Research waste and inefficacy in cardiovascular prediction rule development: exploration of causes and implications



A thesis submitted for the degree of Doctor of Philosophy in Evidence-Based Health Care

Jong-Wook Ban

Kellogg College

Department of Continuing Education

University of Oxford

Hilary Term 2021 49942 words

Abstract

Background: Research waste and inefficiency occur when research studies do not contribute to the decisions of patients, clinicians, and policymakers. In this thesis, I examine research waste and inefficiency throughout the stages of cardiovascular clinical prediction rule (CPRs) development and explore their causes and implications.

Methods: This thesis includes six research projects. A mixed-methods study, which consists of a review of derivation studies, thematic content analysis, and author survey, assesses justifications for deriving a new cardiovascular CPR. Using survival analysis, I evaluate the probability for a CPR for cardiovascular disease risk to have an independent external validation study after its derivation. I assess whether external validation studies update the Framingham Wilson coronary heart disease (CHD) risk rule when it performs poorly, and external validation studies contribute to understanding the performance in meta-analyses. A systematic review evaluates the design, methods, and reporting of studies assessing the impact of cardiovascular CPRs. I conduct an online survey to understand which cardiovascular CPRs are recognised and used by GPs in the UK. Lastly, an exploratory qualitative interview study with thematic analysis explores how perceptions of and experiences with cardiovascular CPR research influence primary care providers' (PCPs) decisions about using cardiovascular CPRs.

Results: I found that the authors of derivation studies do not cite existing evidence when justifying a new cardiovascular CPR even though they believe it is important to do so. Most of cardiovascular CPRs do not have a timely conducted independent external validation study. External validation studies of Framingham Wilson CHD risk rule fail to generate the evidence of generalizability to their full potential because they do not update the CPR when it performs poorly, or do not contribute to understanding the CPR's overall performance. The vast majority of impact studies had flaws in study design, methods or reporting. GPs' use of cardiovascular CPRs in the UK has evolved and is focused on a few CPRs, such as the QRISK scores. PCPs' decisions about using cardiovascular CPRs seem to be influenced by their familiarity, trust, and confidence in using research evidence.

Conclusions: To reduce research waste and inefficiency in cardiovascular CPR development, it is necessary to shift the paradigm from deriving new CPRs to externally validating, assessing the impact of, and promoting the adoption of high-quality existing cardiovascular CPRs. Also, systematic efforts are needed to reduce the flaws in design, methods, and reporting of CPR studies, and strengthen evidence-base to support the use of cardiovascular CPRs.

Statement of contributions

I, Jong-Wook Ban, hereby confirm that this thesis entitled “Research waste and inefficiency in cardiovascular prediction rule development: exploration of causes and implications” was carried out whilst I was a part-time postgraduate student at the University of Oxford. I declare that this thesis is my own work conducted under the supervision of Professors Rafael Perera and Richard Stevens. I confirm that when others’ work is referred to, this is referenced and acknowledged appropriately.

Acknowledgement

Firstly, I would like to express my deepest gratitude to my supervisors, Professors Rafael Perera and Richard Stevens, who guided me throughout my DPhil journey and helped me prepare to become an independent researcher.

Secondly, I am immensely thankful to the late Professor Doug Altman for inspiring and encouraging me to pursue a thesis about the waste and inefficiency in clinical prediction rule research. I am also grateful to him for suggesting me to conduct a survey of authors in chapter 2.

Thirdly, I am thankful to Dr Emma Wallace and Professor Tom Fahey for granting me access to the International Register of Clinical Prediction Rules for Primary Care which was the data source for chapters 2 and 5. I also thank Dr Annette Plüddemann for providing details of methods and data from her 2014 survey study which guided planning and conducting a survey of general practitioners in chapter 6.

Fourthly, I would like to thank Drs Clare Bankhead and Ann Van den Bruel for providing constructive feedback for the thesis structure and written work submitted for the Transfer of Status examination based on chapter 2. I am also grateful to Drs Thomas Fanshawe and James Sheppard for providing invaluable comments for the written work submitted for the Confirmation of Status examination based on chapter 5.

Fifthly, I would like to thank all reviewers who took part in independent extraction and assessment of data: Dr Emma Wallace in chapter 2, Lucy Abel in chapter 4, and Mei Sum Chan, Tonny Brian Muthee, and Arsenio Paez in chapter 5. I thank general practitioners who took part in a pilot survey and giving feedback on the questionnaire in chapter 2: Drs Ann Van den Bruel, Chris Butler, Tim Holt, and David Mant. I am grateful to the Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Oxford for providing funding for the survey in chapter 6. I am also thankful to Dr Allison Cole and Ms Gina Keppel of the WWAMI (Washington, Wyoming, Alaska, Montana, and Idaho) Region Practice and Research Network for helping me recruit primary care providers for interviews in chapter 7. I thank primary care providers who took part in pilot interviews in chapter 7: Drs Paul Bunge, Ciara Smith, and Amir Atabeygi. I am very grateful to Dr Veronika Williams, who shared her expertise in qualitative research and provided crucial advice in chapter 7.

Lastly, I would like to thank my children, parents, and wife for understanding, supporting, and for the most, bearing with me for the past several years.

Table of contents

Abstract I

Statement of contributions III

Acknowledgement IV

Table of contents VI

List of tables XIV

List of figures XVII

List of abbreviations XX

Chapter 1. Background 1

1.1. Clinical prediction rule (CPR) 1

1.1.1. What is a CPR? 1

1.1.2. Stages of CPR development 3

1.2. Value of health research 5

1.2.1. Research waste 5

1.2.2. Research inefficiency 6

1.3. Inefficiency in CPR research 7

1.3.1. Research inefficiency from disparities in the evidence accumulation in stages of CPR development 8

1.3.2. Research inefficiency in the derivation stage 10

1.3.3. Research inefficiency in the external validation stage 13

1.3.4. Research inefficiency related to reporting of performance 14

1.3.5. Research inefficiency in the impact assessment and adoption stage 17

1.4. Aim and scope of the thesis 17

1.4.1. Aim of the thesis..... 17

1.4.2. Scope of the thesis 17

1.5. Structure of the thesis..... 18

1.5.1. Chapter 2. Justification for deriving a new cardiovascular CPR 19

1.5.2. Chapter 3. Independent external validation of cardiovascular risk CPRs.... 19

1.5.3. Chapter 4. Research inefficiency in external validation studies of Framingham CHD risk rule 20

1.5.4. Chapter 5. Design, methods, and reporting of impact studies of cardiovascular CPRs..... 20

1.5.5. Chapter 6. Familiarity and use of cardiovascular CPRs 20

1.5.6. Chapter 7. Influence of research evidence on the uptake of cardiovascular CPRs 21

Chapter 2. Justifications for deriving a new cardiovascular clinical prediction rule 22

2.1. Abstract 22

2.2. Introduction..... 24

2.2.1. Objectives..... 25

2.3. Methods..... 26

2.3.1. Information source 26

2.3.2. Selection of derivation studies..... 27

2.3.3. Review of derivation studies..... 29

2.3.4. Thematic content analysis 30
2.3.5. Survey of authors 32
2.3.6. Statistical analysis 34
2.4. Results 35
2.4.1. Review of derivation studies 35
2.4.2. Content analysis 37
2.4.3. Survey of authors 43
2.5. Discussion 47
2.5.1. Summary of findings 47
2.5.2. Comparison with existing literature 48
2.5.3. Strengths and Limitations 50
2.5.4. Research and clinical implications 53
2.5.5. Conclusion 58
Chapter 3. Independent external validation of prediction rules for cardiovascular disease risk 59
3.1. Abstract 59
3.2. Introduction 61
3.2.1. Objectives 63
3.3. Methods 64
3.3.1. Description of the cohort 64
3.3.2. Assessment of outcome 66
3.3.3. Predictors of an independent external validation study 69
3.3.4. Statistical analysis 74

3.4. Results 79
3.4.1. Cardiovascular CPRs with an independent external validation study 79
3.4.2. Probability for an independent external validation study 85
3.4.3. Univariable analysis 87
3.5. Discussion 93
3.5.1. Summary of findings 93
3.5.2. Comparison with existing literature 94
3.5.3. Strengths and Limitations 95
3.5.4. Research and clinical implications 100
3.5.5. Conclusion 102
Chapter 4. Inefficiencies in external validation studies of Framingham Wilson coronary heart disease risk rule 103
4.1. Abstract 103
4.2. Introduction 104
4.2.1. Objectives 109
4.3. Methods 109
4.3.1. Information source and search for external validation studies 109
4.3.2. Study selection 110
4.3.3. Data collection 113
4.3.4. Data analysis 115
4.3.5. Additional analyses 116
4.4. Results 117
4.4.1. External validation studies of the Framingham Wilson CHD risk rule 117

4.4.2. Inefficiency from failure to update the Framingham Wilson CHD risk rule 121

4.4.3. Inefficiency from the inability to contribute to understanding the overall performance of the Framingham Wilson CHD risk rule 122

4.5. Discussion 129

4.5.1. Summary of findings 129

4.5.2. Comparison with existing literature 130

4.5.3. Strengths and Limitations 132

4.5.4. Research and clinical implications 133

4.5.5. Conclusion 134

Chapter 5. Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules 136

5.1. Abstract 136

5.2. Introduction 138

5.2.1. Objectives 141

5.3. Methods 141

5.3.1. Information source and search 141

5.3.2. Eligibility criteria and study selection 143

5.3.3. Definition of an impact study 145

5.3.4. Assessment of study design 147

5.3.5. Assessment of methods 151

5.3.6. Assessment of reporting 154

5.3.7. Comparison with matched controls 156

5.3.8. Analysis 157

5.4. Results 158
5.4.1. Impact studies of cardiovascular CPR 158
5.4.2. Study design of impact study 161
5.4.3. Risk of bias in impact studies with appropriate study design 163
5.4.4. Reporting of impact studies with appropriate study design 165
5.4.5. Comparison of risk of bias and reporting with matched control studies 167
5.5. Discussion 169
5.5.1. Summary of findings 169
5.5.2. Comparison with existing literature 169
5.5.3. Strengths and Limitations 171
5.5.4. Research and clinical implications 173
5.5.5. Conclusion 175
Chapter 6. Familiarity and use of cardiovascular clinical prediction rules 176
6.1. Abstract 176
6.2. Introduction 177
6.2.1. Objectives 182
6.3. Methods 182
6.3.1. Survey method 182
6.3.2. Questionnaire development 183
6.3.3. Data collection 187
6.3.4. Analysis 189
6.4. Results 190
6.4.1. Survey participants 190

