

RESEARCH ARTICLE

Research inefficiencies in external validation studies of the Framingham Wilson coronary heart disease risk rule: A systematic review

Jong-Wook Ban^{1,2*}, Lucy Abel³, Richard Stevens³, Rafael Perera³

1 Centre for Evidence-Based Medicine, University of Oxford, Oxford, United Kingdom, **2** Department for Continuing Education, University of Oxford, Oxford, United Kingdom, **3** Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

* drjwban@gmail.com



Abstract

Background

External validation studies create evidence about a clinical prediction rule's (CPR's) generalizability by evaluating and updating the CPR in populations different from those used in the derivation, and also by contributing to estimating its overall performance when meta-analysed in a systematic review. While most cardiovascular CPRs do not have any external validation, some CPRs have been externally validated repeatedly. Hence, we examined whether external validation studies of the Framingham Wilson coronary heart disease (CHD) risk rule contributed to generating evidence to their full potential.

Methods

A forward citation search of the Framingham Wilson CHD risk rule's derivation study was conducted to identify studies that evaluated the Framingham Wilson CHD risk rule in different populations. For external validation studies of the Framingham Wilson CHD risk rule, we examined whether authors updated the Framingham Wilson CHD risk rule when it performed poorly. We also assessed the contribution of external validation studies to understanding the Predicted/Observed (P/O) event ratio and *c* statistic of the Framingham Wilson CHD risk rule.

Results

We identified 98 studies that evaluated the Framingham Wilson CHD risk rule; 40 of which were external validation studies. Of these 40 studies, 27 (67.5%) concluded the Framingham Wilson CHD risk rule performed poorly but did not update it. Of 23 external validation studies conducted with data that could be included in meta-analyses, 13 (56.5%) could not fully contribute to the meta-analyses of P/O ratio and/or *c* statistic because these performance measures were neither reported nor could be calculated from provided data.

OPEN ACCESS

Citation: Ban J-W, Abel L, Stevens R, Perera R (2024) Research inefficiencies in external validation studies of the Framingham Wilson coronary heart disease risk rule: A systematic review. PLoS ONE 19(9): e0310321. <https://doi.org/10.1371/journal.pone.0310321>

Editor: Daniel Antwi-Amoabeng, Christus Oschner St. Patrick Hospital, UNITED STATES OF AMERICA

Received: June 24, 2022

Accepted: August 28, 2024

Published: September 13, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0310321>

Copyright: © 2024 Ban et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: The authors received no specific funding for this work.

Competing interests: Rafael Perera receives funding from the NIHR Oxford Biomedical Research Council (BRC), the NIHR Oxford Medtech and In-Vitro Diagnostics Co-operative (MIC), and the Oxford Martin School. Richard Stevens and Rafael Perera receive funding from the NIHR Applied Research Collaboration (ARC) Oxford and Thames Valley.

Discussion

Most external validation studies failed to generate evidence about the Framingham Wilson CHD risk rule's generalizability to their full potential. Researchers might increase the value of external validation studies by presenting all relevant performance measures and by updating the CPR when it performs poorly.

Introduction

External validation is a stage of clinical prediction rule (CPR) development where a CPR derived in a population is evaluated in different populations with a comparable health condition to the health condition of the derivation population (e.g., in another country) [1, 2]. An external validation study can generate evidence about a CPR's generalizability in two distinct ways [3].

Firstly, on a narrow scope, an external validation study can assess whether a CPR is applicable to a new population or setting [4, 5]. When a CPR performs adequately in an external validation study, clinicians can apply the CPR to populations and settings similar to those used in the external validation study, with confidence in its performance [6, 7]. However, CPRs often perform unsatisfactorily when applied to populations different from the ones they are derived, in external validation studies [8–10], which might be due to inflated accuracy from suboptimal design and methods used to derive the CPRs or differences in populations and settings between derivation and external validation studies [4, 11].

Sometimes, authors of external validation studies declare that the CPRs, when they perform poorly, are not generalizable [12–14]. There are several disadvantages for this approach. When no valid CPR exists for the population, clinicians cannot make decisions guided by a CPR until a new accurate CPR is created by often repeating the entire derivation process from ground up, because using poorly performing CPR can lead to decisions harmful to patients [15]. Moreover, this means that the knowledge accumulated while deriving the CPR is simply wasted [11, 16]. Further, this approach might promote the proliferation of many redundant CPRs for the same problem [11, 16] because the poor performance of CPR is one of the most commonly stated justifications by authors who derive a new cardiovascular CPR [17].

Alternatively, authors could attempt to update the CPR while conducting the external validation study when it performs poorly [3, 11, 16, 18]. The performance of the CPR could improve simply by adjusting the calibration intercept or by re-estimating coefficients of predictor variables [11, 19–21]. To the contrary to dismissing the CPR that performs poorly in external validation, previous research and evidence accrued is efficiently used when the CPR is updated [11, 16, 22]. Although the updated CPR still should be externally validated, with this approach, it might be unnecessary for researchers to create a new CPR just for this population by rerunning the entire derivation process [11, 16]. For example, the REGICO function was created by calibrating the Framingham Wilcon CHD risk rule in Spanish cohort [23], which has been successfully validated in various Spanish cohorts [24–26], and adopted by Spanish clinical practice guidelines [27, 28].

Secondly, and more broadly, the results of external validation studies of a CPR can contribute to understanding the CPR's overall performance across an overarching population with the same health condition when meta-analysed in a systematic review [3]. By understanding the overall performance of a CPR from the meta-analysis results, clinicians can guide their use of the CPR in practice. For example, a systematic review and meta-analysis of external

validation studies showed that the American College of Cardiology and American Heart Association (ACC/AHA) pooled cohort equation frequently overestimates cardiovascular disease (CVD) risk [21]. Based on these results, clinicians could use caution when interpreting a person's CVD risk estimated with the ACC/AHA pooled cohort equation.

For the results of an external validation study to contribute to a meta-analysis, relevant performance measures of a CPR must be provided. According to the TRIPOD statement published in 2015, calibration and discrimination measures "should be reported in all prediction model papers" [29]. However, systematic reviews have shown that external validation studies often do not report recommended performance measures, especially measures of calibration [8, 30, 31]. Most studies included in these systematic reviews predate the TRIPOD statement, and authors might not have guidance for appropriate reporting. However, it is still true that when external validation studies do not report relevant performance measures, they might not contribute to understanding a CPR's overall performance in a meta-analysis to their full potential, limiting the value of external validation studies.

While most cardiovascular CPRs do not have a timely conducted independent external validation study [32], there are a small number of cardiovascular CPRs for which many external validation studies have been performed repeatedly. In 1998, Wilson et al. [33] derived a CPR for coronary heart disease (CHD) risk using the Framingham Heart Study cohort. The Framingham Wilson CHD risk rule estimates the 10-year risk of CHD (composite of angina pectoris, acute myocardial infarction, coronary insufficiency, and CHD death) and hard CHD (all CHD outcomes except for angina pectoris) with the following predictors: age, cigarette use, diabetes mellitus status, total and high-density lipoprotein cholesterol categories, and blood pressure categories [33]. Systematic reviews [31, 34] have shown that more external validation studies for the Framingham Wilson CHD risk rule have been conducted than any other cardiovascular CPR. For example, a systematic review found that the Framingham Wilson CHD risk rule was the most frequently validated cardiovascular CPR with 89 external validation studies [31].

Therefore, in this study, we aimed to explore research inefficiencies of external validation studies of the Framingham Wilson CHD risk rule; firstly by assessing whether authors of external validation studies updated the Framingham Wilson CHD risk rule when it performed poorly, and then by evaluating whether external validation studies contributed to understanding the overall performance of the Framingham Wilson CHD risk rule.

Materials and methods

Information source and search for external validation studies

In November 2015, a forward citation search of the derivation study of the Framingham Wilson CHD risk rule [33] was conducted in Scopus. This initial search was updated in June 2020 and December 2022. We conducted our forward citation search using Scopus as a citation index because we have previously demonstrated that this method could efficiently and systematically identify CPR studies [32, 35]. Further, for citation searches in the health science field, Scopus tends to produce more robust results than Web of Science while minimally missing unique references [36, 37]. The searches were not limited to a specific language, date range, access type, or document type.

Study selection

The title and abstract of all identified references were screened, and the full-text article of relevant references was reviewed. A study was included, if

1. the CPR evaluated in the study could be unambiguously identified as the Framingham Wilson CHD risk rule,
2. the Framingham Wilson CHD risk rule was applied to a new population different from the one used in the derivation study,
3. the cardiovascular outcome evaluated was either CHD (a composite of angina pectoris, acute myocardial infarction, coronary insufficiency, and CHD death) or hard CHD (a composite of acute myocardial infarction, coronary insufficiency, and CHD death) with an average follow-up of five years or longer, and
4. a valid performance measure of a CPR such as discrimination or calibration was reported or could be estimated using data provided in the study.

To be unambiguously identified as a study evaluating Framingham Wilson CHD risk rule, authors of the study must have identified the CPR using the term "Framingham" or "Wilson" with a citation to the derivation study of the Framingham Wilson CHD risk rule [33] or presented sufficient details about the CPR including baseline survival function, regression coefficient, hazard ratio, and scoring systems [38] for the verification. When studies defined CHD outcomes using criteria different from the ones used by the Framingham Heart Study [39] due to evolving practice standards over time, we judged whether the modifications were relevant to current clinical practice. We included studies with an average follow-up of five years or longer because, for the vast majority of CPRs for cardiovascular disease risk, this was the intended timeframe for predicting outcomes according to a systematic review [31]. Studies that evaluated modified versions of the Framingham Wilson CHD risk rule were excluded.

