	871	0.99 (0.94-1.05)	35	1.00 (0.89-1.12)
25	Waist/hip ratio	CVD	Per 1 SD higher	5209	562	1.30 (1.23-1.38)	66	1.14 (1.08-1.21)

(continued)

adjusted for a variety of other cardiovascular risk factors (Table 1).

SUMMARY EFFECTS

Overall, 49 (88%) of the eligible meta-analyses reported a nominally statistically significant summary result (Table 1). The largest study had statistically significant results in 41 meta-analyses. The **Figure** shows the estimates of the largest studies against the random effects meta-analysis estimates. The largest study's result was more conservative compared with the summary result in 44 meta-analyses (79%), and most of the largest studies suggested effects of small magnitude (Table 1). Only myo-

cardial metabolic imaging had the reported RR estimate adjusted for cardiovascular risk factors exceeding 3.00 both in the meta-analysis and in the largest study. Brain natriuretic peptide in high-risk populations and coronary artery calcium also reported large effect sizes (RR > 3) in unadjusted analyses. Nonetheless, studies with adjusted analyses have also shown relatively high estimates for these biomarkers.^{42,43}

HETEROGENEITY AND SMALL-STUDY EFFECTS

Twenty-six meta-analyses (46%) had large heterogeneity (I² > 50%) and 9 (16%) had very large heterogeneity (I² > 75%). Evidence for significant small-study ef-

Table 1. Description of 56 Eligible Meta-analyses (continued)

Source	Biomarker	Outcome	Comparison	Events	Events (Largest Study)	Summary RR (CI)	I ²	RR (Largest Study) (CI)
Secondary Prevention (High-Risk Populations, Patients With CVD)								
27	ACR	CVD M	30 mg/g vs lower	2983	1821	1.56 (1.31-1.85) ^a	40	1.43 (1.23-1.65)
59	C-reactive protein	Cardiac M	High vs normal	34	4	4.15 (1.18-14.52) ^c	0	1.04 (0.10-10.35) ^c
29	Cystatin C	CVD	Highest to lowest	2698	113	2.70 (1.97-3.69)	83	1.52 (1.13-2.04) ^b
28	Daytime BP	CVD	Per 10 mm Hg higher	1259	356	1.14 (1.04-1.25)	63	1.08 (0.97-1.19)
27	eGFR	CVD M	45 mL/min/1.73m ² vs lower	2983	1821	1.73 (1.49-2.00) ^a	0	1.66 (1.37-1.99)
30	Masked HT	CVD	Yes vs no	202	156	2.09 (1.55-2.81)	0	1.66 (1.06-2.61)
41	MMI	CVD	Abnormal vs normal	85	13	3.74 (2.09-6.69)	17	4.55 (1.34-15.51)
28	Night-to-day BP ratio	CVD	Per 10 mm Hg higher	551	356	1.24 (1.12-1.36)	0	1.23 (1.08-1.40)
28	Nighttime BP	CVD	Per 10 mm Hg higher	1259	356	1.20 (1.14-1.26)	25	1.20 (1.10-1.31)
28	Nondipping BP	CVD	Yes vs no	631	356	1.40 (1.18-1.66)	0	1.36 (1.09-1.69)
55	Troponin	Cardiac M	Positive vs negative	1330	1290	5.24 (3.56-7.72) ^c	68	2.82 (2.30-3.45) ^{b,c}
50	Troponin T	Cardiac M	>99th percentile	145	26	2.55 (1.93-3.37)	0	1.98 (1.25-3.15) ^b
53	vWF	CHD	Top vs bottom third	723	186	1.64 (1.21-2.22)	0	1.45 (0.86-2.43)
30	White-coat HT	CVD	Yes vs no	162	156	0.96 (0.65-1.42)	0	0.96 (0.50-1.84)