6.4.2. GPs' familiarity with cardiovascular CPRs 192
6.4.3. GPs' use of cardiovascular CPRs 194
6.5. Discussion 200
6.5.1. Summary of findings 200
6.5.2. Comparison with existing literature 202
6.5.3. Strengths and limitations 205
6.5.4. Implications for research and practice 206
6.5.5. Conclusion 207
Chapter 7. Influence of research evidence on the uptake of cardiovascular CPRs 209
7.1. Abstract 209
7.2. Introduction 211
7.2.1. Objectives 214
7.3. Methods 214
7.3.1. Participant selection 214
7.3.2. Data collection 216
7.3.3. Data management and analysis 217
7.4. Results 218
7.4.1. Participants 218
7.4.2. Main themes 220
7.4.3. "Seek and Judge" 222
7.4.4. "Be acquainted and Assume" 232
7.5. Discussion 246

7.5.1. Summary of findings 246
7.5.2. Comparison with existing literature 247
7.5.3. Strengths and limitations 250
7.5.4. Research and clinical implications 252
7.5.5. Conclusion 254
Chapter 8. Discussion and conclusion 255
8.1. Summary of findings 255
8.2. Strengths and limitations 259
8.2.1. Strengths 259
8.2.2. Limitations 261
8.3. Implications of this thesis 265
8.3.1. A framework for CPR research 265
8.3.2. Reducing waste and inefficiency within each stage of CPR development 269
8.3.3. Increasing the uptake of CPRs supported by evidence 270
8.4. Conclusion 271
References 273

List of tables

Table 1.3.1 Performance measures reported in systematic reviews. 16

Table 2.4.1 Characteristics of derivation studies according to the citation of existing cardiovascular clinical prediction rule (CPR). 35

Table 2.4.2 Authors' answers to survey questions according to the citation of existing cardiovascular clinical prediction rule (CPR). 45

Table 2.4.3 Characteristics of derivation studies of cardiovascular clinical prediction rule (CPR) that authors responded, did not respond, and could not be contacted. 46

Table 3.3.1 Potential predictors for a prediction rule for cardiovascular disease risk to have an independent external validation. 73

Table 3.4.1 Characteristics of prediction rules for cardiovascular disease risk and their derivation studies. 79

Table 3.4.2 Number of derivation and independent external validation studies of cardiovascular risk prediction rules published by journals. 84

Table 3.4.3 Univariable Cox proportional hazards regression analyses. 91

Table 4.4.1 Characteristics of included studies. 119

Table 4.4.2 Performance measures of the Framingham Wilson CHD risk rule reported by included studies. 120

Table 4.4.3 Performance measures reported or calculated from data provided in external validation studies of the Framingham Wilson coronary heart disease risk rule. 126

Table 5.2.1 Systematic reviews that evaluated design, methods, or reporting of clinical prediction rule (CPR) studies. 139

Table 5.3.1 Appropriate and inappropriate study design for studies evaluating the effectiveness of a clinical prediction rule adopted from the Cochrane Effective Practice and Organization of Care review guideline (297). 147
Table 5.3.2 Bias domains in the Cochrane Collaboration's risk of bias assessment tools. 153
Table 5.4.1 Cardiovascular clinical prediction rules evaluated in impact studies. 160
Table 5.4.2 Characteristics of impact studies of cardiovascular clinical prediction rule with appropriate and inappropriate study design. 162
Table 5.4.3 Impact studies with appropriate study design but with substantial risk of bias. 164
Table 6.2.1 Studies that evaluated the familiarity or use of cardiovascular clinical prediction rule. 179
Table 6.3.1 Clinical prediction rule (CPRs) included in the 2014 and 2019 surveys. 185
Table 6.4.1 Characteristics of general practitioners (GPs) and their practices. 191
Table 6.4.2 General practitioners (GPs) unfamiliar with clinical prediction rules (CPRs): GPs who had never heard of or who were not sure whether they heard of respective CPRs. 193
Table 6.4.3 Frequency of using cardiovascular clinical prediction rules (CPRs) by all general practitioners (GPs) who participated and GPs who reported that CPR was integrated in electronic health record, n (%). 198
Table 6.4.4 General practitioners (GPs) reporting respective reasons for using clinical prediction rules. 199

Table 7.2.1 Studies that reported the influence of research evidence on using clinical prediction rule (CPRs). 212
Table 7.4.1 Characteristics of primary care providers (PCPs) who completed the online questionnaire. 219
Table 7.4.2 Selection of primary care providers (PCPs) for interview according to the sampling frame. 220

List of figures

Figure 2.4.1 Insufficiencies of existing cardiovascular clinical prediction rules stated by authors in each derivation study. 38

Figure 2.4.2 A flow diagram illustrating how the corresponding or first authors of derivation studies were contacted, and their responses were received. 44

Figure 2.5.1 A framework to reduce the derivation of redundant clinical prediction rules (CPRs) and increase the value of CPR research. 55

Figure 3.3.1 Distribution of sample size variable before and after logarithmic transformation. 75

Figure 3.3.2 Assessment of proportional-hazard assumption for univariable Cox regression models with derivation-related predictor variables using scaled Schoenfeld residuals. 77

Figure 3.3.4 Assessment of proportional-hazard assumption for univariable Cox regression models with reporting and publication-related predictor variables using scaled Schoenfeld residuals 78

Figure 3.4.1 Publication year of derivation studies and independent external validation studies. 81

Figure 3.4.2 A flow diagram. Identifying prediction rules for cardiovascular disease risk with an independent external validation study. 83

Figure 3.4.3 Kaplan-Meier plot. Probability for a prediction rule for cardiovascular disease risk to have an independent external validation study and 95% confidence interval. 86

Figure 3.4.4 Kaplan-Meier plot. Influence of derivation-related variables on the probability for a cardiovascular clinical prediction rule to have an independent external validation study. 88
Figure 3.4.5 Kaplan-Meier plot. Influence of reporting and publication-related variables on the probability for a prediction rule for cardiovascular disease risk to have an independent external validation study. 89
Figure 4.2.1 A conceptual framework of efficient evidence generation from external validation studies of a clinical prediction rule. The figure is an adaptation of Fig 17.1 from “Clinical prediction models: a practical approach to development, validation, and updating” by Steyerberg (72). 105
Figure 4.4.1 Search and selection of studies that evaluated the Framingham Wilson CHD risk rule. 118
Figure 4.4.2 Authors’ interpretations of Framingham Wilson CHR risk rule’s performance in external validation study and consequent actions. 122
Figure 4.4.3 Meta-analyses of Predicted/Observed (P/O) ratio from studies conducted in the USA and Europe..... 125
Figure 4.4.4 Meta-analyses of *c* statistic from studies conducted in the USA and Europe. 128
Figure 5.4.1 A flow diagram summarizing the search and selection of impact studies of cardiovascular clinical prediction rule (CPRs). 159
Figure 5.4.2 Distribution of risk-of-bias judgement in bias domains for (A) randomised controlled trials, (B) cluster randomised trials, and (C) non-randomised experiments. 165

Figure 5.4.3 Frequency and cumulative distribution of the number of reporting domains that impact studies of cardiovascular clinical prediction rules complied with. 166
Figure 5.4.4 Number of reporting domains that impact and control studies complied with according to (A) study design and (B) journal where studies were published 168
Figure 6.4.1 The proportion of general practitioners using clinical prediction rules for cardiovascular disease in the 2014 and 2019 survey. 195
Figure 6.4.2 The proportion of general practitioners using clinical prediction rules for stroke and venous thromboembolism in the 2014 and 2019 survey. 196
Figure 7.4.1 A concept map of the theme "Seek and Judge." 221
Figure 7.4.2 A concept map of the theme "Be acquainted and Assume." 222
Figure 8.3.1 A framework to reduce research waste and inefficiency from disparities in evidence accumulation, increase the value of clinical prediction rule (CPR) research, and promote the implementation of CPRs supported by evidence 266

List of abbreviations

CHD	Coronary heart disease
CVD	Cardiovascular disease
CI	Confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CPM	Clinical prediction model
CPR	Clinical prediction rule
DVT	Deep vein thrombosis
EBM	Evidence-based medicine
EHR	Electronic health record
HR	Hazard ratio
IQR	Interquartile range
GP	General practitioner
NCEP	National Cholesterol Education Program
PE	Pulmonary embolism
RCT	Randomised controlled trial
TRIPOD	Transparent Reporting of a multivariable prediction model for individual Prognosis Or Diagnosis
USA	United States of America
UK	United Kingdom

Chapter 1. Background

1.1. Clinical prediction rule (CPR)

1.1.1. What is a CPR?

Although definitions of a CPR vary in the literature (1-3), it generally refers to a tool that estimates the likelihood of a health outcome using information from health history, physical examination, and test results of a person. For example, Laupacis et al. (4) defined a CPR as,

> "A decision-making tool for clinicians that included 3 or more variables obtained from the history, physical examination, or simple diagnostic tests and that either provided the probability of an outcome or suggested a diagnostic or therapeutic course of action."

CPRs enable clinicians to incorporate multiple factors in making healthcare decisions empirically rather than intuitively (5). Generally, CPRs can be categorised into two groups according to the health outcome that the CPRs evaluate: diagnostic and prognostic CPRs. Diagnostic CPRs assess how likely it is that a person has a health condition at the present time. Prognostic CPRs predict how likely it is for an individual to develop a health outcome in the future (2, 6).

Additionally, CPRs may be categorised as assistive or directive rules according to how predictions are implemented. Assistive rules help clinicians make decisions by simply presenting information such as a predicted probability of an

outcome, whereas directive rules provide corresponding recommendations about how to further evaluate or manage a patient's problem (7-10). Although the "rule" in CPRs might give an impression that they are directive decision rules, the term "CPRs" will apply to both assistive and directive rules throughout the thesis.

Clinicians use CPRs in various ways. Many CPRs are used to guide a diagnostic or therapeutic decision. For example, clinicians evaluating patients with an acute sore throat, which is most often caused by a viral infection, can use the FeverPain score (11) to identify patients with high risk for having group A streptococcal pharyngitis. Limiting antibiotic treatment to high-risk patients according to the FeverPain score can reduce antibiotic use while improving patients' symptoms (12). Similarly, clinicians evaluating patients with community-acquired pneumonia in an outpatient setting can use CURB 65 (13) or Pneumonia Severity Index (14) to identify patients who may benefit from a hospitalisation (15, 16).