Data collection

From each study meeting the inclusion criteria, the following data were collected: bibliographic information, cohort, geographic location, number of participants and events, duration of follow-up, and performance measures. The included studies were categorised into two groups: (1) studies in which one of the stated aims was externally validating Framingham Wilson CHD risk rule, which will be referred to as "external validation studies" in our study, and (2) studies that did not aim to externally validate the Framingham Wilson CHD risk rule but in which a valid performance measure of a CPR was reported or could be calculated using provided data.

For studies whose stated aims included externally validating the Framingham Wilson CHD risk rule, we assessed whether authors concluded that the Framingham Wilson CHD risk rule performed adequately or inadequately. Authors' conclusions were interpreted based on their use of words such as "underestimate", "overestimate", "well", "poorly", "adequate", "inadequate", and other similar terms in the abstract, results, and discussion sections. For external validation studies where authors concluded the Framingham Wilson CHD risk rule performed inadequately, we also examined whether the authors updated the Framingham Wilson CHD risk rule.

We also included studies in which a performance measure of the Framingham Wilson CHD risk rule could be obtained even if the stated aim was not external validation of the Framingham Wilson CHD risk rule, because these studies could contribute to meta-analysing the performance measures of the Framingham Wilson CHD risk rule. The first author (JWB) conducted the forward citation searches, screened the references found, reviewed the full-text articles, assessed the eligibility, and collected data from included studies. A second reviewer (LA)

verified the eligibility of included articles and collected all data independently. Any disagreements were adjudicated through discussions.

Outcome measures

Research inefficiency occurs when research studies fail to “contribute to knowledge or to practice and policy,” to their maximum potential [22, 40]. In our study, we assessed two types of inefficiency among external validation studies of the Framingham Wilson CHD risk rule. Firstly, we evaluated the inefficiency from external validation studies that do not attempt to update the Framingham Wilson CHD risk rule when it performs poorly; thus, missing the opportunity to create a CHD risk rule that could be useful for populations similar to the one used in the external validation. Therefore, one of our principal outcome measures was the proportion of external validation studies that concluded the Framingham Wilson CHD risk rule performed inadequately but did not update the CPR.

Secondly, we assessed the inefficiency from external validation studies that are potentially eligible to be included in a meta-analysis and contribute to understanding the performance of the Framingham Wilson CHD risk rule, but fail to do so because either a P/O ratio or *c* statistic is not reported or cannot be obtained from provided data. Thus, our other main outcome measure was the proportion of eligible external validation studies that could not be included in the meta-analysis of P/O ratio or *c* statistic due to the insufficient reporting of relevant performance measures. We chose to evaluate the P/O ratio because existing systematic reviews [21, 41, 42] of the Framingham Wilson CHD risk rule invariably meta-analyzed the P/O ratio to summarize calibration.

Data analysis

Using all included external validation studies, we calculated the proportion of studies in which authors concluded that the Framingham Wilson CHD risk rule (a) performed adequately, (b) performed inadequately and updated the CPR, and (c) performed inadequately but did not update the CPR.

Arguably, a meta-analysis of external validation studies should include studies conducted in populations that are “plausibly related” [43], which is not always easy to determine clearly. In our study, we took a conservative approach and included in meta-analyses only the studies conducted in the USA and European geographic regions with comparable cardiovascular disease burden, the prevalence of risk factors, and healthcare environment [44]: the UK, Northern Europe, Western Europe, and Southern Europe. We meta-analysed the Predicted/Observed (P/O) event ratio and *c* statistic of the Framingham Wilson CHD risk rule. The definitions of CHD from external validation studies included in the meta-analyses are summarized in [S1 Table](#). Random effects meta-analyses were conducted using methods described by Debray et al. [45], as described in [S1 Appendix](#). When multiple studies were conducted using data from the same cohort, only the study with the largest dataset was included in the meta-analyses. At the suggestion of a reviewer, we display results in forest plots without summary estimates when heterogeneity is high, since summary estimates from meta-analyses are incidental to our study aims rather than directly relevant to our research questions. We presented the proportion of eligible external validation studies that could not be included in the meta-analysis of P/O ratio or *c* statistic because a relevant performance measure was not reported or could not be estimated from provided data.

Additionally, we conducted post hoc citation analyses to evaluate the potential impact of external validation studies in which authors concluded that the Framingham Wilson CHD risk rule performed inadequately but did not update it. In Scopus, we ran forward citation

searches of these external validation studies, which were updated in December 2022. We searched for studies that derived a new cardiovascular CPR while using the Framingham Wilson CHD risk rule's poor performance as one of the justifications and citing one of these external validation studies.

We prepared this report following all applicable recommendations from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [46] as summarised in [S2 Appendix](#). Data from our study are provided in [S1 Dataset](#).

Results

Included studies of the Framingham Wilson CHD risk rule

A total of 98 studies that evaluated the Framingham Wilson CHD risk rule from 84 publications were included, [Fig 1](#).

Characteristics of the included studies are summarised in [Table 1](#). For 40 (40.8%) of 98 included studies, one of the aims was externally validating the Framingham Wilson CHD risk rule. For 58 (59.2%) studies, although aims did not include externally validating the Framingham Wilson CHD risk rule, a performance measure of the Framingham Wilson CHD risk rule was reported or could be estimated using data provided in the study. The lists of included articles, excluded articles, and journals that published the included studies are provided in [S3 Appendix](#), [S2](#) and [S3](#) Tables, respectively.

Performance measures of the Framingham Wilson CHD risk rule reported by the included studies are summarised in [Table 2](#). Of 40 studies that aimed to externally validate the Framingham Wilson CHD risk rule, 17 (42.5%) did not report any calibration measure. Only 14 (35.0%) studies reported calibration according to the recommendations from the TRIPOD statement: 5 (12.5%) studies with both calibration table and plot, 8 (20.0%) studies with a calibration plot only, one (2.5%) study with a calibration table only. A measure of discrimination was not reported in a minority (11 of 40, 27.5%) of studies that aimed to externally validate the Framingham Wilson CHD risk rule. However, *c* statistic was reported with a 95% confidence interval only in 17 of 40 (42.5%) external validation studies, adhering to the recommendation in the TRIPOD statement.

Inefficiency from failure to update the Framingham Wilson CHD risk rule

Authors' interpretations of the Framingham Wilson CHD risk rule's performance in external validation study and their consequent actions are summarised in [Fig 2](#). Of 40 external validation studies of the Framingham Wilson CHD risk rule, authors of 27 external validation studies (67.5%) concluded that the performance was inadequate but did not update the CPR. On the other hand, the authors of four external validation studies (10.0%) that concluded the Framingham Wilson CHD risk rule performed poorly updated the CPR: two studies by recalibrating [47, 48] and two studies by re-estimating parameters and modifying predictor variables [49, 50]. In three of these four studies, the performance of the Framingham Wilson CHD risk rule was judged to be improved after updating it.

In a post hoc citation analysis, a total of 1,341 references were found in forward citation searches of 27 external validation studies that concluded the performance of the Framingham Wilson CHD risk rule was inadequate but did not update it. From these references, we found 20 studies that derived a new cardiovascular CPR using the poor performance of the Framingham Wilson CHD risk rule in the external validation study as one of the justifications, with a citation to the external validation study. Of these 20 new cardiovascular CPRs, four were derived by overlapping authors with external validation studies, and 16 were derived by different authors, as presented in [S4 Table](#).