Abbreviations: ACR, albumin to creatinine ratio; Apo(a), apolipoprotein (a); BNP, brain natriuretic peptide; BP, blood pressure; CAC, coronary artery calcium; CHD, coronary heart disease; CVD, cardiovascular disease; eGFR, estimated glomerular filtration rate; ESR, erythrocyte sedimentation rate; HbA1c, glycosylated hemoglobin; HDL, high-density lipoprotein; HT, hypertension; IMT, intima-media thickness; Lp-PLA2, lipoprotein-associated phospholipase; M, mortality; MACE, major adverse cardiac event; MET, metabolic equivalent; MI, myocardial infarction; MMI, myocardial metabolic imaging; NT-proBNP, N-terminal fragment of the prohormone brain natriuretic peptide; PWV, pulse wave velocity; RAC, retinal arteriolar caliber; RR, relative risk; RVC, retinal venular caliber; vWF, von Willebrand factor.

^aMeta-analysis reporting linear splines and RR of only 1 knot point is shown here.

^bMeta-analysis with significant small-study effects by the Egger test ($P < .10$).

^cReported RR is unadjusted for other cardiovascular risk factors.

fects was noted in 13 meta-analyses (23%). Meta-analyses of fibrinogen, selenium, and apolipoprotein B with CHD and brain natriuretic peptide, aortic pulse wave velocity, and cystatin C with CVD showed very large heterogeneity and evidence of small-study effects. The meta-analyses of coronary artery calcium and cardiorespiratory fitness with CVD and troponin with 30-day cardiac death in patients with acute coronary syndrome showed large heterogeneity ($I^2 = 50\%-75\%$) and evidence of small-study effects (Table 1).

EXCESS SIGNIFICANCE

In 29 meta-analyses (52%) there was a significant excess of observed “positive” studies compared with those expected (Table 2). Table 3 shows aggregate data from all the meta-analyses and according to different subgroups. Among 919 studies included in 56 meta-analyses, 472 (51%) had nominally statistically significant results, while the expected number was 317. The difference between the observed and expected was significant ($P < .001$). The excess of significant findings was documented across all the examined subgroups (Table 3).

SIGNIFICANT BIOMARKER ASSOCIATIONS WITHOUT HINTS OF BIAS

Of the 56 meta-analyses, 18 (32%) had nominally statistically significant summary associations per random effects calculations and had no evidence of small-study effects ($P \geq .10$), not very large heterogeneity ($I^2 \leq 75\%$), and no evidence for excess significance.

Overall, 13 of the 18 associations (72%) were based on cumulative evidence of more than 1000 cardiovascular events. This included 9 biomarkers with associations with CHD in general populations (non-high-density lipoprotein cholesterol, serum albumin, *Chlamydia pneumoniae* IgG titers, apolipoprotein B/AI ratio, glycosylated hemoglobin, lipoprotein-associated phospholipase mass and activity, erythrocyte sedimentation rate, and nonfasting insulin) and 2 biomarkers (estimated glomerular filtration rate [eGFR] and albumin to creatinine ratio) with associations with CVD mortality in high-risk and general populations. Across these 13 associations, the RR per 1-SD increase (assuming the top vs bottom tertile comparison corresponds to approximately 2 SD) had a median (range) of 1.2 (1.1-1.5), excluding the eGFR and albumin to creatinine meta-analyses. The latter meta-analyses in general and in high-risk populations reported RR using linear splines with knots at different levels of biomarkers. The RR was greater than 2 only in extreme comparisons—for example, the meta-analysis between eGFR and CVD mortality in general populations reported an RR of 2.66 comparing eGFR of 15 mL/min/1.73 m² (0.4% of the population) vs the reference category (95 mL/min/1.73 m²); the RR in less extreme categories was more conservative (RR = 1.40 for 60 vs 95 mL/min/1.73 m²).

COMMENT

A systematic evaluation of 56 meta-analyses of emerging cardiovascular biomarkers suggests that many results are prone to biases. We found strong evidence to

suggest the effect of biomarkers is exaggerated because the largest studies—which one would expect to produce the most stable estimates—consistently showed smaller effects. In most meta-analyses, too many single studies had reported “positive” results compared with what would be expected on the basis of the results of the

largest studies. This suggests that small studies with “negative” results remain unpublished or that their results are distorted during analysis and reporting to seem more prominent.