Some CPRs are used to guide decisions on the prevention of disease. For example, major guidelines on the primary prevention of cardiovascular disease in the United Kingdom (UK) (17, 18) and the United States of America (USA) (19) recommend using CPRs such as QRISK2 (20), ASSIGN score (21), and American College of Cardiology/American Heart Association (ACC/AHA) pooled cohort equation (22) to estimate the risk of cardiovascular disease. Clinicians often use these CPRs to communicate cardiovascular disease risk to patients

and decide whether to start an intervention (3, 19), such as the HMG-CoA reductase inhibitor.

1.1.2. Stages of CPR development

CPR development is a multi-stage process: derivation, external validation, impact assessment, and adoption. In each CPR development stage, unique evidence about a CPR is generated (1, 7, 23). The evidence accumulated through these stages of CPR development forms the scientific premise for using CPR in practice.

In the derivation stage, researchers identify variables potentially predictive of a health outcome and build a CPR typically using mathematical methods such as logistic or Cox proportional hazards regression with data from a defined population and setting (24). How well the CPR predicts the health outcome, or the accuracy of the CPR, is presented using performance measures such as discrimination (e.g. *c* statistic) and calibration (e.g. calibration intercept and slope) (25-27). Researchers can conduct an internal validation to assess whether the CPR's accuracy is reproducible within parts of the same data used to derive the CPR and present a more realistic performance measure of the CPR (24, 28).

In the external validation stage, a CPR is tested in populations and settings different from the ones used to derive the CPR (29, 30). When the CPR

accurately predicts the health outcome in new populations and settings, the CPR gains evidence of generalizability (26). The accuracy of CPRs confirmed in various populations and settings might be better trusted than the accuracy of CPRs without or with limited evidence of generalizability (1, 7). When a CPR performs less well than desired in different populations and settings, external validation provides opportunities to update the CPR using various methods (2, 31, 32).

In the impact assessment stage, whether using a CPR improves a patient outcome, care process, or cost of care is tested similarly to evaluating the effect of implementing other types of complex interventions (23). Therefore, the evidence about the CPR's effectiveness is generated in this stage, ideally using experimental study design such as cluster randomised trials (2, 23, 33). As a CPR successfully completes these three stages, the CPR accumulates evidence of accuracy, generalizability, and effectiveness that forms the evidence-base for using the CPR in practice (1, 7).

The purpose of a CPR is fulfilled when clinicians and patients use the CPR and hence the decisions made using the CPR improve patient care. Even when a CPR has accumulated promising evidence of accuracy, generalizability, and effectiveness, clinicians might not be inclined to use the CPR or a short-term success in improved practice from the CPR implementation might not be sustained due to various barriers (1, 2, 23). Therefore, in the adoption stage,

which is the last stage of CPR development, key evidence evaluated is a sustained uptake of the CPR and consequent improvement in practice.

1.2. Value of health research

Health research is valuable when it creates new knowledge that improves health care decisions of patients, clinicians, or policymakers, or when other researchers can use it to develop such knowledge later on, for example, in a systematic review (34). Unfortunately, it has been suggested that the majority of health research studies add little value to society because the results of the studies do not contribute to creating knowledge useful to clinical practice or policy decision (35, 36).

1.2.1. Research waste

The concept of waste in health research can be found in a 1994 paper titled “the scandal of poor medical research” by Doug Altman, where he drew attention to the issue of exhausting an enormous amount of resources on health research conducted with methodological flaws (37). In 2009, Chalmers and Glasziou (36) first used the term “research waste” and described how avoidable waste is generated in four phases of the research process: planning, design and conduct, publication, and reporting. Chalmers and Glasziou (35, 36) also estimated that 85% of all health research studies were research waste and resources required to carry out these studies were needlessly wasted. Later, Chalmers et al. (38) defined research waste as “research projects that

contribute nothing or very little to knowledge or to practice and policy” in a paper that was published as part of a 2014 Lancet series that described key sources of research waste and recommended solutions (38-42).

There have been several initiatives to tackle waste in health research. For example, James Lind Alliance Priority Setting Partnerships involved patients and clinicians in identifying ten research priorities in various clinical areas so that research studies answer questions that are most relevant to patients and clinicians (43). The Evidence-Based Research Network has been promoting the use of systematic review of existing evidence before starting a new study to avoid wasteful duplication (44). The EQUATOR (Enhancing the QUAlity and Transparency Of health Research) Network encourages the development and use of high-quality reporting guidelines to advance reporting quality of health research studies and make research reports useful to patients, clinicians, and other researchers (45).

1.2.2. Research inefficiency

The term “research inefficiency” was used alongside the term “research waste” in two papers (40, 42) of the previously mentioned Lancet series and one commentary (46) to the Lancet series. In addition, the term “efficiency” was used by authors of other papers from the Lancet series when discussing resource use (38, 39), conducting systematic reviews (38, 42), the impact of high-quality study protocols (38), research reporting (42), using existing evidence to design a new study (39), and replication of research (39). However,

the authors of these papers neither provided a definition of research inefficiency nor explained a distinction between research inefficiency and research waste.

For this thesis, I define research inefficiency to occur when research studies fail to contribute to knowledge or to practice and policy, to their maximum potential. I use this definition of research inefficiency to capture research practices that need to be changed to increase the value of research but do not meet the definition of research waste. For example, when researchers of a study provide insufficient details in a report or fail to publish the primary data or statistical codes used, the study cannot deliver the maximum contribution to creating knowledge. But the study still created new knowledge that might be useful to patients, clinicians, and other researchers if the design and methods were appropriate, and therefore does not meet the definition of research waste. In this context, research waste can be viewed as the most severe subset of research inefficiency.

1.3. Inefficiency in CPR research

Inefficiencies may arise from any phase of the health research process: planning, design and conduct, publication, and reporting. (36, 46). CPR research is no exception, and this principle applies. In addition, disparities in the evidence accumulation for the different stages of CPR development are another important source of inefficiencies in CPR research.

1.3.1. Research inefficiency from disparities in the evidence accumulation in stages of CPR development

CPRs that have not successfully accumulated high-quality evidence of accuracy, generalizability, and effectiveness through the CPR development stages (1, 7) cannot contribute to fully evidence-based clinical decision making.

Unfortunately, CPR research has been disproportionately focused on the derivation stage compared with external validation and impact assessment stage (7, 47), leaving many CPRs without complete evidence-base for using them in practice. Systematic reviews have shown that the lack of evidence of generalizability or effectiveness is one of the key barriers in using CPRs by clinicians (48, 49). Therefore, CPRs for which evidence of generalizability and effectiveness is not established in external validation and impact studies, represent research inefficiency because they are unlikely to be trusted and used by clinicians and, thus, cannot support clinical decision making to their maximum potential. Furthermore, only a limited number of CPRs seem to be used in clinical practice (3) and, although discussed in the literature (23), there is no clear example of a CPR with evidence of sustained practice improvement.

The research inefficiency from disparities in the evidence accumulation in stages of CPR development was described in three early studies that surveyed CPR publications in general medical journals (4, 7, 50). Wasson et al. (50) were the first that reported this research inefficiency in 1985. They found, of 33 CPR studies published in four journals (Annals of Internal Medicine, British Medical

Journal, Journal of the American Medical Association, and New England Journal of Medicine) between 1981 and 1984, only five presented the performance of a CPR in external validation, and two reported the clinical impact of using a CPR (50). A decade later, Laupacis et al. (4) reported that, of 30 CPR studies published in the same journals published between 1991 and 1994, 11 prospectively validated a CPR and one assessed the impact of implementing a CPR. A study by Reilly et al. (7) in 2006 showed that there were 16 broad validation studies and six impact studies among 41 CPR studies published in those journals between 2000 and 2003.

More recent systematic reviews and CPR registries illustrate a broader picture of this research inefficiency, not limited to certain general medical journals (47, 51-53). A systematic review conducted to develop an International Register of Clinical Prediction Rules for Primary Care found that 196 (21.9%) of 895 CPR studies included in the International Register were broad validation studies and only 16 (1.8%) were impact studies. Although initial systematic reviews (51, 52) for the Tufts PACE Clinical Predictive Model Registry excluded validation studies, the registry currently reports that 575 (41.6%) of 1,382 cardiovascular CPRs have a validation study (54). Lastly, a systematic review by Damen et al. (53) found that only 132 (36.4%) of 363 CPRs for cardiovascular disease risk have been externally validated.

While most CPRs do not have an external validation study, some CPRs have been externally validated many times (47, 53). For example, a systematic

review (47) found 34 validation studies of Alvarado score for acute appendicitis (55) and another systematic review (53) found 89 studies that externally validated Framingham Wilson coronary heart disease (CHD) risk rule (56).

It is unknown why researchers keep creating new CPRs when there are many CPRs for the same problem. It is also unclear why some CPRs are externally validated, sometimes repeatedly, and others are not. It might be hypothesized that researchers often do not examine why it is necessary to derive or externally validate a CPR. Therefore, these disparities in evidence accumulations in different CPR development stages correspond to research inefficiency due to choosing “research question irrelevant to clinicians and patients” in the planning phase of the research process (36, 46). Even though these disparities in evidence accumulations in different stages of CPR development have been described (4, 7, 47, 50, 53), to the best of my knowledge, no research has been done to evaluate the cause of this research inefficiency.

1.3.2. Research inefficiency in the derivation stage

Several systematic reviews have shown that many derivation studies are conducted using suboptimal methods and provide insufficient details in the report (57-67). Although missing information on predictor variables or outcomes should be addressed using multiple imputation so that the risk of bias in results can be minimised (68-70), mishandling of missing data is common among derivation studies (57-64, 67, 71). For example, systematic reviews have shown that multiple imputation was rarely applied: only in 1 of 39 (2.6%) studies

deriving prediction rules for prevalent or incident type 2 diabetes mellitus (59), 2 of 11 (18.2%) studies deriving prediction rules for chronic kidney disease (62), and 5 of 48 (10.4%) derivation studies published in six high-impact general medicine journals (60).

On the other hand, systematic reviews have found that complete case analysis was commonly used: in 29 of 61 (47.5%) studies deriving prognostic rules for breast cancer (58), 21 of 39 (53.8%) studies deriving prediction rules for prevalent or incident type 2 diabetes mellitus (59), 9 of 15 (60.0%) studies deriving prognostic rules for prostate cancer (67), and 33 of 48 (68.8) derivation studies published in six high-impact general medicine journals (60). These derivation studies with complete case analyses may lead to an unnecessary reduction in sample size and biased estimation of predictive performance measures (68, 72), representing research inefficiency.