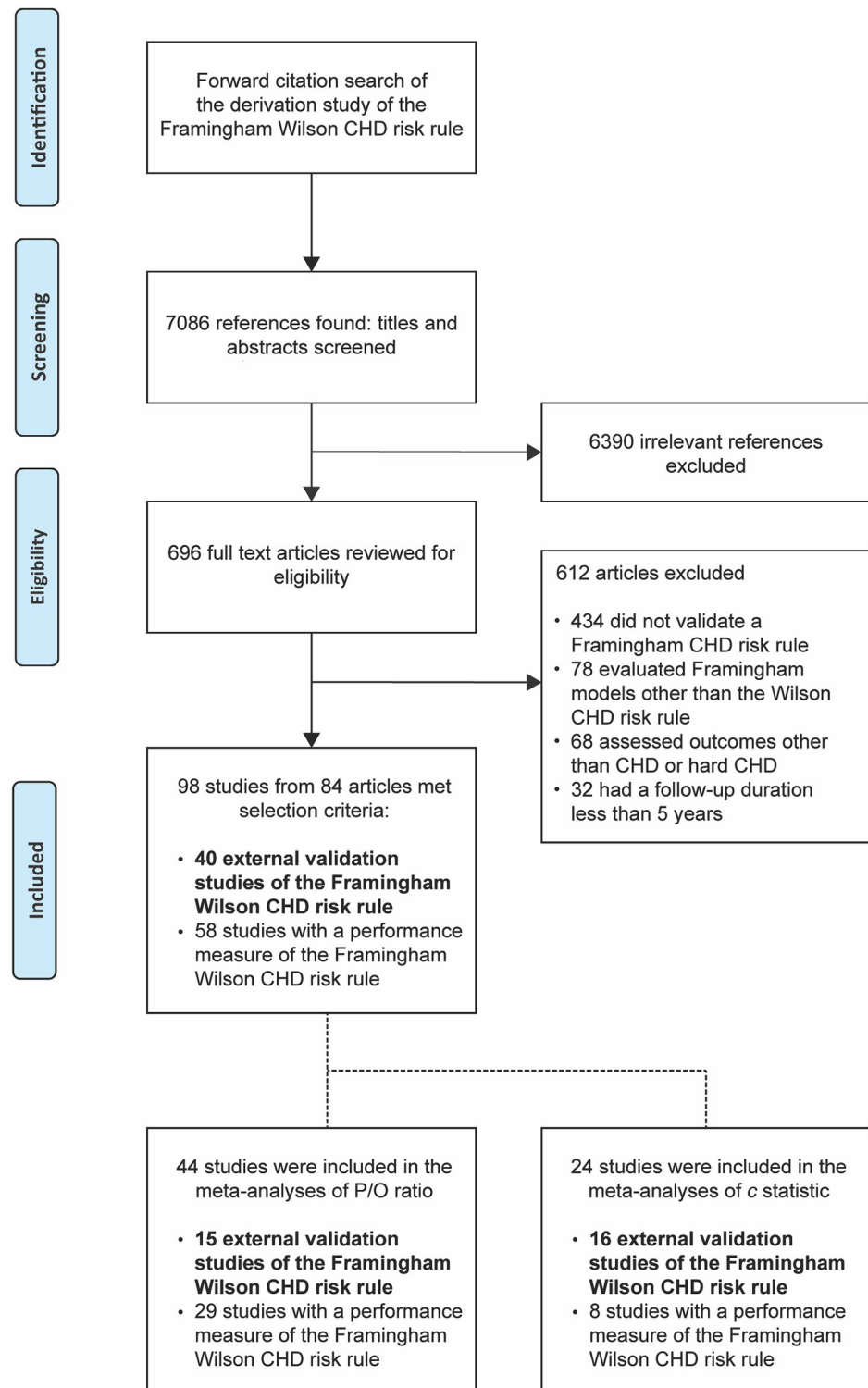


Fig 1. Search and selection of studies that evaluated the Framingham Wilson coronary heart disease (CHD) risk rule. P/O ratio: Predicted/Observed event ratio.

<https://doi.org/10.1371/journal.pone.0310321.g001>

Table 1. Characteristics of the included studies of the Framingham Wilson coronary heart disease (CHD) risk rule.

Characteristic	External validation studies of the Framingham Wilson CHD risk rule, n = 40	Other studies with a performance measure of the Framingham Wilson CHD risk rule, n = 58	All included studies, n = 98
Publication year, median (IQR)	2008 (2006–2012)	2010 (2008–2014)	2009 (2007–2014)
*Journal impact factor, median (IQR)	4.213 (3.133–8.754)	13.2825 (5.099–51.598)	7.05 (3.35–35.855)
Study location			
• Within the USA	7 (17.5)	21 (36.2)	28 (28.6)
• Outside the USA	33 (82.5)	37 (63.8)	70 (71.4)
Sample size, median (IQR)	1,437.5 (625.5–5,611)	3,124.5 (1,392–6,604)	2,474.5 (847–5,732)
Number of events, median (IQR)	131.5 (62–294)	204 (117–499)	187 (93–460)
Duration of follow-up in years, median (IQR)	10.0 (7.65–10.0)	8.6 (6.4–10.9)	9.65 (6.4–10.0)
Description of CPR being validated			
• Risk equation ^a	7 (17.5)	17 (29.3)	24 (24.5)
• Risk score ^b	4 (10.0)	9 (15.5)	13 (13.3)
• Unclear	29 (72.5)	32 (55.2)	61 (62.2)
Handling of missing data			
• Complete case analysis	30 (75.0)	27 (46.6)	57 (58.2)
• Imputation	0 (0.0)	16 (27.6)	16 (16.3)
• Unclear	10 (25.0)	15 (25.9)	25 (25.5)

Values are numbers (%) unless indicated otherwise.

* 2021 Web of Science Journal Citation Report.

^a Risk equation: regression equations that estimate the 10-year risk of developing CHD using coefficients of predictor variables.

^b Risk score: simplified scoring systems that categorise the 10-year risk of developing CHD based on the sum of points assigned to applicable predictor variables.

<https://doi.org/10.1371/journal.pone.0310321.t001>

Inefficiency from the inability to contribute to understanding the overall performance of the Framingham Wilson CHD risk rule

A total of 44 studies were included in the meta-analysis of the P/O ratio: 15 studies with an explicit aim of externally validating the Framingham Wilson CHD risk rule and 29 studies where a P/O ratio could be obtained from data. As presented in Fig 3A and 3B, the P/O ratios that were obtained showed high heterogeneity, with I^2 statistic greater than 80% in all analyses.

A total of 24 studies were included in the meta-analysis of c statistic: 16 studies with an explicit aim of externally validating the Framingham Wilson CHD risk rule and 8 studies where a c statistic could be obtained from data. Forest plots of the c statistic of the Framingham Wilson CHD risk rule are presented in Fig 4. Studies conducted in the USA, Northern Europe, and Southern Europe had high heterogeneity with a corresponding I^2 statistic of 96.98%, 96.50%, and 79.06%, respectively. The summary c statistic (95% confidence interval) from the meta-analysis of three studies conducted in the UK was 0.699 (0.680–0.718) with an I^2 statistic of 30.33%. The summary c statistic (95% confidence intervals) from the meta-analyses of five Western European studies was 0.692 (0.660–0.722) with an I^2 statistic of 49.00%.

Of 40 studies that aimed to externally validate the Framingham Wilson CHD risk rule, 23 studies were conducted in the USA and Europe: therefore could have been included in the meta-analyses as presented in Table 3. All 23 studies were eligible to be included in the meta-analysis of c static. However, only 21 of 23 studies were eligible for the meta-analysis of the P/O ratio; two studies, Vaidya 2007 [51] and Simmons 2008 [52], were ineligible because they were conducted using a subset of data from the same cohort used by another eligible study. Six

Table 2. Performance measures of the Framingham Wilson coronary heart disease (CHD) risk rule reported by included studies.

Performance measure	External validation studies of the Framingham Wilson CHD risk rule, n = 40	Other studies with a performance measure of the Framingham Wilson CHD risk rule, n = 58	All included studies, n = 98
Any calibration	23 (57.5)	9 (15.5)	32 (32.7)
• P/O ratio with 95% CI	3 (7.5)	0 (0.0)	3 (3.1)
• P/O ratio without 95% CI	5 (12.5)	0 (0.0)	5 (5.1)
• Calibration plot	13 (32.5)	1 (1.7)	14 (14.3)
• Calibration table	6 (15.0)	0 (0.0)	6 (6.1)
• Calibration intercept	2 (5.0)	0 (0.0)	2 (2.0)
• Calibration slope	1 (2.5)	1 (1.7)	2 (2.0)
• Goodness of fit test	9 (22.5)	7 (12.1)	16 (16.3)
Any discrimination	29 (72.5)	21 (36.2)	50 (51.0)
• c statistic with 95% CI	17 (42.5)	9 (15.5)	26 (26.5)
• c statistic without 95% CI	11 (27.5)	12 (20.7)	23 (23.5)
• ROC curve	16 (40.0)	5 (8.6)	21 (21.4)
• Discrimination slope	1 (2.5)	1 (1.7)	2 (2.0)
Any classification	9 (22.5)	0 (0.0)	9 (9.2)
• Sensitivity or specificity	9 (22.5)	0 (0.0)	9 (9.2)
• Positive or negative predictive value	8 (20.0)	0 (0.0)	8 (8.2)
• Positive or negative likelihood ratio	5 (12.5)	0 (0.0)	5 (5.1)
• Diagnostic odds ratio	4 (10.0)	0 (0.0)	4 (4.1)
• Accuracy	4 (10.0)	0 (0.0)	4 (4.1)
Other*	1 (2.5)	3 (5.2)	4 (4.1)

Values are the number of studies (%).

ROC: receiver operating characteristic.

* Includes measures such as R^2 and Brier score.

<https://doi.org/10.1371/journal.pone.0310321.t002>

external validation studies of the Framingham Wilson CHD risk rule could not be included the meta-analysis of P/O ratio because they neither reported a P/O ratio nor provided sufficient information about the data to estimate it. Similarly, seven external validation studies of the Framingham Wilson CHD risk rule could not contribute to the meta-analysis of c statistic because they neither reported a c statistic nor provided data to estimate it.

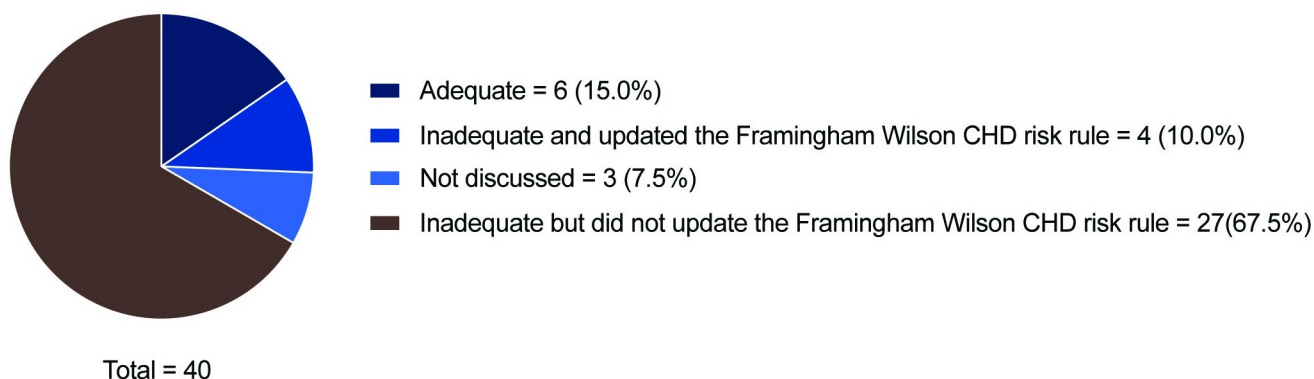


Fig 2. Authors' interpretations of the Framingham Wilson coronary heart disease (CHD) risk rule's performance in external validation study and consequent actions.