Bias from preferential reporting of positive findings or “positive analyses” has been postulated as a major problem in clinical investigation and biomarker research in particular.⁶⁰⁻⁶² Perhaps only a fraction of potentially available data are eventually published in this field. For example, the number of studies in the meta-analyses of triglycerides, body mass index, C-reactive protein, homocysteine, and apolipoprotein B, which all examined general populations and CHD, varies extensively: 68, 39, 31, 26, and 21 studies, respectively. These biomarkers are relatively inexpensive and easy to measure routinely; thus, they should be available in most epidemiologic data sets with cardiovascular outcomes. Differences in the number of published studies may result in part from prioritization for publication of positive results. In addition, some promising biomarker results may simply reflect false-positives owing to multiple testing of many biomarkers.^{63,64} Selective analysis reporting bias emerges when there are many analyses that can be performed and only some of them, the ones with the “best” results, are presented.^{65,66} Scientists may have examined a wide range of biomarkers in the same population, dif-

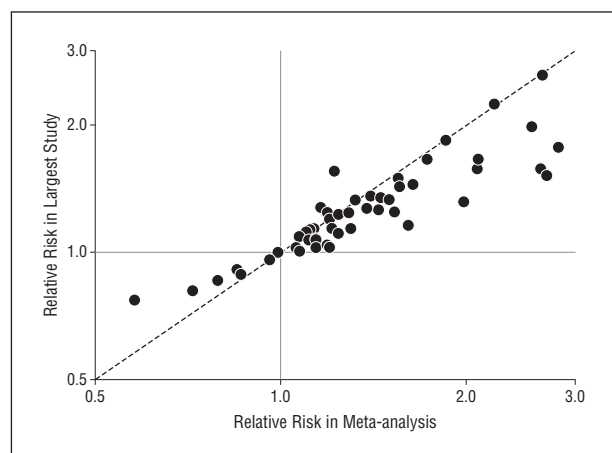


Figure. Correlation between the summary relative risk in each meta-analysis (random effects) and the relative risk in the largest study. Five outliers with relative risk greater than 3 in the meta-analysis and/or the largest study are not shown. Axes are in logarithmic scale.

Table 2. Observed and Expected Number of Significant Studies in the 29 Meta-analyses With a Significant Excess of Significant Studies

Source	Biomarker	Studies	Observed Significant	Expected Significant	Observed to Expected	P Value ^a
38	Triglycerides	68	5	1.7	2.94	.01
51	Fibrinogen	31	27	21.9	1.23	.04
59	C-reactive protein	4	1	0.2	5.00	.07
54	Homocysteine	26	17	6.2	2.74	<.001
52	Serum uric acid	23	5	1.6	3.13	.005
47	Selenium	30	12	3.8	3.16	<.001
35	C-reactive protein	31	16	10.1	1.58	.02
39	Lipoprotein(a)	30	10	5.6	1.79	.04
49	Apolipoprotein B	21	15	10.1	1.49	.02
49	Apolipoprotein AI	19	11	4.8	2.29	.001
31	Postload glucose	15	6	0.8	7.50	<.001
37	BNP	40	37	30.3	1.22	.01
50	Troponin T	8	6	2.3	2.61	.003
45	Fasting insulin	14	4	0.5	8.00	<.001
29	Cystatin C	11	10	7.0	1.43	.06
25	Body mass index	39	17	8.6	1.98	.001
25	Abdominal obesity	39	21	12.2	1.72	.003
25	Waist/hip ratio	39	16	4.4	3.64	<.001
28	Nighttime BP ^b	8	5	1.7	2.94	.003
28	Daytime BP ^b	8	2	0.6	3.33	.04
28	Night-to-day BP ratio ^b	3	3	0.8	3.75	.006
28	Nighttime BP ^c	8	6	0.7	8.57	<.001
28	Daytime BP ^c	8	7	0.9	7.78	<.001
28	Night-to-day BP ratio ^c	8	2	0.5	4.00	.02
40	RVC	12	2	0.3	6.67	.004
30	Masked HT	6	4	0.8	5.00	.001
46	IMT	5	5	0.4	12.50	<.001
36	Aortic PWV	15	10	0.4	25.00	<.001
44	Cardiorespiratory fitness	24	21	2.2	9.55	<.001