Further, systematic reviews have found that internal validation using methods such as bootstrapping or cross-validation is infrequently presented in derivation studies (57, 59, 60, 63-67, 73). Not offering internal validation is a missed opportunity for researchers to test the reproducibility and provide a more realistic performance of a CPR to clinicians and other researchers.

The TRIPOD statement recommends derivation studies “present the full prediction model to allow predictions for individuals” (6). Logistic regression models should be presented using an intercept and coefficients for predictor

variables included in the final model (6, 74). For survival models, regression coefficients must be presented with baseline survival at relevant time points (6, 74). In addition, how continuous variables were measured (e.g. kg) and how categorical variables were coded (e.g. non-smoker = 0, smoker = 1) should be clearly described to calculate an individual's risk (6, 74).

Unfortunately, systematic reviews have shown that derivation studies do not always clearly report the final model (53, 58, 60, 64, 73, 75-78). For example, only 13 of 54 (24.1%) prognostic models for patients with breast cancer presented regression coefficients for all predictor variables (58), and only 6 of 17 (35.3%) prediction models for incidence of breast cancer reported the final model (76). Similarly, 11 of 41 prognostic models for mortality in very premature infants did not report the final model (75). In addition, although regression coefficients were provided in 10 of 11 logistic regression models for the prognosis of patients with aneurysmal subarachnoid haemorrhage in a systematic review, the intercept was reported only in 2 models (77). Finally, Damen et al. (53) found that only 167 of 363 (46.0%) CPRs for cardiovascular disease risk were presented using a complete regression formula.

When derivation studies do not describe a full model, researchers and clinicians cannot evaluate the model further or use the model in clinical practice. Therefore, these derivation studies are also research inefficiency or waste that do not contribute to knowledge or practice to their full potential.

Some systematic reviews have found a higher proportion of studies reporting the full prediction model. For example, Mallet et al. (73) found that 34 of 47 (72.3%) prognostic models in cancer, 44 of which were Cox models, reported regression coefficients. Similarly, Bouwmeester et al. (60) have shown that 11 of 14 (79%) derivation studies in general medicine journals reported the final model. However, these systematic reviews only searched high-impact journals and may not accurately reflect reporting practices in the fields.

1.3.3. Research inefficiency in the external validation stage

Research inefficiency in the external validation stage has been less well studied than the research inefficiency in the derivation stage. Systematic reviews of external validation studies have shown that the quality of methods and reporting was often poor (58, 60, 65, 66, 79). For example, many external validation studies include an insufficient number of participants with and without outcome events (60, 79) which can result in biased and imprecise estimation of predictive performance measures (80, 81). Despite this, external validation studies infrequently describe how the sample size is justified (58, 65) and also complete case analyses are common (60, 79), which can lead to a smaller sample size. Because external validation studies infrequently present both discrimination and calibration measures of CPRs (65, 79), the results of external validation studies often cannot be incorporated in a meta-analysis. It has been widely argued that dismissing a CPR that performs poorly in an external validation study is an inefficient use of evidence and researchers of the external validation study should update the CPR that performs poorly (2, 8, 29, 72, 82). However, it is not

yet known whether CPRs are updated when they perform poorly in external validation studies.

1.3.4. Research inefficiency related to reporting of performance

The TRIPOD statement recommends “presenting calibration and discrimination in all studies deriving or externally validating prediction models” (74). Calibration is a measure of agreement between predictions made by a CPR and actual outcome events observed (25). According to the TRIPOD statement, calibration should be reported graphically using a calibration plot or alternatively using a calibration table (74). Although less ideal, calibration of a prediction rule might be presented with calibration-in-a-large or with calibration intercept and slope (27, 83). Although statistical tests for goodness of fit, such as the Hosmer-Lemeshow test, are often used to report calibration, the TRIPOD statement (74) and existing literature (27, 83, 84) recommends against using them because they are influenced by sample size and grouping (85-87) and uninformative about how much a CPR overestimates or underestimates the risk (74).

Discrimination assesses a CPR's ability to categorise individuals with and without the outcome of interest (25). The TRIPOD statement recommends reporting discrimination of a CPR using *c* statistic (74). Equivalent to the area under the receiver operating characteristic curve, *c* statistic is the probability that a CPR predicts a higher risk of an outcome for the individual with the outcome when two randomly selected individuals with and without the outcome are presented (88). Classification measures such as sensitivity and specificity

are not recommended because they require dichotomising a prediction rather than generating a probability (74). Lastly, overall performance measures, "distance between the observed status and the predicted probability" (89), such as R^2 and Brier score, may be presented in addition to calibration and discrimination measures (25, 74).

As shown in Table 1.3.1, many systematic reviews have found that prediction rule studies do not always report recommended performance measures (53, 59, 60, 62, 73, 79). For example, external validation studies reporting a prediction rule's calibration measure are uncommon: only 12.5% for prognostic rules for cancer (16) and 58.4% for prediction rules for cardiovascular disease risk (53). In addition, a systematic review found that a calibration plot, which the TRIPOD statement recommends, is presented only in 14.1% of studies externally validating a prediction rule (79).

When a study does not report how well a CPR predicts an outcome, it will be hard for other researchers and clinicians to judge whether to evaluate the prediction rule further or use it in clinical practice. Therefore, studies evaluating a clinical prediction rule without fully describing relevant performance measures are research inefficiency or waste that cannot contribute to knowledge and practice to their full potential.

Table 1.3.1 Performance measures reported in systematic reviews.

Performance measure	Mallet et al. (73) 2010, n = 42 [a]	Mallet et al. (73) 2010, n = 16 [c]	Collins et al. (59) 2011, n = 39 [b]	Bouwmeester et al. (60) 2012, n = 15 [a]	Collins et al. (62) 2013, n = 11 [b]	Collins et al. (79) 2014, n = 78 [c]	Damen et al. (53) 2016, n = 363 [a]	Damen et al. (53) 2016, n = 474 [c]
Any calibration	2 (4.8)	2 (12.5)	10 (25.6)	-	6 (54.5)	25 (32.1)	116 (32.0)	277 (58.4)
Calibration plot	-	-	2 (5.1)	4 (26.7)	3 (27.3)	11 (14.1)	31 (8.5)	122 (25.7)
Calibration table	-	-	-	-	-	-	34 (9.4)	62 (13.1)
Calibration intercept and slope	-	-	-	0 (0.0)	-	-	-	-
Predicted/observed ratio	-	-	-	-	-	-	12 (3.3)	124 (26.2)
Goodness of fit test	-	-	8 (20.5)	4 (26.7)	6 (54.5)	16 (20.5)	67 (18.5)	68 (14.3)
Any discrimination	-	-	39 (100.0)	-	9 (81.8)	57 (73.1)	163 (44.9)	306 (64.6)
C statistic	8 (19.0)	10 (62.5)	39 (100.0)	12 (80.0)	9 (81.8)	57 (73.1)	143 (39.4)	303 (63.9)
D statistic	0 (0.0)	0 (0.0)	1 (2.6)	-	-	-	5 (1.4)	45 (9.5)
Any overall performance	2 (4.8)	4 (25.0)	-	-	-	-	35 (9.6)	68 (14.3)
Brier score	-	-	-	1 (6.7)	-	5 (6.4)	15 (4.1)	45 (9.5)
R^2	-	-	-	2 (13.3)	-	13 (16.7)	13 (3.6)	49 (10.3)
Any other measures	-	-	-	-	-	-	-	-
Net reclassification	0 (0.0)	0 (0.0)	2 (5.1)	6 (40.0)	-	-	-	-
Sensitivity or specificity	-	-	-	1 (6.7)	-	-	-	-
Receiver operating curve	-	-	-	-	-	23 (29.5)	-	-
Clinical utility	-	-	-	-	-	1 (1.3)	-	-
Other measures	2 (4.8)	2 (12.5)	31 (79.5)	5 (33.3)	-	-	-	-

Values are the number of studies (%). [a] Derivation studies. [b] Primarily derivation studies, some of which included validation studies. [c] Validation studies.

1.3.5. Research inefficiency in the impact assessment and adoption stage

For the vast majority of CPRs, no impact study has been conducted (47, 53, 90). Systematic reviews of impact studies have shown that many impact studies have a substantial risk of bias (91-93). However, no study has directly evaluated research inefficiency in the impact assessment and adoption stage of CPR development.

1.4. Aim and scope of the thesis

1.4.1. Aim of the thesis

I aim to describe research waste and inefficiency in each stage of CPR development and explore some of the potential causes for research waste and inefficiency and their implications to CPR research and clinical practice.

1.4.2. Scope of the thesis

At the initial research proposal stage of the thesis, I planned to study CPRs in all clinical areas. As I began developing protocols for individual research projects, I decided to limit the scope of the thesis to one clinical area because it became clear that including CPRs in all clinical areas was less practical. I chose

to focus on cardiovascular CPRs for the following reasons. Firstly, there was a relative abundance of CPRs for cardiovascular conditions than other clinical areas (47). Therefore, it was more likely for me to find sufficient research materials to work with and be able to answer my research questions by focusing on cardiovascular CPRs compared with CPRs in other clinical areas. Secondly, cardiovascular CPRs were more commonly recognised and used by clinicians than CPRs in other clinical areas (3), thus focusing on cardiovascular CPRs would likely produce findings that have broader relevance to clinical practice.

I also concentrated on CPRs used in primary care. I am a General Internist practising full time ambulatory primary care medicine for adult patients in the USA, and it was logical for me to work with CPRs relevant to my clinical practice. Also, the thesis was conducted at the Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, where available expertise and resources were related to the primary care setting.

1.5. Structure of the thesis

I structured this thesis in accordance with the stages of CPR development because key research waste and inefficiency are closely related to having a multi-stage process, and each stage of CPR development contains unique challenges. There are six chapters with original research projects. These six chapters address specific questions about the waste and inefficiency of cardiovascular CPR research in each stage of CPR development: derivation (chapter 2), external validation (chapters 3 and 4), impact assessment (chapter

5), and adoption (chapters 6 and 7). I briefly summarise aims, objectives, and methods for each chapter with the original research project.

1.5.1. Chapter 2. Justification for deriving a new cardiovascular CPR

In chapter 2, I examine whether authors of a new cardiovascular CPR assess existing evidence and justify the need for deriving the new CPR. To achieve this, firstly, I review the introduction section of derivation studies to evaluate whether the authors cite an existing CPR when justifying deriving a new cardiovascular CPR. Secondly, I explore why authors derive new cardiovascular CPRs when other CPRs for the same problem exist by conducting a thematic content analysis. Lastly, I survey the authors of new cardiovascular CPRs to understand why some authors cite existing cardiovascular CPRs and others do not.