<https://doi.org/10.1371/journal.pone.0310321.g002>

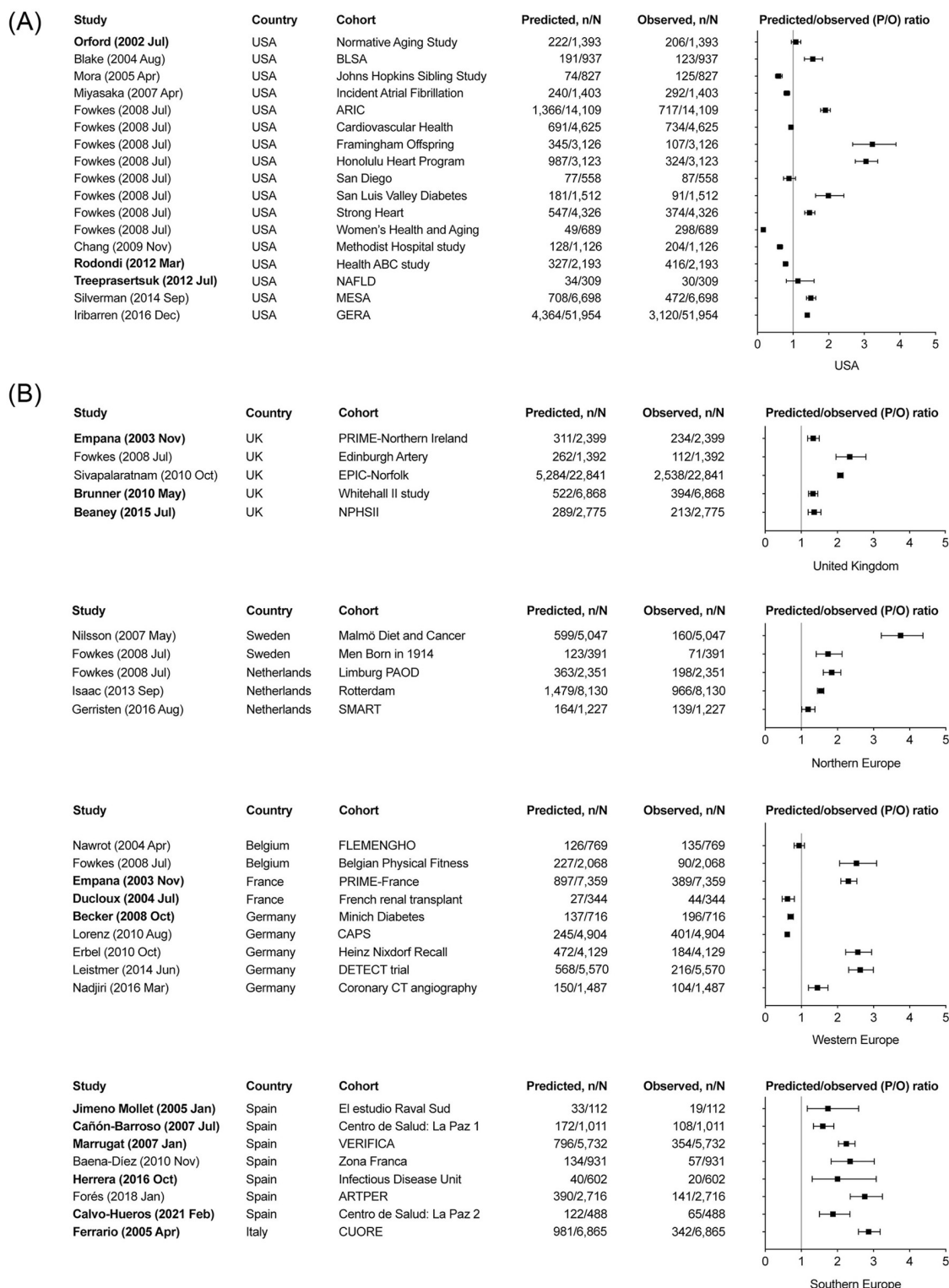


Fig 3. A. Meta-analyses of Predicted/Observed (P/O) ratio from studies conducted in the USA. Studies with an explicit aim to externally validate the Framingham Wilson coronary heart disease rule are presented in bold characters. n: the number of participants with the outcome, N: the total number of participants. **B.** Meta-analyses of Predicted/Observed (P/O) ratio from studies conducted in Europe. Studies with an explicit aim to externally validate the Framingham Wilson coronary heart disease rule are presented in bold characters. n: the number of participants with the outcome, N: the total number of participants.

<https://doi.org/10.1371/journal.pone.0310321.g003>

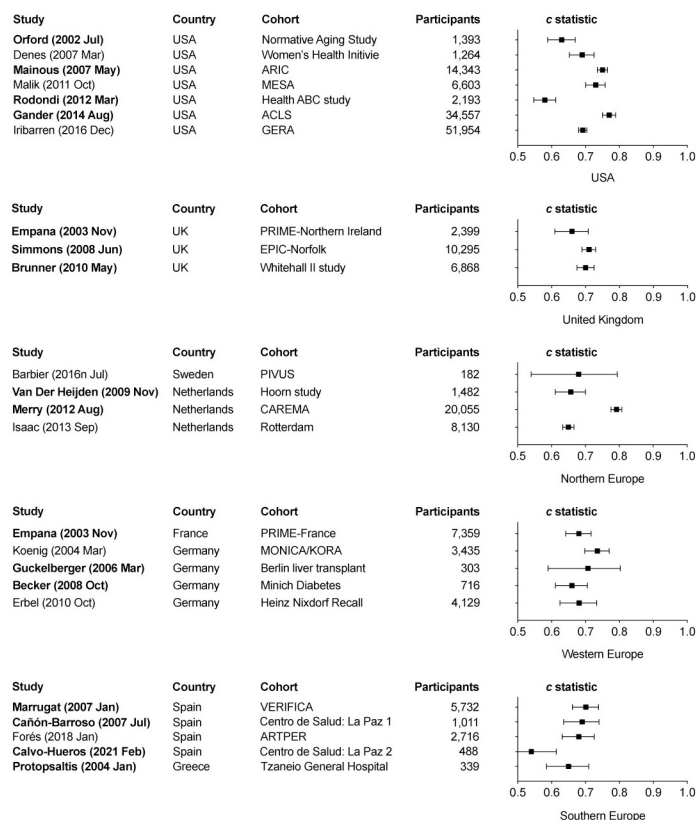


Fig 4. Meta-analyses of c statistic from studies conducted in the United States and Europe. Studies with an explicit aim to externally validate the Framingham Wilson coronary heart disease rule are presented in bold characters.

<https://doi.org/10.1371/journal.pone.0310321.g004>

In summary, of 23 studies that aimed to externally validate the Framingham Wilson CHD risk rule, only ten (43.5%) contributed to understanding its performance in eligible meta-analyses of P/O ratio and c statistic. On the other hand, 13 (56.5%) studies that aimed to externally validate the Framingham Wilson CHD risk rule did not contribute to eligible meta-analyses to their full potential because they neither reported relevant performance measures nor provided data to estimate them.

Discussion

In this study, we evaluated whether external validation studies of the Framingham Wilson CHD risk rule generated useful evidence for new populations or settings, and whether they contributed to understanding the Framingham Wilson CHD risk rule's overall performance in overarching populations.

Summary of findings

Many (67.5%) external validation studies found that the Framingham Wilson CHD risk rule performed inadequately but did not take the opportunity to update the CPR, which we consider an inefficient use of research and data. Subsequently, 20 new cardiovascular CPRs were derived using the poor performance of the Framingham Wilson CHD risk rule in these external validation studies as one of the justifications.

Table 3. Performance measures reported or calculated from data provided in 23 external validation studies of the Framingham Wilson coronary heart disease risk rule, conducted in the USA and Europe.

Study	Location	Cohort	Performance measure reported or calculated from data	
			P/O ratio	c statistic
Orford 2002 [53]	USA	Normative Aging Study	Calculated from data	Reported
Mainous 2007 [54]	USA	ARIC	Not available*	Reported ^b
Vaidya 2007 [51]	USA	Johns Hopkins Sibling study	Ineligible ^a	Not available*
Rodondi 2012 [49]	USA	Health ABC study	Calculated from data	Reported
Treeprasertsuk 2012 [55]	USA	NAFLD	Calculated from data	Not available*
Gander 2014 [56]	USA	ACLS	Not available*	Reported ^b
Emapana 2003 [57]	UK	PRIME-UK	Reported ^b	Reported
Simmons 2008 [52]	UK	EPIC-Norfolk	Ineligible ^a	Reported ^b
Brunner 2010 [58]	UK	Whitehall II study	Calculated from data	Reported ^b
Beaney 2015 [59]	UK	NPHSII	Calculated from data	Not available*
Van Der Heijden 2009 [14]	Northern Europe	Hoorn study	Not available*	Reported ^b
Merry 2012 [60]	Northern Europe	CAREMA	Not available*	Reported ^b
Empana 2003 [61]	Western Europe	PRIME- France	Reported ^b	Reported
Ducloux 2004 [62]	Western Europe	French renal transplant cohort	Calculated from data	Not available*
Guckelberger 2006 [63]	Western Europe	Charité-Universitätsmedizin Berlin liver transplant	Not available*	Reported
Becker 2008 [64]	Western Europe	Munich Diabetes	Calculated from data	Reported ^b
Jimeno-Mollet 2005 [65]	Southern Europe	El estudio Raval Sud	Calculated from data	Not available*
Ferrario 2005 [47]	Southern Europe	CUORE	Calculated from data	Not available*
Marrugat 2007 [24]	Southern Europe	VERIFICA	Calculated from data	Reported
Cañón-Barroso 2007 [66]	Southern Europe	Centro de Salud: La Paz 1	Calculated from data	Reported ^b
Herrera 2016 [26]	Southern Europe	Infectious Disease Unit at Hospital del Mar	Calculated from data	Not available*
Calvo-Hueros [67]	Southern Europe	Centro de Salud: La Paz 2	Reported	Reported ^b
Protopsaltis 2004 [68]	Southern Europe	Tzaneio General Hospital	Not available*	Reported

P/O ratio: Predicted/Observed event ratio.