Abbreviations are explained in the first footnote to Table 1.

^aP value of the excess significance test.

^bGeneral populations.

^cPatients with cardiovascular disease.

Table 3. Observed and Expected Number of Significant Studies Across All Meta-analyses and in Subgroups

Variable	Studies	Meta-analyses	Expected Significant	Observed Significant	Observed to Expected	P Value ^a
All meta-analyses	919	56	317	472	1.49	<.001
Heterogeneity						
I ² < 50%	347	30	106	136	1.28	<.001
I ² > 50%	572	26	211	336	1.59	<.001
Type of meta-analysis						
MPL	522	42	197	300	1.52	<.001
MIPD	397	14	120	172	1.43	<.001
Statistically significant summary result						
Yes	792	49	308	455	1.48	<.001
No	127	7	9	17	1.89	.007
Type of biomarker						
Biological fluid measurement	661	37	269	343	1.28	<.001
Physical measurement imaging	258	19	48	129	2.69	<.001
Type of population						
Primary prevention	816	42	274	416	1.52	<.001
Secondary prevention	103	14	43	56	1.30	.007
Small-study effects						
Yes	238	13	115	192	1.67	<.001
No	681	43	202	280	1.39	<.001

Abbreviations: MIPD, meta-analysis of individual participant data; MPL, meta-analysis of published literature.

^aP value of the excess significance test.

ferent outcomes, different cut-off points, and different covariate adjustments for a single biomarker but then report only 1 or a few of these analyses.

In our evaluation, the excess of significant findings was seen across all prespecified subgroups. A cluster of meta-analyses, including those of fibrinogen, selenium, and apolipoprotein B with CHD and brain natriuretic peptide and aortic pulse wave velocity with CVD, had very large heterogeneity and small-study effects. Of note, the Egger test is particularly difficult to interpret in the presence of prominent between-study heterogeneity.¹⁹ Genuine heterogeneity of the strength of association in diverse settings and populations may be confused for selective reporting bias. However, in all of these cases there was also a marked excess of studies with positive results. Therefore, the overall picture is more consistent with bias and suggests that the claimed effect sizes are inflated. The meta-analyses themselves sometimes addressed these issues on individual biomarkers, calculated heterogeneity, assessed presence of publication bias, and presented subgroup and meta-regression analyses to identify differences between studies.

Hints of bias cannot exclude that these biomarkers have any association with cardiovascular outcomes. It is difficult to differentiate whether the underlying effect is small or null and whether genuine heterogeneity exists. There are also several associations in the literature of cardiovascular biomarkers that did not show any evidence of biases. These included the association of eGFR and albumin to creatinine ratio with CVD mortality in general and high-risk populations as well as the association of non-high-density lipoprotein cholesterol, serum albumin, *Chlamydia pneumoniae* IgG, glycosylated hemoglobin, nonfasting insulin, erythrocyte sedimentation rate, apolipoprotein B/AI ratio, and lipoprotein-associated phospholipase with CHD. Most of these associations were weak in terms of the magnitude of the RR. Therefore, their abil-

ity to improve cardiovascular risk assessment might be limited when used in isolation.

Several caveats should be considered in interpreting our findings. First, both asymmetry and excess significance tests offer hints of bias, not proof thereof. The exact estimation of excess significance is influenced by the choice of plausible effect size and the calculation of power. We selected the largest study's effect as the plausible effect. These studies may have inherent biases themselves. Finally, we performed power calculations assuming that all studies within each meta-analysis reported the same RR metric. This is a common assumption in meta-analyses of CVD outcomes, which typically consider the different reported metrics to be similar assuming CVD incidence is less than 10%.