1.5.2. Chapter 3. Independent external validation of cardiovascular risk CPRs

In chapter 3, I explore why some cardiovascular CPRs are externally validated by independent researchers when most others are not. I use survival analysis to estimate the probability for a cardiovascular CPR to have an independent external validation study after the CPR is derived. I also assess whether features of derivation, reporting, and publication of cardiovascular CPRs are associated with having an independent external validation study using Cox proportional hazards regression analyses.

1.5.3. Chapter 4. Research inefficiency in external validation studies of Framingham CHD risk rule

In chapter 4, I quantify research inefficiency among studies externally validating Framingham Wilson CHD risk rule. Firstly, I assess whether authors of external validation studies update Framingham Wilson CHD risk rule when the CPR performs poorly. Then, I evaluate whether external validation studies make a significant contribution to understanding the performance of the Framingham Wilson CHD risk rule.

1.5.4. Chapter 5. Design, methods, and reporting of impact studies of cardiovascular CPRs

In chapter 5, I assess research waste and inefficiency in impact studies of cardiovascular CPRs. I systematically review the design, method, and reporting of impact studies of cardiovascular CPRs. I also compare the quality of methods and reporting between impact studies of cardiovascular CPRs with appropriate study design and matched control studies evaluating other types of nonpharmacologic interventions.

1.5.5. Chapter 6. Familiarity and use of cardiovascular CPRs

In chapter 6, I evaluate which cardiovascular CPRs are recognised and used in primary care. I conduct a survey to understand which cardiovascular CPRs are

recognised and used by general practitioners (GPs) in the UK and to assess how the GPs' familiarity and use of cardiovascular CPRs have evolved by comparing the results of the survey with the results of the 2014 study by Plüddemann et al. (3).

1.5.6. Chapter 7. Influence of research evidence on the uptake of cardiovascular CPRs

In chapter 7, I explore how research evidence influences the use of cardiovascular CPRs in primary care. I conduct qualitative interviews of primary care providers in the USA and explore how they perceive and experience cardiovascular CPR research, and how their perceptions and experiences influence the use of cardiovascular CPRs in practice.

Chapter 2. Justifications for deriving a new cardiovascular clinical prediction rule

A manuscript based on this chapter was published in PLoS One:

> 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. (94)

2.1. Abstract

Background: Many cardiovascular clinical prediction rule (CPRs) have been derived, often for the same problems. Although the need for any new research should be determined by systematically examining existing evidence, it is unknown whether researchers evaluate existing CPRs to justify deriving a new cardiovascular CPR. In this chapter, I assess whether authors of cardiovascular CPRs cite existing CPRs, why authors think existing CPRs are insufficient, and why some authors do not cite existing CPRs.

Methods: Derivation studies of cardiovascular CPRs included in the International Register of Clinical Prediction Rules for Primary Care were evaluated. I reviewed the derivation studies' introduction section to determine whether the authors cited existing CPRs when justifying a new cardiovascular CPR. Using thematic content analysis, I explored the stated reasons for determining existing cardiovascular CPRs were insufficient. I surveyed the

authors of derivation studies via e-mail and post and asked them whether they were aware of any existing cardiovascular CPRs at the time of derivation, how they searched for existing CPRs, and whether they thought it was important to cite existing CPRs. I assessed the association between authors' citations of an existing CPR and their reported awareness of existing CPRs, search for existing CPRs, or importance of citing existing CPRs.

Results: Of 85 derivation studies included, 48 (56.5%) cited existing CPR, 33 (38.8%) did not cite any existing CPR, and four (4.7%) declared there was no existing CPR to cite. Content analysis identified five categories of existing CPRs' insufficiency related to: (1) derivation (6 studies; 12.5%), (2) construct (32 studies; 66.7%), (3) performance (11 studies; 22.9%), (4) transferability (13 studies; 27.1%), and (5) evidence (9 studies; 18.8%). Authors of 54 derivation studies (71.1% of 76 authors contacted) responded to the survey. Twenty-five authors (46.3%) reported they were aware of an existing CPR at the time of derivation. Twenty-nine authors (53.7%) declared they conducted a systematic search to identify existing CPRs. Forty-nine authors (90.7%) indicated citing an existing CPR in derivation studies was important. The citation of an existing CPR was not clearly associated with the authors' reported awareness of existing CPRs, search for existing CPRs, or importance of citing existing CPRs.

Discussion: Cardiovascular CPRs are often derived without citing existing CPRs to justify the new CPRs' need, although most authors agree it is important to do so. The most common justification for a new CPR concerned existing CPRs'

construct, including the choice of predictor variables or relevance of outcomes, some of which could have been addressed by externally validating or updating existing CPRs. Researchers should clearly justify why new CPRs are needed with reference to existing CPRs to avoid unnecessary duplication.

2.2. Introduction

Conducting a derivation study is the first of four stages of CPR development (1, 7, 23). In a derivation study, candidate predictors for a diagnostic or prognostic outcome are identified, and a CPR is built typically using statistical methods such as logistic or Cox regression (24). The accuracy of the CPR is estimated in a population, and the “apparent accuracy” can be tested and adjusted using a part of the same population in an internal validation (28).

Many cardiovascular CPRs have been created (47, 51). However, the derivation of cardiovascular CPRs has been disproportionately focused on some clinical areas such as coronary artery disease and congestive heart failure (51). For example, a systematic review found there were 363 CPRs for assessing cardiovascular disease risk (53). Another systematic review identified 64 similar prognostic CPRs for congestive heart failure and 50 modifications of those CPRs that consisted of many overlapping predictors (95). Many of these cardiovascular CPRs have not been externally validated, and only a few have been assessed in an impact study (47, 96). Clinicians might be hesitant to use

these redundant CPRs for the same problem with uncertainties in their generalizability or effectiveness (49, 97).

Before any new research is undertaken, what is known about the subject should be systematically assessed to identify pertinent research questions, choose an optimal study design, and avoid unnecessary duplications (38, 44, 98). By the same token, researchers deriving a new CPR should clearly establish the need by examining existing CPRs for the clinical problem. The Transparent Reporting of a multivariable prediction model for individual Prognosis Or Diagnosis (TRIPOD) statement is a reporting guideline for CPR studies published in 2015, and it recommends that authors of a derivation study,

> "Explain the medical context (including whether diagnostic or prognostic) and rationale for developing … the multivariable prediction model, including references to existing models." (6)

Deriving CPRs without assessing what is already available might lead to many redundant CPRs for the same problems. In this chapter, I examine whether authors of derivation studies justify the need for creating a new cardiovascular CPR by evaluating existing evidence.

2.2.1. Objectives

Three specific objectives of this chapter are:

(1) to review whether authors of derivation studies cite existing CPRs when justifying a new cardiovascular CPR,

(2) to explore why authors derive a new CPR when CPRs for the same problem exist, and

(3) to assess why some authors cite existing cardiovascular CPRs and others do not.

2.3. Methods

2.3.1. Information source

I used the International Register of Clinical Prediction Rules for Primary Care (47), which will be referred to “the International Register” from now on, as the source of cardiovascular CPRs evaluated in this chapter. I chose the International Register because it was the most comprehensive and accessible source of cardiovascular CPRs pertinent to primary care at the time. Although Tufts Predictive Analytics and Comparative Effectiveness Clinical Prediction Model Database had a larger collection of cardiovascular CPRs (51), I did not use it as the information source because it was not readily accessible and included CPRs targeted for specialists such as cardiovascular surgeons.

The International Register was published in 2014 to promote the integration of CPRs in primary care practice and facilitate future CPR research (47), which contained 745 articles deriving, validating, and assessing the impact of 434

CPRs categorised according to the clinical domains of the International Classification of Primary Care, Second Edition (99). The CPRs included in the International Register were identified by reviewing 73,305 references collected from searching Medline for CPR literature published in 30 journals between 1980 and 2009 using a unique search strategy for CPRs relevant to primary care, searching secondary sources, and contacting experts in CPR research (47).

Because the International Register included studies published before 2010, I considered updating the search to collect contemporary studies using the same methods used by the authors of the International. However, at the protocol development stage, it was discovered that the number of references that would be retrieved and need to be reviewed was substantial due to the exponential growth of prediction rule literature. Therefore, I determined it was impractical to update the search using the same methods given limited time and resources to complete the project as a chapter of a doctorate thesis. I also considered collecting more recently published CPRs using a less comprehensive search strategy. Although the project using the simpler search would have been more attainable, I chose not to use this approach because the results would have been difficult to interpret.

2.3.2. Selection of derivation studies

I reviewed all references for cardiovascular CPRs found in the International Register and identified eligible derivation studies. For a derivation study to be

eligible, authors must have derived a de novo CPR for a cardiovascular outcome by selecting the CPR's predictor variables from candidate variables (e.g. backward elimination or full model approach) and constructing the final multivariable model with those variables using a model-building method (e.g. logistic or Cox proportional hazards regression) (24, 27, 100). The eligibility was not restricted to any derivation year, population, setting or language.

A derivation study was excluded if:

(1) it updated an existing cardiovascular CPR using methods such as re-calibration, revision or extension (2, 32, 101),

- re-calibration: the intercept and coefficients of an existing CPR were adjusted using the calibration intercept and slope
- re-estimation: the intercept and coefficients of an existing CPR were recalculated
- modification: some predictors were added to, removed from, or replaced in an existing CPR

(2) it presented a cardiovascular CPR without describing how it was derived (e.g. guideline), or

(3) it lacked an introduction section (e.g. conference abstracts).

I excluded updating studies because authors built on an existing CPR rather than deriving an entirely new CPR. Studies without an introduction section were

also excluded because I evaluated the introduction section of derivation studies to identify authors' justifications for deriving a new CPR and citations to existing CPRs.

2.3.3. Review of derivation studies

I reviewed the introduction section of derivation studies that met selection criteria and examined whether authors cited an existing CPR when they justified why it was necessary to derive a new CPR. I considered that authors of a derivation study cited an existing CPR when the introduction section included a sentence discussing an existing CPR with a relevant bibliographic reference. Rather than a citation for a systematic review of CPRs, a citation for an existing CPR was examined because it was consistent with what is recommended by the TRIPOD statement. In addition, systematic reviews of CPRs were much less commonly done before 2010 compared with systematic reviews of intervention or diagnostic test accuracy studies. Therefore, it was unrealistic to expect for many authors of derivation studies to have cited or conducted a systematic review of existing CPRs.