* Not reported by the authors of external validation studies and could not be calculated from data provided in the external validation studies despite being conducted using data that could contribute to the meta-analyses.

^a These studies were ineligible because they were conducted using a subset of data from the same cohort used by another eligible study.

^b Reported with a 95% confidence interval.

<https://doi.org/10.1371/journal.pone.0310321.t003>

Many external validation studies of the Framingham Wilson CHD risk rule did not take the opportunity to report relevant calibration and discrimination measures. Even when authors of the external validation studies reported these performance measures of the Framingham Wilson CHD risk rule, most did not adhere to the recommendations from the TRIPOD statement. Consequently, the majority (56.5%) of external validation studies of the Framingham Wilson CHD risk rule fail to contribute to understanding the CPR's overall performance to full potential because they neither reported relevant performance measures nor provided sufficient information about the data to estimate them. These research waste and inefficiencies limit the value of external validation studies.

The heterogeneity between studies suggests that even when many previous validations have been conducted in a particular geographic location, there is still potential to learn something new due to differences in populations, changes in predictor and outcome definitions, and temporal evolution of clinical practice [69]. In particular, since the Framingham Heart Study defined these outcomes several decades ago, standard practices for diagnosing CHD have

evolved greatly [70–72]. Therefore, we do not argue that the validation studies were wasteful in themselves. Rather, we encourage authors of validation studies to maximise the value of their study: by publishing comprehensive validation statistics and by publishing an updated (“recalibrated”) rule when indicated.

Comparison with existing literature

Numerous CPRs for cardiovascular conditions have been derived [31, 34, 73, 74]. But, for some clinical conditions such as coronary heart disease or cerebrovascular accident, researchers keep deriving new CPRs despite many existing CPRs for the conditions. For instance, 363 distinct CPRs for assessing the risk of cardiovascular disease general population were found in a systematic review [31]. We found that, when authors of external validation studies concluded that the Framingham Wilson CHD risk rule performed poorly but did not attempt to update it, their conclusions were used as a rationale for deriving multiple, potentially redundant cardiovascular CPRs. These results are compatible with findings from a previous study that showed one of the most common justifications for deriving a new cardiovascular CPR was the poor performance in external validation [17]. Further, these new cardiovascular CPRs, added to the pool of hundreds of existing cardiovascular CPRs, are unlikely to be recognized or used by clinicians because most clinicians are familiar with and use only a limited number of cardiovascular CPRs [75, 76].

Many systematic reviews have pointed out various methodological and reporting issues in external validation studies. For example, systematic reviews have shown that many external validation studies include an insufficient number of participants with and without outcome events [8, 77], which can result in biased and imprecise estimation of predictive performance measures [78, 79]. Despite this, external validation studies infrequently describe how the sample size is justified [80, 81] and missing data in external validation studies are often handled by complete case analysis, which is deprecated [8, 77, 82]. Also, we found that many external validation studies of the Framingham Wilson CHD risk rule did not report calibration and discrimination measures as recommended by the TRIPOD statement. These findings are consistent with systematic reviews in diabetes, cancer, chronic kidney disease, and cardiovascular disease that showed external validation studies often do not report recommended performance measures, especially for calibration [8, 30, 31, 83, 84]. For example, Collins et al. [83] systematically reviewed studies externally validating CPRs from various clinical domains, and found that the majority (67.9%) of external validation studies did not report any calibration measures while a measure of discrimination was reported in most studies (73.1%). In addition, we found that performance measures of the Framingham Wilson CHD risk rule are often presented without their confidence intervals, although the TRIPOD statement recommends “reporting measures with confidence intervals” [38]. This is also in line with findings from existing systematic reviews [8, 84, 85]. Other systematic reviews have also shown that, while timely conducted independent external validation studies are scarce [9, 32], efforts to externally validate CPRs are disproportionately focused on a small number of CPRs such as the Framingham risk scores [10, 31, 34].

Strengths and limitations

We included 98 studies from 84 publications (40 external validation studies and 58 studies with a performance measure of the Framingham Wilson CHD risk rule) that could potentially contribute to meta-analyses of Framingham Wilson CHD risk rule by conducting a forward citation search, which is considerably more than those included in a similar systematic review [21] that used exhaustive search strategy. This re-demonstrates the efficiency and robustness of the forward citation search used in this study as a method to identify external validation studies.

On the other hand, we only relied on Scopus to conduct our forward citation search. Although the use of a single citation index is a very common practice [86], and Scopus performs well compared with Web of Science in the health science field [36, 37], searching multiple sources might have been more ideal to ensure none of the potentially relevant references was missed.

Our study has the following limitations. We used the Framingham Wilson CHD risk rule as an example and it is uncertain that findings are generalisable to other CPRs that have many external validation studies. In addition, our investigation into studies that cite studies in our review (those that conclude Framingham is inadequate, but do not update) was a post hoc addition; after discovering that most external validation studies simply dismissed the Framingham Wilson CHD risk rule when it performed poorly, we considered it necessary to understand the impact of these studies.

We assessed whether external validation studies of the Framingham Wilson CHD risk rule reported its performance measure following the recommendations from the TRIPOD statement, which was published in 2015. Because most external validation studies (all except for three) that we evaluated predate the publication of the TRIPOD statement, our findings should not be interpreted as an assessment of reporting quality against the recommendations from the TRIPOD statement. Our intention was to investigate whether external validation studies provide sufficient information so that their results could contribute to the understanding the overall performance of the Framingham Wilson CHD risk rule.

We only considered overall calibration (P/O ratio) and discrimination (*c* statistic) in particular; good validation should consider three or more dimensions [87, 88]. However, focussing on two commonly reported aspects of validation did not prevent us from demonstrating research inefficiencies in this literature. Also, we followed the approaches of Debray et al. [45], as described in [S1 Appendix](#), for extracting P/O ratios; arguably, a systematic reviewer with sufficient software and mathematical knowledge could also estimate P/O ratios from published calibration plots, albeit inexactly; this would increase the number of external validation studies of the Framingham Wilson CHD risk rule that contributed to understanding the performance in eligible meta-analyses of P/O ratio and *c* statistic from 10 of 23 (43.5%) to 11 (47.8%), but not change our overall conclusions.

Lastly, some of the external validation studies that we included were conducted using data from the same cohorts. For example, nine external validation studies were produced using data from Centro de Salud: La Paz [12, 66, 89–93], which had a total sample size of 1,011 with only 108 CHD events. Although it is arguably wasteful to repeatedly publish external validation studies of a prediction rule using a small dataset, we did not attempt to quantify whether these studies constitute research inefficiency or waste in our study.

Research and clinical implications

In this study, we identified two key inefficiencies in the external validation step of cardiovascular CPR development. Future studies (e.g. qualitative interviews or analysis of CPR registry) could further explore causes and mechanisms for these inefficiencies. Understanding factors that influence these inefficiencies to occur might help develop targeted ways to prevent them and increase the chance for external validation studies to fully contribute to knowledge and practice.

Researchers planning to externally validate a CPR should systematically evaluate existing evidence and determine whether a new external validation study is needed. When the CPR performs well, clinicians are more likely to have confidence in its prediction and adopt the CPR in practice [94, 95]. But, when the CPR performs less than satisfactorily in an external validation study, researchers should avoid simply dismissing the CPR. Instead, researchers should attempt to update the CPR and evaluate whether a clinically meaningful predictive performance can be

achieved so that the evidence generated is useful to the relevant populations and settings. Journal editors should encourage authors to provide a clear justification for conducting an external validation study of a CPR, ideally based on a systematic examination of existing evidence, and to update the CPR when it performs poorly. Further, journals should adopt the TRIPOD statement in their instructions to authors and require authors to present all relevant performance measures in external validation studies so that other researchers and clinicians could use the results of external validation studies to further evaluate and implement the CPR.

Conducting a systematic review or using evidence from systematic reviews to determine whether a new external validation study is justified can be challenging [96, 97]. Creating a living evidence system for a CPR that continuously updates meta-analyses with the latest evidence of generalizability about the CPR [98] might be useful for researchers considering an external validation study as well as clinicians and other stakeholders of CPR research. Clinicians should be made aware that a CPR that performed poorly in an external validation study might have an updated version that performs well in similar populations [24, 26, 99].

Conclusion

The focus of cardiovascular CPR research should shift from deriving many new CPRs to externally validating existing CPRs and updating them when needed because the potential end-users, such as patients, clinicians, researchers, and decision-makers, need to understand the generalizability of the CPRs. However, we have shown that conducting more external validation studies of a CPR does not always strengthen the CPR's evidence of generalizability. Further, we demonstrated that dismissing a CPR when it performs poorly can lead to the creation of new, potentially redundant CPRs. Therefore, in addition to shifting the focus of CPR research, the authors of external validation studies should ensure that they are using appropriate design and methods and providing sufficient information about the results so that the evidence they generate can help the end-users about the generalizability of the CPRs.

Supporting information

S1 Table. The definition of coronary heart disease (CHD) outcome provided by external validation studies included in the meta-analyses.

(DOCX)

S2 Table. List of excluded articles and categories for the exclusion.

(DOCX)

S3 Table. Journals that published the included studies of the Framingham Wilson coronary heart disease risk rule.

(DOCX)

S4 Table. Studies that derived a new cardiovascular clinical prediction rule using the poor performance of the Framingham Wilson coronary heart disease risk rule in the external validation study as one of the justifications.