With acknowledgment of these caveats, our evaluation maps the status of the evidence on the associations between popular cardiovascular biomarkers with CVD outcomes. Similar conclusions have been drawn from investigations in other fields, such as Alzheimer disease and mental health conditions.^{21,22} Cardiovascular biomarkers may be as prone to bias as biomarkers in other fields. Single biomarkers with large effect sizes are probably rare. Thus, the much-awaited improvement in CVD risk prediction will require evaluation of composites of numerous biomarkers in large-scale consortia using standardized approaches to minimize biases.⁴

Accepted for Publication: November 5, 2012.

Published Online: March 25, 2013. doi:10.1001/jamainternmed.2013.3018

Correspondence: John P. A. Ioannidis, MD, DSc, Stanford Prevention Research Center, 1265 Welch Rd, Medical School Office Building, Room X306, Stanford, CA 94350 (jioannid@stanford.edu).

Author Contributions: Dr Ioannidis had full access to all the data in the study and takes responsibility for the

integrity of the data and the accuracy of the data analysis. *Study concept and design*: Tzoulaki, Siontis, and Ioannidis. *Acquisition of data*: Tzoulaki, Siontis, and Evangelou. *Analysis and interpretation of data*: All authors. *Drafting of the manuscript*: Tzoulaki, Siontis, and Ioannidis. *Critical revision of the manuscript for important intellectual content*: Siontis and Evangelou. *Statistical analysis*: All authors. *Administrative, technical, and material support*: Siontis. *Study supervision*: Ioannidis. **Conflict of Interest Disclosures**: None reported. **Online-Only Material**: The eTable and eFigure are available at <http://www.jamainternalmed.com>.