For derivation studies that did not cite any existing CPR, I reviewed the introduction section and assessed whether authors justified deriving a new cardiovascular CPR by declaring there was no existing CPR that addressed the cardiovascular problem. Then, derivation studies were grouped into ones that "cited existing CPR", "did not cite any existing CPR", and declared there was "no CPR to cite."

I also extracted the following characteristics from each eligible derivation study:

(1) bibliographic data (authors and publication year),

(2) the country where the CPR was derived,

(3) the type of CPR (diagnostic or prognostic), and

(4) the clinical domain of the International Register that the CPR belonged to.

Another reviewer (Emma Wallace) independently determined whether authors of derivation studies cited existing CPRs when justifying a new cardiovascular CPR and extracted all additional data. We compared data and adjudicated disagreements through discussions.

2.3.4. Thematic content analysis

For derivation studies that cited an existing CPR, I examined authors' justifications for deriving a new cardiovascular CPR despite the fact that CPRs for the same purpose already existed. I used a thematic content analysis (102) to produce a typology of reasons why an existing CPR was insufficient to address the problem stated by authors of these derivation studies. I focused on manifest content (what texts describe) rather than latent content (the underlying meaning of texts) (103, 104) because it was impossible to reliably interpret the implicit underlying meaning of the texts by reading authors' statements. Since

there was no prior knowledge about why new cardiovascular CPRs are created, I used an inductive approach to develop a coding scheme by allowing themes to emerge from the examination of data (105).

The introduction section of the derivation studies that cited existing cardiovascular CPRs were imported to NVivo for Mac (QSR International Pty Ltd). I read the introduction section of each derivation study repeatedly until I was familiar with data. Then, I identified sentences that contained any reason why an existing CPR was insufficient to address the problem. A preliminary code was created using key words from each sentence. After creating preliminary codes for 30% of randomly sampled derivation studies, a coding scheme was generated by combining preliminary codes with an overlapping theme. The remaining derivation studies were coded according to the coding scheme. A new code was generated when there was content that described an insufficiency of existing cardiovascular CPRs but was not compatible with any existing codes. Previously coded derivation studies were re-coded using the updated coding scheme. Data within each code were reviewed to confirm that they were consistent with the assigned code. The code was divided when inconsistencies in data were found. Similar codes were organised into categories.

I summarized the content analysis results by presenting each category with data extracts from derivation studies to illustrate the essence of the category

(105). I also reported the number of derivation studies that provided each category of reasons why existing cardiovascular CPRs were insufficient.

2.3.5. Survey of authors

I conducted a quantitative survey of authors to explore why some authors cited existing cardiovascular CPRs in their derivation studies while others did not. An online survey supplemented by postal survey enabled me to quickly and cost-effectively reach out to many authors. Although a deeper understanding about this research question might have been gained by collecting rich data from an in-depth interview or focus group of derivation study authors, I chose not to use these qualitative methods because of time and funding constraints.

Primary target participants for the survey were the corresponding authors of included derivation studies because these authors were likely to have been intimately involved in conceptualizing, conducting, and writing up their derivation studies. I also contacted the first author of a derivation study when the corresponding author, who was not the first author, could not be reached.

I identified the target participants' e-mail and postal addresses by reviewing the derivation studies and searching the authors in PubMed and Google Scholar. In November 2015, I sent a personalized e-mail with a link to an online questionnaire and a participant invitation letter to each target participant. A second e-mail was sent if the author did not respond within eight weeks. For

authors whose e-mail address was not known or who did not respond to a third e-mail invitation, I sent a personalized survey packet by post. The survey packet included a participant invitation letter, a questionnaire, and a stamped return envelope. The participant invitation letter outlined the purpose of the survey, why the author was chosen, the length of the survey, implied consent by completing, the confidentiality of data collected, and how to raise any concern. The participant invitation letter is provided in Appendix 2.1. Due to financial constraints, I could only send the survey packets to authors in the USA, Canada, and the UK.

I developed the survey questionnaire by considering potential reasons why authors of derivation studies might not have cited existing cardiovascular CPRs. These include:

(1) authors were not aware of any existing cardiovascular CPR,

(2) authors did not search for existing cardiovascular CPRs, and

(3) authors did not believe citing exiting cardiovascular CPR was important.

The survey questionnaire was piloted among academic GPs from the Nuffield Department of Primary Care Health Sciences, University of Oxford, who had experiences in deriving CPRs. The survey questionnaire was updated after receiving feedback from the academic GPs who participated in the pilot. The final version of the survey consisted of three closed-ended questions, which can be found in Appendix 2.2.

The Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) of the University of Oxford conducted an ethical review and approved the survey (reference number: MS-IDREC-C1-2015-161). Participants were asked to confirm that they had read and understood the information about the survey by checking a box. To prevent multiple submissions from the same participant, each author was required to enter a unique alphanumeric code before starting the survey provided in the participant invitation letter. No incentive was offered for participating.

2.3.6. Statistical analysis

The characteristics of derivation studies are presented with numbers and proportions for categorical variables, and medians and interquartile ranges (IQRs) for non-normally distributed continuous variables. The answers to survey questions are reported with numbers and proportions. I used Fisher's exact test to assess the null hypothesis that answers to survey questions are same between derivation studies that cited existing CPR, did not cite any existing CPR, and declared there was no CPR to cite. Stata (Release 14. College Station, TX: StataCorp LP) was used for all statistical analyses.

2.4. Results

2.4.1. Review of derivation studies

There were 131 derivation studies of 138 cardiovascular CPRs across 17 cardiovascular clinical domains in the International Register. From these, 46 derivation studies were excluded: 35 that updated existing cardiovascular CPRs, ten practice guidelines and narrative reviews that presented a CPR without describing how it was derived, and one conference abstract that did not have an introduction section.

There were 85 derivation studies that met the selection criteria. Of these derivation studies, 48 (56.5%) cited at least one existing CPR, 33 (38.8%) did not cite any existing CPR, and 4 (4.7%) declared there was no existing CPR to cite. Characteristics of these included derivation studies are presented in Table 2.41.

Table 2.4.1 Characteristics of derivation studies according to the citation of existing cardiovascular clinical prediction rule (CPR).

Characteristics	Cited existing CPR	Did not cite existing CPR	No existing CPR to cite	All derivation studies
Publication year				
Before 1980	1 (2.1)	0 (0.0)	0 (0.0)	1 (1.2)
1980 - 1989	6 (12.5)	6 (18.2)	1 (25.0)	13 (15.3)
1990 - 1999	11 (22.9)	11 (33.3)	0 (0.0)	22 (25.9)
2000 - 2009	30 (62.5)	16 (48.5)	3 (75.0)	49 (57.6)
Location				
USA	20 (41.7)	15 (45.5)	3 (75.0)	38 (44.7)
UK	7 (14.6)	5 (15.2)	0 (0.0)	12 (14.1)
Europe	6 (12.5)	5 (15.2)	1 (25.0)	12 (14.1)
Canada	4 (8.3)	1 (3.0)	0 (0.0)	5 (5.9)

Characteristics	Cited existing CPR	Did not cite existing CPR	No existing CPR to cite	All derivation studies
Multiple	7 (14.6)	5 (15.2)	0 (0.0)	12 (14.1)
Other	4 (8.3)	2 (6.1)	0 (0.0)	6 (7.1)
Type of CPR				
Diagnostic	22 (45.8)	19 (57.6)	1 (25.0)	42 (49.4)
Prognostic	26 (54.2)	14 (42.4)	3 (75.0)	43 (50.6)
Clinical domain				
Stroke/cerebrovascular accident	8 (16.7)	5 (15.2)	1 (25.0)	14 (16.5)
Risk factor cardiovascular disease	8 (16.7)	3 (9.1)	0 (0.0)	11 (12.9)
Heart failure	2 (4.2)	7 (21.2)	0 (0.0)	9 (10.6)
Pulmonary embolism	6 (12.5)	2 (6.1)	1 (25.0)	9 (10.6)
Ischemic heart disease with angina	6 (12.5)	1 (3.0)	0 (0.0)	7 (8.2)
Acute myocardial infarction	6 (12.5)	1 (3.0)	0 (0.0)	7 (8.2)
Phlebitis/thrombophlebitis	4 (8.3)	2 (6.1)	0 (0.0)	6 (7.1)
Other	8 (16.7)	12 (36.4)	2 (50.0)	22 (25.9)
Total	48 (100.0)	33 (100.0)	4 (100.0)	85 (100.0)

Values are numbers (%) of the derivation study.

The median sample size and the median number of events were 1,253 (IQR, 526 to 5,711) and 213.5 (IQR, 64 to 462) for all included derivation studies: 1,274 (IQR, 533.5 to 5,591) and 217 (IQR, 62 to 509) for studies that cited existing CPR, 933 (IQR, 466 to 6,422) and 135 (IQR, 65 to 274) for studies that did not cite existing CPR, and 3,703 (IQR, 1,505.5 to 10,147.5) and 760 (IQR, 27 to 1,091) for studies that declared there was no CPR to cite.