(DOCX)

S1 Appendix. Methods for meta-analysis of predicted and observed (P/O) ratio and *c* statistic.

(DOCX)

S2 Appendix. PRISMA checklist.

(DOCX)

S3 Appendix. List of the included articles of the Framingham Wilson coronary heart disease risk rule.

(DOCX)

S1 Dataset.

(XLSX)

Author Contributions**Conceptualization:** Jong-Wook Ban, Lucy Abel, Richard Stevens, Rafael Perera.**Data curation:** Jong-Wook Ban.**Formal analysis:** Jong-Wook Ban, Lucy Abel.**Investigation:** Jong-Wook Ban, Lucy Abel, Richard Stevens, Rafael Perera.**Methodology:** Jong-Wook Ban.**Project administration:** Jong-Wook Ban.**Resources:** Jong-Wook Ban.**Supervision:** Richard Stevens, Rafael Perera.**Writing – original draft:** Jong-Wook Ban.**Writing – review & editing:** Jong-Wook Ban, Lucy Abel, Richard Stevens, Rafael Perera.**References**

1. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *Bmj*. 2009; 338:b605. Epub 2009/05/30. <https://doi.org/10.1136/bmj.b605> PMID: 19477892.
2. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *Journal of clinical epidemiology*. 2016; 69:245–7. Epub 2015/05/20. <https://doi.org/10.1016/j.jclinepi.2015.04.005> PMID: 25981519.
3. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. 2nd ed. Cham, Switzerland: Springer; 2019.
4. Altman DG, Royston P. What do we mean by validating a prognostic model? *Statistics in medicine*. 2000; 19(4):453–73. [https://doi.org/10.1002/\(sici\)1097-0258\(20000229\)19:4<453::aid-sim350>3.0.co;2-5](https://doi.org/10.1002/(sici)1097-0258(20000229)19:4<453::aid-sim350>3.0.co;2-5) PMID: 10694730.
5. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Annals of internal medicine*. 1999; 130(6):515–24. <https://doi.org/10.7326/0003-4819-130-6-199903160-00016> PMID: 10075620.
6. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Annals of internal medicine*. 2006; 144(3):201–9. Epub 2006/02/08. <https://doi.org/10.7326/0003-4819-144-3-200602070-00009> PMID: 16461965.
7. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA: the journal of the American Medical Association*. 2000; 284(1):79–84. Epub 2000/06/29. <https://doi.org/10.1001/jama.284.1.79> PMID: 10872017.
8. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC medical research methodology*. 2014; 14:40. <https://doi.org/10.1186/1471-2288-14-40> PMID: 24645774.
9. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *Journal of clinical epidemiology*. 2015; 68(1):25–34. Epub 2014/12/03. <https://doi.org/10.1016/j.jclinepi.2014.09.007> PMID: 25441703.
10. Wessler BS, Nelson J, Park JG, McGinnes H, Gulati G, Brazil R, et al. External Validations of Cardiovascular Clinical Prediction Models: A Large-Scale Review of the Literature. *Circ Cardiovasc Qual Outcomes*. 2021:CIRCOUTCOMES121007858. Epub 2021/08/04. <https://doi.org/10.1161/CIRCOUTCOMES.121.007858> PMID: 34340529.

11. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. *Journal of clinical epidemiology*. 2008; 61(11):1085–94. Epub 2009/02/12. <https://doi.org/10.1016/j.jclinepi.2008.04.008> PMID: 19208371.
12. Buitrago F, Calvo-Hueros JL, Canon-Barroso L, Pozuelos-Estrada G, Molina-Martinez L, Espigares-Arroyo M, et al. Original and REGICOR Framingham functions in a nondiabetic population of a Spanish health care center: a validation study. *Annals of family medicine*. 2011; 9(5):431–8. Epub 2011/09/14. <https://doi.org/10.1370/afm.1287> PMID: 21911762.
13. Davis WA, Colagiuri S, Davis TM. Comparison of the Framingham and United Kingdom Prospective Diabetes Study cardiovascular risk equations in Australian patients with type 2 diabetes from the Fremantle Diabetes Study. *Med J Aust*. 2009; 190(4):180–4. Epub 2009/02/18. <https://doi.org/10.5694/j.1326-5377.2009.tb02684.x> PMID: 19220181.
14. van der Heijden AA, Ortegón MM, Niessen LW, Nijpels G, Dekker JM. Prediction of coronary heart disease risk in a general, pre-diabetic, and diabetic population during 10 years of follow-up: accuracy of the Framingham, SCORE, and UKPDS risk functions: The Hoorn Study. *Diabetes care*. 2009; 32(11):2094–8. Epub 2009/10/31. <https://doi.org/10.2337/dc09-0745> PMID: 19875606.
15. Gulati G, Upshaw J, Wessler BS, Brazil RJ, Nelson J, van Klaveren D, et al. Generalizability of Cardiovascular Disease Clinical Prediction Models: 158 Independent External Validations of 104 Unique Models. *Circ Cardiovasc Qual Outcomes*. 2022; 15(4):e008487. Epub 2022/04/01. <https://doi.org/10.1161/CIRCOUTCOMES.121.008487> PMID: 35354282.
16. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012; 98(9):691–8. <https://doi.org/10.1136/heartjnl-2011-301247> PMID: 22397946.
17. Ban J-W, Wallace E, Stevens R, Perera R. Why do authors derive new cardiovascular clinical prediction rules in the presence of existing rules? A mixed methods study. *PLoS One*. 2017; 12(6):e0179102. <https://doi.org/10.1371/journal.pone.0179102> PMID: 28591223.
18. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC medical research methodology*. 2013; 13:33. <https://doi.org/10.1186/1471-2288-13-33> PMID: 23496923.
19. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Statistics in medicine*. 2004; 23(16):2567–86. <https://doi.org/10.1002/sim.1844> PMID: 15287085.
20. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. *Journal of clinical epidemiology*. 2008; 61(1):76–86. <https://doi.org/10.1016/j.jclinepi.2007.04.018> PMID: 18083464.
21. Damen JA, Pajouheshnia R, Heus P, Moons KGM, Reitsma JB, Scholten R, et al. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis. *BMC Med*. 2019; 17(1):109. Epub 2019/06/14. <https://doi.org/10.1186/s12916-019-1340-7> PMID: 31189462.
22. Ban J-W. Research waste and inefficiency in cardiovascular prediction rule development: exploration of causes and implications: University of Oxford; 2021.
23. Marrugat J, Solanas P Fau—D'Agostino R, D'Agostino R Fau—Sullivan L, Sullivan L Fau—Ordovas J, Ordovas J Fau—Cordón F, Cordon F Fau—Ramos R, et al. [Coronary risk estimation in Spain using a calibrated Framingham function]. 2003;(0300–8932 (Print)).
24. Marrugat J, Subirana I, Comin E, Cabezas C, Vila J, Elosua R, et al. Validity of an adaptation of the Framingham cardiovascular risk function: the VERIFICA Study. *J Epidemiol Community Health*. 2007; 61(1):40–7. Epub 2006/12/22. <https://doi.org/10.1136/jech.2005.038505> PMID: 17183014.
25. Comin E, Solanas P, Cabezas C, Subirana I, Ramos R, Gene-Badia J, et al. [Estimating cardiovascular risk in Spain using different algorithms]. *Rev Esp Cardiol*. 2007; 60(7):693–702. Epub 2007/08/01. PMID: 17663853.
26. Herrera S, Guelar A, Sorli L, Vila J, Molas E, Grau M, et al. The Framingham function overestimates the risk of ischemic heart disease in HIV-infected patients from Barcelona. *HIV Clin Trials*. 2016; 17(4):131–9. Epub 20160512. <https://doi.org/10.1080/15284336.2016.1177266> PMID: 27169692.
27. Rodríguez T, Fernando R, Mairata S. RIESGO CARDIOVASCULAR GUÍA DE ACTUACIÓN EN ATENCIÓN PRIMARIA2006.
28. San Vicente Blanco R, Pérez Irazusta I, Ibarra Amarica J, Berraondo Zabalegui I, Uribe Oyarbide F, Urraca Garcia de Madinabeitia J, et al. Lipidoen erabilerari buruzko praktika klinikoko gidaliburua arrisku kardiobaskularreko faktore bezala2008.
29. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine*. 2015; 162(1):W1–73. <https://doi.org/10.7326/M14-0698> PMID: 25560730.

30. Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. *BMC medicine*. 2010; 8:21. <https://doi.org/10.1186/1741-7015-8-21> PMID: 20353579.
31. Damen JA, Hooft L, Schuit E, Debray TP, Collins GS, Tzoulaki I, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *Bmj*. 2016; 353:i2416. <https://doi.org/10.1136/bmj.i2416> PMID: 27184143.
32. Ban JW, Stevens R, Perera R. Predictors for independent external validation of cardiovascular risk clinical prediction rules: Cox proportional hazards regression analyses. *Diagn Progn Res*. 2018; 2:3. Epub 2018/02/06. <https://doi.org/10.1186/s41512-018-0025-6> PMID: 31093553.
33. 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):1837–47. <https://doi.org/10.1161/01.cir.97.18.1837> PMID: 9603539.
34. Keogh C, Wallace E, O'Brien KK, Galvin R, Smith SM, Lewis C, et al. Developing an international register of clinical prediction rules for use in primary care: a descriptive analysis. *Annals of family medicine*. 2014; 12(4):359–66. <https://doi.org/10.1370/afm.1640> PMID: 25024245.
35. Ban JW, Chan MS, Muthee TB, Paez A, Stevens R, Perera R. Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules are suboptimal: a systematic review. *J Clin Epidemiol*. 2021; 133:111–20. Epub 2021/01/31. <https://doi.org/10.1016/j.jclinepi.2021.01.016> PMID: 33515655.
36. Martín-Martín A, Orduna-Malea E, Thelwall M, Delgado López-Cózar E. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *J Informetr*. 2018; 12(4):1160–77. <https://doi.org/10.1016/j.joi.2018.09.002>
37. Martín-Martín A, Thelwall M, Orduna-Malea E, Delgado Lopez-Cozar E. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics*. 2021; 126(1):871–906. Epub 2020/09/29. <https://doi.org/10.1007/s11192-020-03690-4> PMID: 32981987.
38. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Annals of internal medicine*. 2015; 162(1):55–63. <https://doi.org/10.7326/M14-0697> PMID: 25560714.
39. Kannel WB, Wolf PA, Garrison RJ, Cupples LA, D'Agostino RB, National Heart L, et al. The Framingham study: an epidemiological investigation of cardiovascular disease / Section 34: Some risk factors related to the annual incidence of cardiovascular disease and death using pooled repeated biennial measurements: Framingham heart study, 30 year followup / L. Adrienne Cupples and Ralph B. D'Agostino. [Bethesda, Md.]: National Heart, Lung and Blood Institute [Bethesda, Md.]; 1987.
40. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, et al. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014; 383(9912):156–65. [https://doi.org/10.1016/S0140-6736\(13\)62229-1](https://doi.org/10.1016/S0140-6736(13)62229-1) PMID: 24411644.
41. Brindle P, Beswick A, Fahey T, Ebrahim S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. *Heart*. 2006; 92(12):1752–9. Epub 2006/04/20. <https://doi.org/10.1136/hrt.2006.087932> PMID: 16621883.
42. Eichler K, Puhon MA, Steurer J, Bachmann LM. Prediction of first coronary events with the Framingham score: a systematic review. *American heart journal*. 2007; 153(5):722–31, 31 e1-8. <https://doi.org/10.1016/j.ahj.2007.02.027> PMID: 17452145.
43. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009. xxviii, 497 p. p.
44. Timmis A, Townsend N, Gale C, Torbica A, Lettino M, Petersen S, et al. European Society of Cardiology: Cardiovascular Disease Statistics 2019. *European Heart Journal*. 2019; 41(1):12–85.
45. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res*. 2019; 28(9):2768–86. Epub 2018/07/24. <https://doi.org/10.1177/0962280218785504> PMID: 30032705.
46. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009; 6(7):e1000097. Epub 2009/07/21. <https://doi.org/10.1371/journal.pmed.1000097> PMID: 19621072.
47. Ferrario M, Chiodini P, Chambless LE, Cesana G, Vanuzzo D, Panico S, et al. Prediction of coronary events in a low incidence population. Assessing accuracy of the CUORE Cohort Study prediction equation. *Int J Epidemiol*. 2005; 34(2):413–21. <https://doi.org/10.1093/ije/dyh405> PMID: 15659467.
48. Reissigova J, Zvarova J. The Framingham risk function underestimated absolute coronary heart disease risk in Czech men. *Methods Inf Med*. 2007; 46(1):43–9. PMID: 17224979.
49. Rodondi N, Locatelli I, Aujesky D, Butler J, Vittinghoff E, Simonsick E, et al. Framingham risk score and alternatives for prediction of coronary heart disease in older adults. *PLoS One*. 2012; 7(3):e34287. Epub 2012/03/28. <https://doi.org/10.1371/journal.pone.0034287> PMID: 22470551.

50. Nishimura K, Okamura T, Watanabe M, Nakai M, Takegami M, Higashiyama A, et al. Predicting coronary heart disease using risk factor categories for a Japanese urban population, and comparison with the framingham risk score: the suita study. *J Atheroscler Thromb*. 2014; 21(8):784–98. Epub 20140325. <https://doi.org/10.5551/jat.19356> PMID: 24671110.
51. Vaidya D, Yanek LR, Moy TF, Pearson TA, Becker LC, Becker DM. Incidence of coronary artery disease in siblings of patients with premature coronary artery disease: 10 years of follow-up. *Am J Cardiol*. 2007; 100(9):1410–5. <https://doi.org/10.1016/j.amjcard.2007.06.031> PMID: 17950799.
52. Simmons RK, Sharp S, Boekholdt SM, Sargeant LA, Khaw KT, Wareham NJ, et al. Evaluation of the Framingham risk score in the European Prospective Investigation of Cancer-Norfolk cohort: does adding glycated hemoglobin improve the prediction of coronary heart disease events? *Archives of internal medicine*. 2008; 168(11):1209–16. <https://doi.org/10.1001/archinte.168.11.1209> PMID: 18541829.
53. Orford JL, Sesso HD, Stedman M, Gagnon D, Vokonas P, Gaziano JM. A comparison of the Framingham and European Society of Cardiology coronary heart disease risk prediction models in the normative aging study. *Am Heart J*. 2002; 144(1):95–100. Epub 2002/07/03. <https://doi.org/10.1067/mhj.2002.123317> PMID: 12094194.
54. Mainous AG 3rd, Koopman RJ, Diaz VA, Everett CJ, Wilson PW, Tilley BC. A coronary heart disease risk score based on patient-reported information. *Am J Cardiol*. 2007; 99(9):1236–41. <https://doi.org/10.1016/j.amjcard.2006.12.035> PMID: 17478150.
55. Treeprasertsuk S, Leverage S, Adams LA, Lindor KD, St Sauver J, Angulo P. The Framingham risk score and heart disease in nonalcoholic fatty liver disease. *Liver Int*. 2012; 32(6):945–50. Epub 2012/02/04. <https://doi.org/10.1111/j.1478-3231.2011.02753.x> PMID: 22299674.
56. Gander J, Sui X, Hazlett LJ, Cai B, Hebert JR, Blair SN. Factors related to coronary heart disease risk among men: validation of the Framingham Risk Score. *Prev Chronic Dis*. 2014; 11:E140. <https://doi.org/10.5888/pcd11.140045> PMID: 25121352.
57. Empana JP, Ducimetiere P, Arveiler D, Ferrieres J, Evans A, Ruidavets JB, et al. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. *Eur Heart J*. 2003; 24(21):1903–11. Epub 2003/10/31. <https://doi.org/10.1016/j.ehj.2003.09.002> PMID: 14585248.
58. Brunner EJ, Shipley MJ, Marmot MG, Kivimaki M, Witte DR. Do the Joint British Society (JBS2) guidelines on prevention of cardiovascular disease with respect to plasma glucose improve risk stratification in the general population? Prospective cohort study. *Diabetic Medicine*. 2010; 27(5):550–5. <https://doi.org/10.1111/j.1464-5491.2010.02921.x> PMID: 20536951
59. Beaney KE, Cooper JA, Ullah Shahid S, Ahmed W, Qamar R, Drenos F, et al. Clinical Utility of a Coronary Heart Disease Risk Prediction Gene Score in UK Healthy Middle Aged Men and in the Pakistani Population. *PLoS One*. 2015; 10(7):e0130754. Epub 2015/07/03. <https://doi.org/10.1371/journal.pone.0130754> PMID: 26133560.
60. Merry AH, Boer JM, Schouten LJ, Ambergen T, Steyerberg EW, Feskens EJ, et al. Risk prediction of incident coronary heart disease in The Netherlands: re-estimation and improvement of the SCORE risk function. *European journal of preventive cardiology*. 2012; 19(4):840–8. Epub 2011/05/10. <https://doi.org/10.1177/1741826711410256> PMID: 21551214.
61. Empana JP, Tafflet M, Escolano S, Vergnaud AC, Bineau S, Ruidavets JB, et al. Predicting CHD risk in France: a pooled analysis of the D.E.S.I.R., Three City, PRIME, and SU.VI.MAX studies. *Eur J Cardiovasc Prev Rehabil*. 2011; 18(2):175–85. Epub 2011/04/01. <https://doi.org/10.1177/1741826710389354> PMID: 21450663.
62. Ducloux D, Kazory A, Chalopin JM. Predicting coronary heart disease in renal transplant recipients: a prospective study. *Kidney Int*. 2004; 66(1):441–7. Epub 2004/06/18. <https://doi.org/10.1111/j.1523-1755.2004.00751.x> PMID: 15200454.
63. Guckelberger O, Mutzke F, Glanemann M, Neumann UP, Jonas S, Neuhaus R, et al. Validation of cardiovascular risk scores in a liver transplant population. *Liver Transpl*. 2006; 12(3):394–401. Epub 2006/02/25. <https://doi.org/10.1002/lt.20722> PMID: 16498651.
64. Becker A, Leber AW, Becker C, von Ziegler F, Tittus J, Schroeder I, et al. Predictive value of coronary calcifications for future cardiac events in asymptomatic patients with diabetes mellitus: a prospective study in 716 patients over 8 years. *BMC Cardiovasc Disord*. 2008; 8:27. Epub 2008/10/14. <https://doi.org/10.1186/1471-2261-8-27> PMID: 18847481.
65. Jimeno Mollet J, Molist Brunet N, Franch Nadal J, Serrano Borraz V, Serrano Barragan L, Gracia Gimenez R. [Variability in the calculation of coronary risk in type-2 diabetes mellitus]. *Aten Primaria*. 2005; 35(1):30–6. Epub 2005/02/05. <https://doi.org/10.1157/13071042> PMID: 15691452.
66. Canon-Barroso L, Diaz Herrera N, Cruces Muro E, Nieto Hernandez T, Garrote Florencio T, Buitrago F. [Predictive capacity, comparison and clinic consequences of the Framingham-Wilson and REGICOR tables in persons for whom care has been provided at a health care center, Badajoz, Spain]. *Rev Esp*