REFERENCES

- Hlatky MA, Greenland P, Arnett DK, et al; American Heart Association Expert Panel on Subclinical Atherosclerotic Diseases and Emerging Risk Factors and the Stroke Council. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association [published correction appears in *Circulation*. 2009;119(25):e606]. *Circulation*. 2009;119(17):2408-2416.
- Wang TJ. Assessing the role of circulating, genetic, and imaging biomarkers in cardiovascular risk prediction. *Circulation*. 2011;123(5):551-565.
- Gilstrap LG, Wang TJ. Biomarkers and cardiovascular risk assessment for primary prevention: an update. *Clin Chem*. 2012;58(1):72-82.
- Ioannidis JP, Tzoulaki I. Minimal and null predictive effects for the most popular blood biomarkers of cardiovascular disease. *Circ Res*. 2012;110(5):658-662.
- Wilson PW. Challenges to improve coronary heart disease risk assessment. *JAMA*. 2009;302(21):2369-2370.
- Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*. 2010;121(15):1768-1777.
- Tzoulaki I, Liberopoulos G, Ioannidis JP. Use of reclassification for assessment of improved prediction: an empirical evaluation. *Int J Epidemiol*. 2011;40(4):1094-1105.
- Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction beyond the Framingham Risk Score. *JAMA*. 2009;302(21):2345-2352.
- Ioannidis JP, Tzoulaki I. What makes a good predictor? the evidence applied to coronary artery calcium score. *JAMA*. 2010;303(16):1646-1647.
- Tzoulaki I, Siontis KC, Ioannidis JP. Prognostic effect size of cardiovascular biomarkers in datasets from observational studies versus randomised trials: meta-epidemiology study. *BMJ*. 2011;343:d6829. doi:10.1136/bmj.d6829.
- Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. *Circulation*. 2006;113(19):2335-2362.
- Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97(18):1837-1847.
- Greenland P, Alpert JS, Beller GA, et al; American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2010;122(25):e584-e636.
- Apple FS, Wu AH, Mair J, et al; Committee on Standardization of Markers of Cardiac Damage of the IFCC. Future biomarkers for detection of ischemia and risk stratification in acute coronary syndrome. *Clin Chem*. 2005;51(5):810-824.
- Naghavi M, Libby P, Falk E, et al. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: part I. *Circulation*. 2003;108(14):1664-1672.
- Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med*. 1997;127(9):820-826.
- Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*. 2007;335(7626):914-916.
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634.
- Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002. doi:10.1136/bmj.d4002.
- Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clin Trials*. 2007;4(3):245-253.
- Kavvoura FK, McQueen MB, Khoury MJ, Tanzi RE, Bertram L, Ioannidis JP. Evaluation of the potential excess of statistically significant findings in published genetic association studies: application to Alzheimer's disease. *Am J Epidemiol*. 2008;168(8):855-865.
- Ioannidis JP. Excess significance bias in the literature on brain volume abnormalities. *Arch Gen Psychiatry*. 2011;68(8):773-780.
- Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. *Stat Med*. 1982;1(2):121-129.
- Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med*. 2000;19(22):3127-3131.
- Wormser D, Kaptoge S, Di Angelantonio E, et al; Emerging Risk Factors Collaboration. Separate and combined associations of body mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet*. 2011;377(9771):1085-1095.
- Santos-Oliveira R, Purdy C, da Silva MP, dos Anjos Carneiro-Leão AM, Machado M, Einarson TR. Haemoglobin A1c levels and subsequent cardiovascular disease in persons without diabetes: a meta-analysis of prospective cohorts. *Diabetologia*. 2011;54(6):1327-1334.
- van der Velde M, Matsushita K, Coresh J, et al; Chronic Kidney Disease Prognosis Consortium. Lower estimated glomerular filtration rate and higher albuminuria are associated with all-cause and cardiovascular mortality: a collaborative meta-analysis of high-risk population cohorts. *Kidney Int*. 2011;79(12):1341-1352.
- Hansen TW, Li Y, Boggia J, Thijs L, Richart T, Staessen JA. Predictive role of the nighttime blood pressure. *Hypertension*. 2011;57(1):3-10.
- Lee M, Saver JL, Huang WH, Chow J, Chang KH, Ovbiagele B. Impact of elevated cystatin C level on cardiovascular disease risk in predominantly high-cardiovascular-risk populations: a meta-analysis. *Circ Cardiovasc Qual Outcomes*. 2010;3(6):675-683.
- Pierdomenico SD, Cuccurullo F. Prognostic value of white-coat and masked hypertension diagnosed by ambulatory monitoring in initially untreated subjects: an updated meta analysis. *Am J Hypertens*. 2011;24(1):52-58.
- Sarwar N, Aspelund T, Eiriksdottir G, et al. Markers of dysglycaemia and risk of coronary heart disease in people without diabetes: Reykjavik prospective study and systematic review. *PLoS Med*. 2010;7(5):e1000278. doi:10.1371/journal.pmed.1000278.
- Matsushita K, van der Velde M, Astor BC, et al; Chronic Kidney Disease Prognosis Consortium. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet*. 2010;375(9731):2073-2081.
- Erqou S, Thompson A, Di Angelantonio E, et al. Apolipoprotein(a) isoforms and the risk of vascular disease: systematic review of 40 studies involving 58,000 participants. *J Am Coll Cardiol*. 2010;55(19):2160-2167.
- Thompson A, Gao P, Orfei L, et al; Lp-PLA(2) Studies Collaboration. Lipoprotein-associated phospholipase A₂ and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. *Lancet*. 2010;375(9725):1536-1544.
- Kaptoge S, Di Angelantonio E, Lowe G, et al; Emerging Risk Factors Collaboration. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet*. 2010;375(9709):132-140.
- Vlachopoulos C, Aznaouridis K, Stefanadis C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-analysis. *J Am Coll Cardiol*. 2010;55(13):1318-1327.
- Di Angelantonio E, Chowdhury R, Sarwar N, et al. B-type natriuretic peptides and cardiovascular risk: systematic review and meta-analysis of 40 prospective studies. *Circulation*. 2009;120(22):2177-2187.
- Di Angelantonio E, Sarwar N, Perry P, et al; Emerging Risk Factors Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA*. 2009;302(18):1993-2000.
- Erqou S, Kaptoge S, Perry PL, et al; Emerging Risk Factors Collaboration. Lipoprotein (a) concentration and the risk of coronary heart disease, stroke, and nonvascular mortality. *JAMA*. 2009;302(4):412-423.
- McGeechan K, Liew G, Macaskill P, et al. Meta-analysis: retinal vessel caliber and risk for coronary heart disease. *Ann Intern Med*. 2009;151(6):404-413.
- Inaba Y, Bergmann SR. Prognostic value of myocardial metabolic imaging with BMIPP in the spectrum of coronary artery disease: a systematic review. *J Nucl Cardiol*. 2010;17(1):61-70.
- Ryding AD, Kumar S, Worthington AM, Burgess D. Prognostic value of brain natriuretic peptide in noncardiac surgery: a meta-analysis. *Anesthesiology*. 2009;111(2):311-319.
- Sarwar A, Shaw LJ, Shapiro MD, et al. Diagnostic and prognostic value of absence of coronary artery calcification [published correction appears in *JACC Cardiovasc Imaging*. 2010;3(10):1089]. *JACC Cardiovasc Imaging*. 2009;2(6):675-688.
- Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. *JAMA*. 2009;301(19):2024-2035.
- Sarwar N, Sattar N, Gudnason V, Danesh J. Circulating concentrations of insulin markers and coronary heart disease: a quantitative review of 19 Western prospective studies. *Eur Heart J*. 2007;28(20):2491-2497.