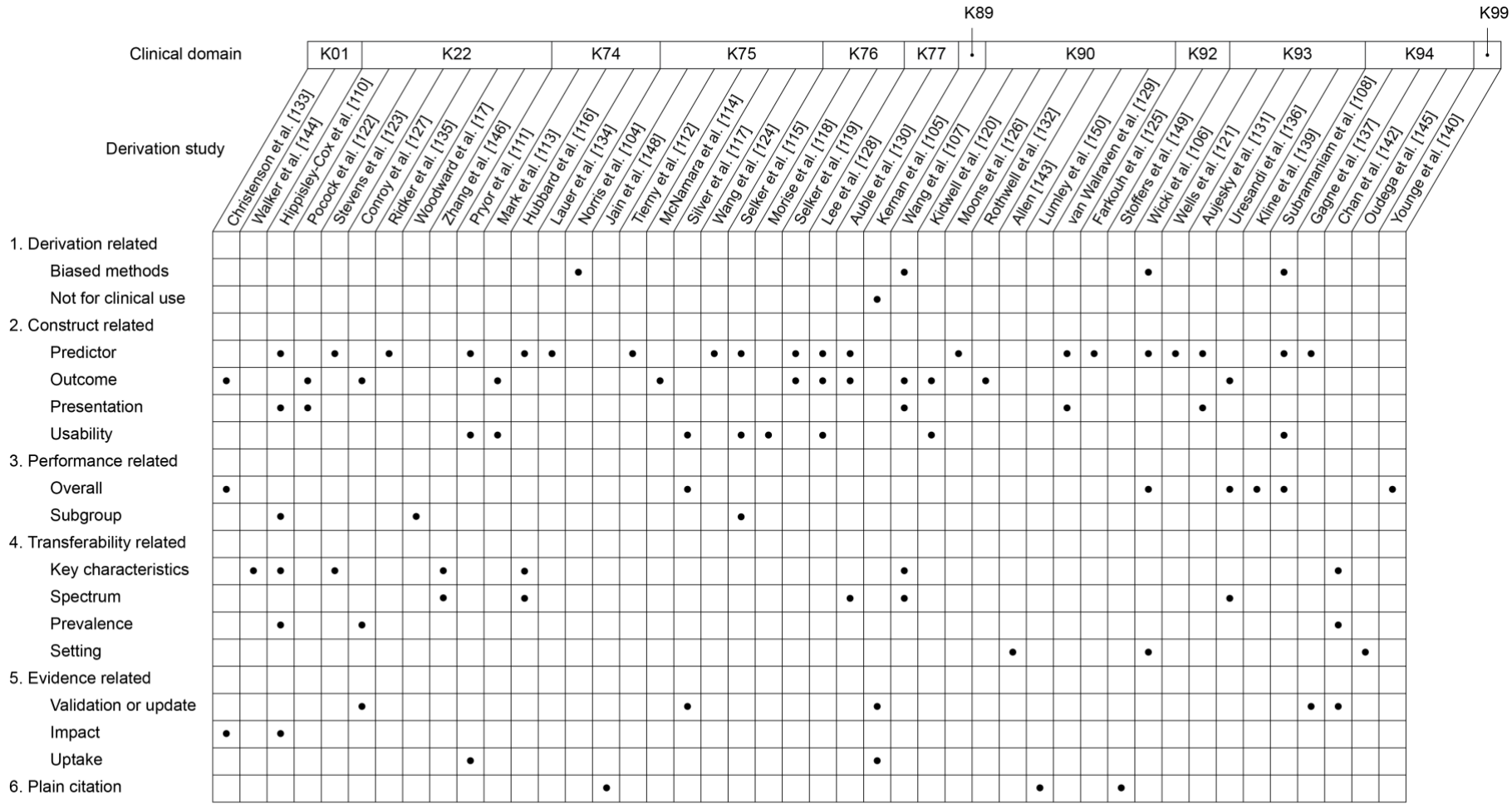
Studies conducted in the US and UK comprised 58.8% (50 of 85) of included derivation studies and contained 87.1% (2,184,447 of 2,508,631) of total participants. The median sample size and the median number of events were 1,735.5 (IQR, 540 to 5,711) and 291 (IQR, 64 to 766) for studies from the US,

2,896.5 (IQR, 207 to 10,516) and 78.5 (IQR, 50 to 1,165) for studies from the UK, 756 (IQR, 340.5 to 1,875) and 251 (IQR, 93 to 301) for studies from Europe, 769 (IQR, 662 to 1,239) and 77 (IQR, 17 to 217) for studies from Canada, 3,368.5 (IQR, 672 to 7,422.5) and 251 (IQR, 93 to 301) for studies from multiple countries, and 641.5 (IQR, 176 to 4,400) and 157 (IQR, 35 to 213) for studies from other areas.

Twelve studies presented more than one version of a CPR with varying numbers of predictor variables. The median number of predictor variables included in cardiovascular CPRs was 6 (IQR of 4 to 8 when counting the smallest final models and 4.5 to 9 when calculating the largest final models).

2.4.2. Content analysis

In the thematic content analysis of 48 derivation studies of cardiovascular CPRs that cited existing CPR, I identified five categories of reasons why an existing CPR was insufficient to address the problem stated by the authors of derivation studies. These categories of existing CPRs' insufficiency described by authors in each derivation study are presented in Figure 2.4.1.



Clinical domains according to the International Classification of Primary Care, Second Edition: K01 = Heart pain, K22 = Risk factor cardiovascular disease, K74 = Ischaemic heart disease with angina, K75: Acute myocardial infarction, K76 = Ischaemic heart disease without angina, K77 = Heart failure, K89 = Transient cerebral ischemia, K90 = Stroke/cerebrovasular accident, K92 = Atherosclerosis/PVD, K93 = Pulmonary embolism, K94 = Phlebitis/thrombophlebitis, and K99 = Cardiovascular disease other.

Figure 2.4.1 Insufficiencies of existing cardiovascular clinical prediction rules stated by authors in each derivation study.

(1) Derivation related insufficiency

Authors of six derivation studies (12.5%) stated that existing cardiovascular CPRs had insufficiencies related to their derivation process (106-111). Authors of four studies noted that the derivations were conducted using biased methods (e.g. selection bias) (106, 108-110) and authors of two studies stated that the existing CPRs were created without the intention of clinical use (107, 111). For example, Wang et al. (109) reported that the methods used to derive one of the prognostic rules for atrial fibrillation (AF) was flawed:

> "They tested this scheme using Medicare claims data from patients who were hospitalized for AF but did not receive anticoagulation therapy. A potential limitation of this approach is selection bias, because clinical features associated with nonuse of warfarin or hospitalization for AF are likely to influence stroke risk. Also, some strokes may be missed by hospital discharge data if they are small, immediately lethal, or improperly coded."

(2) Construct related insufficiency.

Authors of 32 derivation studies (66.7%) stated that existing CPRs had insufficiencies related to their construct (108-110, 112-140). There were 20 derivation studies (41.7%) that mentioned problems with predictor variables of existing CPRs. Some of these derivation studies discussed issues with predictor variables that might influence the accuracy of CPRs, such as the inaccuracy or dichotomization of predictor variables (114, 125, 137). Other derivation studies pointed out predictor variable related insufficiencies that might hinder the

implementation of CPRs, including predictors were missed (110, 112, 118, 121, 125, 127, 128, 136, 137), unavailable in practice (117, 126, 130, 132, 133, 139), clinically insensible (110, 117, 131), irreproducible (108, 113), not standardized (108), and not clearly defined (123). The following authors' quotes illustrate some of these issues related to predictor variables:

> "The Framingham algorithm does not include factors such as social deprivation, body mass index, family history of cardiovascular disease, and current treatment with antihypertensives." (112)

> "A more fundamental limitation is that the MMPS, and other currently available analogous mortality predictors, use clinical information from the first 24 hours or more not available upon presentation to the hospital, … they cannot be used immediately in the real-time clinical setting." (117)

Authors of 13 derivation studies argued that the existing CPRs' outcomes were irrelevant (109, 115, 116, 121, 122, 124, 129, 130, 132, 134, 135, 138, 140). One example is shown here:

> "Available risk scores do not include mortality as an end point although studies have indicated that AF is an independent risk factor for death as well as stroke and therapies for AF may affect mortality." (109)

Five derivation studies described issues with the presentation of CPRs (109, 112, 124, 131, 133). One study pointed out existing CPRs did not estimate the absolute risk of patients:

> "Although existing risk schemes provide absolute event rates for 3 to 7 levels of risk, it may be beneficial to have more specific estimates of risk." (109)

Eight studies discussed difficulties of using existing CPRs due to complex calculations, need for a calculator or computer, and large numbers of predictor variables (110, 113, 115, 117, 119, 120, 122, 130). One example is shown here:

> "Second, previous treadmill test scores have relied on either complex regression model calculations or microcomputer-based measurements that require the use of special equipment, which makes them difficult to apply in routine clinical practice." (115)

(3) Performance related insufficiency

Authors of 11 derivation studies (22.9%) stated that existing CPRs performed disappointingly (21, 108, 110, 112, 117, 119, 135, 138, 141-143). Three of these studies discussed the inadequate performance of existing CPRs in a subgroup of population (e.g. socially deprived) (21, 112, 117). Subramaniam et al. (110) justified deriving the Hamilton score by highlighting the inaccuracy of Wells score for Deep Vein Thrombosis (DVT):

> "The modified Wells score has limitations in discriminating patients likely to have DVT and those unlikely to have DVT. In a study involving 1,096 ambulatory outpatients … 495 patients (45.2%) were identified as likely to have DVT using the modified Wells score. This is despite the ambulatory population, which is expected to have a lower risk for DVT than hospital inpatients."

(4) Transferability related insufficiency

Authors of 13 derivation studies (27.1%) stated that existing CPRs had limitations in their transferability (108, 109, 112, 118, 125, 129, 132, 138, 144-

148). Seven derivation studies discussed transferability constraints due to differences in participants' key characteristics such as age, sex, and co-morbidities (109, 112, 118, 125, 144, 146, 148). Five studies reported that differences in disease spectrum might hinder CPR transferability (109, 118, 132, 138, 148). Some studies discussed the dissimilarity in disease prevalence (112, 129, 144) and settings (108, 145, 147) as a potential obstacle. Authors of the following derivation study discussed concern about using an existing CPR created in a different setting:

> "These studies have been made of selected groups of patients in rehabilitation units. The results therefore may not be applicable to general hospital stroke admissions." (145)

(5) Evidence related insufficiency

Authors of nine derivation studies (18.8%) mentioned the insufficient evidence of existing CPRs as the reason to create a new cardiovascular CPR (107, 112, 113, 119, 129, 135, 139, 144, 149). Authors of derivation studies frequently pointed out the lack of external validation of existing cardiovascular CPRs (107, 119, 129, 139, 144). Uncertainties regarding the impact (112, 135, 149) and lack of uptake of existing CPRs in clinical practice (107, 113, 149) were also mentioned. Lack of validation of existing CPRs was the reason for deriving a new CPR in the following example:

> "Although pregnancy is recognized as a risk factor for venous thrombosis, no prospective studies validate the use of current diagnostic strategies for DVT." (144)

(6) Plain citation

In addition to these five categories of existing CPRs' insufficiencies related to (1) derivation, (2) construct, (3) performance, (4) transferability and (5) evidence, some studies simply mentioned existing cardiovascular CPRs without discussing why they were insufficient. Authors of three derivation studies (6.3%) just cited existing CPRs without clearly explaining why these CPRs were insufficient, and a new CPR was needed (150-152). For example, authors of the derivation study in the following example just listed available prediction rules without stating why they needed to derive a new CPR:

> "Two stroke risk prediction models have been published: the Framingham Study model, and the Israeli Ischemic Heart Disease Project (IIHD) model." (152)

2.4.3. Survey of authors

Of the 85 derivation studies included in the review, I attempted to contact the corresponding or first authors of 76 derivation studies whose email or postal address was found, as summarised in Figure 2.4.2. Of those authors, the authors of 54 derivation studies responded (39 by e-mail and 15 by post) giving a response rate of 71.1%.