- Salud Publica. 2007; 81(4):353–64. Epub 2007/11/29. <https://doi.org/10.1590/s1135-57272007000400003> PMID: 18041538.
67. Calvo-Hueros JI, Canon-Barroso L, Morales-Gabardino JA, Buitrago F. Cardiovascular risk and validation of cardiovascular risk prediction functions in a cohort of patients with type 2 diabetes followed for 10 years in Badajoz (SPAIN). AN observational study. *Prim Care Diabetes*. 2021; 15(1):115–20. Epub 20200815. <https://doi.org/10.1016/j.pcd.2020.07.014> PMID: 32811775.
 68. Protosaltis ID, Konstantinopoulos PA, Kamaratos AV, Melidonis AI. Comparative study of prognostic value for coronary disease risk between the U.K. prospective diabetes study and Framingham models. *Diabetes care*. 2004; 27(1):277–8. Epub 2003/12/25. <https://doi.org/10.2337/diacare.27.1.277> PMID: 14694007.
 69. Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. *BMC medicine*. 2023; 21(1):70. <https://doi.org/10.1186/s12916-023-02779-w> PMID: 36829188
 70. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: executive summary: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *Circulation*. 2012; 126(25):3097–137. Epub 20121119. <https://doi.org/10.1161/CIR.0b013e3182776f83> PMID: 23166210.
 71. Task Force M, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *European heart journal*. 2013; 34(38):2949–3003. Epub 20130830. <https://doi.org/10.1093/eurheartj/ehd296> PMID: 23996286.
 72. Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2021; 144(22):e368–e454. Epub 20211028. <https://doi.org/10.1161/CIR.0000000000001030> PMID: 34709928.
 73. Wessler BS, Lai Yh L, Kramer W, Cangelosi M, Raman G, Lutz JS, et al. Clinical Prediction Models for Cardiovascular Disease: Tufts Predictive Analytics and Comparative Effectiveness Clinical Prediction Model Database. *Circ Cardiovasc Qual Outcomes*. 2015; 8(4):368–75. <https://doi.org/10.1161/CIRCOUTCOMES.115.001693> PMID: 26152680.
 74. Wessler BS, Paulus J, Lundquist CM, Aijan M, Natto Z, Janes WA, et al. Tufts PACE Clinical Predictive Model Registry: update 1990 through 2015. *Diagn Progn Res*. 2017; 1:20. Epub 2017/12/21. <https://doi.org/10.1186/s41512-017-0021-2> PMID: 31093549.
 75. Pluddemann A, Wallace E, Bankhead C, Keogh C, Van der Windt D, Lasserson D, et al. Clinical prediction rules in practice: review of clinical guidelines and survey of GPs. *The British journal of general practice: the journal of the Royal College of General Practitioners*. 2014; 64(621):e233–42. <https://doi.org/10.3399/bjgp14X677860> PMID: 24686888.
 76. Ban J-W, Perera R, Stevens R. GPs' familiarity with and use of cardiovascular clinical prediction rules: a UK survey study. *BJGP Open*. 2020:bjgpopen20X101081. <https://doi.org/10.3399/bjgpopen20X101081> PMID: 33023870
 77. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS medicine*. 2012; 9(5):1–12. Epub 2012/05/26. <https://doi.org/10.1371/journal.pmed.1001221> PMID: 22629234.
 78. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multi-variable prognostic model: a resampling study. *Statistics in medicine*. 2016; 35(2):214–26. Epub 2015/11/11. <https://doi.org/10.1002/sim.6787> PMID: 26553135.
 79. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *Journal of clinical epidemiology*. 2005; 58(5):475–83. Epub 2005/04/23. <https://doi.org/10.1016/j.jclinepi.2004.06.017> PMID: 15845334.
 80. Heus P, Damen J, Pajouheshnia R, Scholten R, Reitsma JB, Collins GS, et al. Poor reporting of multi-variable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. *BMC Med*. 2018; 16(1):120. Epub 2018/07/20. <https://doi.org/10.1186/s12916-018-1099-2> PMID: 30021577.
 81. Altman DG. Prognostic models: a methodological framework and review of models for breast cancer. *Cancer Invest*. 2009; 27(3):235–43. Epub 2009/03/18. <https://doi.org/10.1080/07357900802572110> PMID: 19291527.

82. Sperrin M, Martin GP, Sisk R, Peek N. Missing data should be handled differently for prediction than for description or causal explanation. *J Clin Epidemiol*. 2020; 125:183–7. <https://doi.org/10.1016/j.jclinepi.2020.03.028> PMID: 32540389
83. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC medicine*. 2011; 9:103. <https://doi.org/10.1186/1741-7015-9-103> PMID: 21902820.
84. Collins GS, Omar O, Shanyinde M, Yu LM. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *Journal of clinical epidemiology*. 2013; 66(3):268–77. <https://doi.org/10.1016/j.jclinepi.2012.06.020> PMID: 23116690.
85. Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. *Bmj*. 2012; 344:e3318. <https://doi.org/10.1136/bmj.e3318> PMID: 22628003.
86. Briscoe S, Bethel A, Rogers M. Conduct and reporting of citation searching in Cochrane systematic reviews: A cross-sectional study. *Res Synth Methods*. 2020; 11(2):169–80. Epub 20190704. <https://doi.org/10.1002/jrsm.1355> PMID: 31127978.
87. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European heart journal*. 2014; 35(29):1925–31. Epub 20140604. <https://doi.org/10.1093/eurheartj/ehu207> PMID: 24898551.
88. Stevens RJ, Poppe KK. Reply to "Calibration slope versus discrimination slope: shoes on the wrong feet": validation stands on three feet, not two. *Journal of clinical epidemiology*. 2020; 125:162–3. Epub 20200701 <https://doi.org/10.1016/j.jclinepi.2020.06.001> PMID: 32622554.
89. Canon-Barroso L, Cruces-Muro E, Fernandez-Ochoa G, Nieto-Hernandez T, Garcia-Vellido A, Buitrago F. [Validation of 3 equations of coronary risk in diabetic population of a primary care center]. *Med Clin (Barc)*. 2006; 126(13):485–90. Epub 2006/04/21. <https://doi.org/10.1157/13086846> PMID: 16624226.
90. Canon-Barroso L, Diaz Herrera N, Calvo Hueros JI, Cruces Muro E, Nieto Hernandez T, Buitrago Ramirez F. [Incidence of cardiovascular disease and validity of equations of coronary risk in diabetic patients with metabolic syndrome]. *Med Clin (Barc)*. 2007; 128(14):529–35. Epub 2007/04/17. <https://doi.org/10.1157/13101162> PMID: 17433206.
91. Calvo Hueros JI, Canon Barroso L, Gomez Jimenez C, Martin Hidalgo-Barquero MV, Espigares Arroyo M, Buitrago Ramirez F. [Cardiovascular risk in patients with renal failure]. *Med Clin (Barc)*. 2008; 131(2):41–6. Epub 2008/07/01. <https://doi.org/10.1157/13123485> PMID: 18588827.
92. Buitrago F, Calvo-Hueros JI, Gomez-Jimenez C, Canon-Barroso L, Suarez-Gonzalez F, Robles NR. Hidden chronic renal insufficiency and cardiovascular events in patients with hypertension in a primary care center. *Ren Fail*. 2010; 32(7):757–65. Epub 2010/07/29. <https://doi.org/10.3109/0886022X.2010.493982> PMID: 20662687.
93. Canon-Barroso L, Muro EC, Herrera ND, Ochoa GF, Hueros JI, Buitrago F. Performance of the Framingham and SCORE cardiovascular risk prediction functions in a non-diabetic population of a Spanish health care centre: a validation study. *Scandinavian journal of primary health care*. 2010; 28(4):242–8. Epub 2010/09/30. <https://doi.org/10.3109/02813432.2010.518407> PMID: 20873973.
94. Graham ID, Stiell IG, Laupacis A, O'Connor AM, Wells GA. Emergency physicians' attitudes toward and use of clinical decision rules for radiography. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*. 1998; 5(2):134–40. <https://doi.org/10.1111/j.1553-2712.1998.tb02598.x> PMID: 9492134.
95. Muthee TB, Kimathi D, Richards GC, Etyang A, Nunan D, Williams V, et al. Factors influencing the implementation of cardiovascular risk scoring in primary care: a mixed-method systematic review. *Implementation science: IS*. 2020; 15(1):57. Epub 2020/07/22. <https://doi.org/10.1186/s13012-020-01022-x> PMID: 32690051.
96. Lund H, Brunnhuber K, Juhl C, Robinson K, Leenaars M, Dorch BF, et al. Towards evidence based research. *BMJ*. 2016; 355:i5440. Epub 20161021. <https://doi.org/10.1136/bmj.i5440> PMID: 27797786.
97. Lund H, Juhl CB, Norgaard B, Draborg E, Henriksen M, Andreasen J, et al. Using an evidence-based research approach before a new study is conducted to ensure value. *Journal of clinical epidemiology*. 2020. Epub 2020/09/29. <https://doi.org/10.1016/j.jclinepi.2020.07.019>
98. Adibi A, Sadatsafavi M, Ioannidis JPA. Validation and Utility Testing of Clinical Prediction Models: Time to Change the Approach. *JAMA: the journal of the American Medical Association*. 2020; 324(3):235–6. Epub 2020/03/07. <https://doi.org/10.1001/jama.2020.1230> PMID: 32134437.
99. Baena-Diez JM, Vidal-Solsona M, Byram AO, Gonzalez-Casafont I, Ledesma-Ulloa G, Marti-Sans N. The epidemiology of cardiovascular disease in primary care. the Zona Franca Cohort study in Barcelona, Spain. *Rev Esp Cardiol*. 2010; 63(11):1261–9. Epub 2010/11/13. [https://doi.org/10.1016/s1885-5857\(10\)70251-6](https://doi.org/10.1016/s1885-5857(10)70251-6) PMID: 21070722.