46. Lorenz MW, Markus HS, Bots ML, Rosvall M, Sitzer M. Prediction of clinical cardiovascular events with carotid intima-media thickness: a systematic review and meta-analysis. *Circulation*. 2007;115(4):459-467.
47. Flores-Mateo G, Navas-Acien A, Pastor-Barriuso R, Guallar E. Selenium and coronary heart disease: a meta-analysis. *Am J Clin Nutr*. 2006;84(4):762-773.
48. Sattar N, Wannamethee G, Sarwar N, et al. Adiponectin and coronary heart disease: a prospective study and meta-analysis. *Circulation*. 2006;114(7):623-629.
49. Thompson A, Danesh J. Associations between apolipoprotein B, apolipoprotein AI, the apolipoprotein B/AI ratio, and coronary heart disease: a literature-based meta-analysis of prospective studies. *J Intern Med*. 2006;259(5):481-492.
50. Khan NA, Hemmelgarn BR, Tonelli M, Thompson CR, Levin A. Prognostic value of troponin T and I among asymptomatic patients with end-stage renal disease: a meta-analysis. *Circulation*. 2005;112(20):3088-3096.
51. Danesh J, Lewington S, Thompson SG, et al; Fibrinogen Studies Collaboration. Plasma fibrinogen level and the risk of major cardiovascular diseases and non-vascular mortality: an individual participant meta-analysis. *JAMA*. 2005;294(14):1799-1809.
52. Wheeler JG, Juzwishin KD, Eiriksdottir G, Gudnason V, Danesh J. Serum uric acid and coronary heart disease in 9,458 incident cases and 155,084 controls: prospective study and meta-analysis. *PLoS Med*. 2005;2(3):e76. doi:10.1371/journal.pmed.0020076.
53. Whincup PH, Danesh J, Walker M, et al. von Willebrand factor and coronary heart disease: prospective study and meta-analysis. *Eur Heart J*. 2002;23(22):1764-1770.
54. Homocysteine Studies Collaboration. Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. *JAMA*. 2002;288(16):2015-2022.
55. Ottani F, Galvani M, Nicolini FA, et al. Elevated cardiac troponin levels predict the risk of adverse outcome in patients with acute coronary syndromes. *Am Heart J*. 2000;140(6):917-927.
56. Danesh J, Whincup P, Walker M, et al. *Chlamydia pneumoniae* IgG titres and coronary heart disease: prospective study and meta-analysis. *BMJ*. 2000;321(7255):208-213.
57. Danesh J, Whincup P, Walker M, et al. Low-grade inflammation and coronary heart disease: prospective study and updated meta-analyses. *BMJ*. 2000;321(7255):199-204.
58. Danesh J, Collins R, Peto R, Lowe GD. Haematocrit, viscosity, erythrocyte sedimentation rate: meta-analyses of prospective studies of coronary heart disease. *Eur Heart J*. 2000;21(7):515-520.
59. Padayachee L, Rodseth RN, Biccard BM. A meta-analysis of the utility of C-reactive protein in predicting early, intermediate-term, and long-term mortality and major adverse cardiac events in vascular surgical patients. *Anaesthesia*. 2009;64(4):416-424.
60. Ioannidis JP, Panagiotou OA. Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA*. 2011;305(21):2200-2210.
61. Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One*. 2008;3(8):e3081. doi:10.1371/journal.pone.0003081.
62. Andre F, McShane LM, Michiels S, et al. Biomarker studies: a call for a comprehensive biomarker study registry. *Nat Rev Clin Oncol*. 2011;8(3):171-176.
63. Ioannidis JP, Tarone R, McLaughlin JK. The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*. 2011;22(4):450-456.
64. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124. doi:10.1371/journal.pmed.0020124.
65. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ*. 2005;330(7494):753. doi:10.1136/bmj.38356.424606.8F.
66. Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457-2465.

INVITED COMMENTARY

Biomarkers in Cardiovascular Medicine

The Shame of Publication Bias

The term *evidence-based medicine* has been defined as “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.”^{1(p71)} Few other concepts in contemporary medical practice have achieved such universal adulation among practitioners, academicians, and payers. As a profession, we have placed evidence-based medicine on a golden pedestal as the ultimate expression of our desire to make clinical decisions in a systematic and scientific fashion. In an ideal world, evidence-based medicine drives the content of clinical guidelines and informs decisions by payers about what tests or procedures should be performed and reimbursed. In developing guidelines, authors typically integrate the findings from all available published studies to provide the best possible advice to clinicians. Accordingly, the quality of guidelines is only as good as the published studies on which they are based.

Unfortunately, there is a dark secret that corrupts nearly every aspect of our profession and undermines societal efforts to promote evidence-based medicine. Many relevant studies that inform evidence-based medicine are never published, a phenomenon often termed *publication bias* or *positive publication bias*. In 1979, this practice was graphically described by psychologist Robert Rosenthal as the “file drawer problem.”² Rosenthal wrote that “the extreme view

of the file drawer problem is that journals are filled with the 5% of the studies that show Type I errors, while the file drawers are filled with the 95% of the studies that show nonsignificant results.”^{2(p638)} In commercially sponsored pharmaceutical trials, the phenomenon of publication bias has been well described. Pediatric studies of antidepressant drugs showing increased suicidality in children and adolescents went systematically unpublished until public outcry resulted in disclosure and ultimately regulatory action.³ In a 2007 meta-analysis evaluating the cardiovascular risks of rosiglitazone, 35 of 42 trials conducted by the drug maker were unpublished and became available only after a lawsuit by the State of New York.⁴ The problem of publication bias became so pervasive in drug research that legislative action was ultimately required, resulting in congressional action to create a clinical trial registry maintained by the National Library of Medicine that now includes reporting of results.⁵

The article by Tzoulaki et al⁶ in this issue of *JAMA Internal Medicine* documents another troubling aspect of publication bias—selective reporting of associations between biomarkers and cardiovascular outcomes. Their findings are striking and disturbing, demonstrating that most meta-analyses of biomarkers commonly used in cardiovascular medicine show evidence of publication bias. Some of the biomarkers with strong evidence for selective reporting are