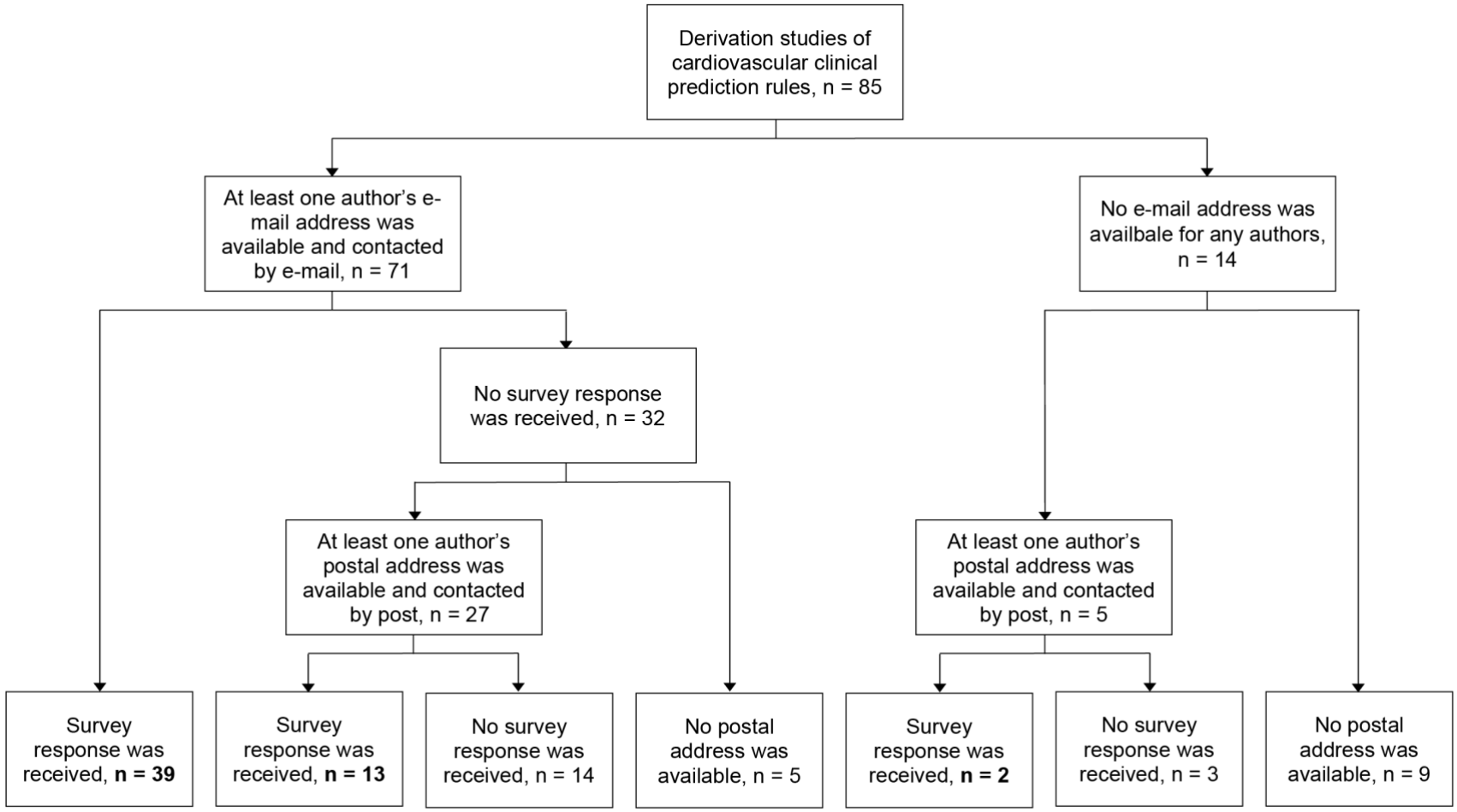


Figure 2.4.2 A flow diagram illustrating how the corresponding or first authors of derivation studies were contacted, and their responses were received.

The results of the author survey are presented in Table 2.4.2. The citation of existing CPRs was not clearly associated with whether authors were aware of existing CPRs (Fisher's exact $p = 0.11$) nor how they searched for existing CPRs (Fisher's exact $p = 0.19$). Although 90.7% (49 of 54) of authors indicated it was important to cite existing CPRs for the same problem when deriving a new CPR, their perceived importance was not associated with citing an existing CPR (Fisher's exact $p = 0.71$).

Table 2.4.2 Authors' answers to survey questions according to the citation of existing cardiovascular clinical prediction rule (CPR).

Question and response	Cited existing CPR	Did not cite existing CPR	No existing CPR to cite	All authors
1. *"At the time of derivation, were you aware of any existing prediction rules that addressed the same problem?"*				
Yes	19 (55.9)	6 (33.3)	0 (0.0)	25 (46.3)
No	14 (41.2)	12 (66.7)	2 (100.0)	28 (51.9)
No response	1 (2.9)	0 (0.0)	0 (0.0)	1 (1.9)
2. *"How did you become aware of existing prediction rules that addressed the same clinical problem?*				
Systematic review	18 (52.9)	9 (50.0)	2 (100.0)	29 (53.7)
No systematic review	7 (20.6)	8 (44.4)	0 (0.0)	15 (27.8)
No search	8 (23.5)	1 (5.6)	0 (0.0)	9 (16.7)
No response	1 (2.9)	0 (0.0)	0 (0.0)	1 (1.9)
3. *"How important do you think it is to cite existing prediction rules for the same problem when deriving a new prediction rule?"*				
Important	30 (88.2)	17 (94.4)	2 (100.0)	49 (90.7)
Unimportant	4 (11.8)	1 (5.6)	0 (0.0)	5 (9.3)
Total	34 (100.0)	18 (100.0)	2 (100.0)	54 (100.0)

Values are numbers (%) of derivation studies.

Characteristics of derivation studies that a response from an author was received, authors were contacted, but no response was received, and no author could be contacted are presented in Table 2.4.3.

Table 2.4.3 Characteristics of derivation studies of cardiovascular clinical prediction rule (CPR) that authors responded, did not respond, and could not be contacted.

Characteristics	Authors contacted by e-mail or post		No author could be contacted	All derivation studies
	Responded	Did not respond		
Publication year				
Before 1980	0 (0.0)	0 (0.0)	1 (11.1)	1 (1.2)
1980 - 1989	6 (11.1)	6 (27.3)	1 (11.1)	13 (15.3)
1990 - 1999	14 (25.9)	7 (31.8)	1 (11.1)	22 (25.9)
2000 - 2009	34 (63.0)	9 (40.9)	6 (66.7)	49 (57.6)
Location				
USA	22 (40.7)	14 (63.6)	2 (22.2)	38 (44.7)
UK	8 (14.8)	3 (13.6)	1 (11.1)	12 (14.1)
Europe	7 (13.0)	1 (4.5)	4 (44.4)	12 (14.1)
Canada	5 (9.3)	0 (0.0)	0 (0.0)	5 (5.9)
Multiple	10 (18.5)	1 (4.5)	1 (11.1)	12 (14.1)
Other	2 (3.7)	3 (13.6)	1 (11.1)	6 (7.1)
Type of CPR				
Diagnostic	24 (44.4)	15 (68.2)	3 (33.3)	42 (49.4)
Prognostic	30 (55.6)	7 (31.8)	6 (66.7)	43 (50.6)
Clinical domain				
Stroke/cerebrovascular accident	11 (20.4)	1 (4.5)	2 (22.2)	14 (16.5)
Risk factor cardiovascular disease	8 (14.8)	2 (9.1)	1 (11.1)	11 (12.9)
Heart failure	5 (9.3)	4 (18.2)	0 (0.0)	9 (10.6)
Pulmonary embolism	6 (11.1)	2 (9.1)	1 (11.1)	9 (10.6)
Ischemic heart disease with angina	4 (7.4)	2 (9.1)	1 (11.1)	7 (8.2)
Acute myocardial infarction	1 (1.9)	5 (22.7)	1 (11.1)	7 (8.2)
Phlebitis/thrombophlebitis	5 (9.3)	1 (4.5)	0 (0.0)	6 (7.1)
Other	14 (25.9)	5 (22.7)	3 (33.3)	22 (25.9)
Citation of existing CPR				
Cited existing CPR	34 (63.0)	9 (40.9)	5 (55.6)	48 (56.5)
Did not cite existing CPR	18 (33.3)	12 (54.5)	3 (33.3)	33 (38.8)
No existing CPR to cite	2 (3.7)	1 (4.5)	1 (11.1)	4 (4.7)
Total	54 (100.0)	22 (100.0)	9 (100.0)	85 (100.0)

Values are numbers (%) of derivation study.

2.5. Discussion

In this chapter, I reviewed whether authors cited existing CPRs when justifying a new cardiovascular CPR, explored authors' justifications for creating a new cardiovascular CPR when CPRs for the same problem exist, and evaluated why some authors cited existing CPRs and others did not.

2.5.1. Summary of findings

I found that only slightly more than half the authors of derivation studies cited at least one existing CPR and justified deriving a new cardiovascular CPR. Authors justified creating a new cardiovascular CPR by pointing out existing CPRs' insufficiencies related to derivation, construct, performance, transferability, and evidence. The most commonly cited category of insufficiency (70.4%) was about the existing CPRs' construct, such as issues in the choice of predictor variables, the relevance of outcomes, clinical usability, or presentation format. The second most common category (29.5%) was the insufficiency of existing CPRs' transferability due to the difference in populations or care settings.

The results of the author survey did not clearly explain the reasons why some authors cited existing CPRs in their derivation studies, and others did not;

neither being aware of existing CPRs, having conducted a systematic review of existing CPRs, nor believing it is important to cite existing CPRs was associated with citing existing cardiovascular CPRs. Given that the vast majority of authors indicated citing existing cardiovascular CPRs in derivation studies was important, it is difficult to defend the low citation rate found in this study.

2.5.2. Comparison with existing literature

Although the overall rate of the citation of existing CPRs was suboptimal in derivation studies of cardiovascular CPRs, the citation of existing CPRs seems to have improved over time, especially for derivation studies published after 2000. A cross-sectional study of ten journals with the highest impact factor in various clinical domains by Heus et al. (65) found that 74.0% of derivation studies published in 2012 adhered to the TRIPOD item 3a “explain the medical context and rationale for developing the multivariable prediction model including reference to existing models” (6) which is higher than the rate of the citation of existing CPRs in this chapter. Further, Zamanipoor Najafabadi et al. (153) evaluated the quality of reporting and methods between CPR studies published in seven general medicine journals with the highest impact factor before (2012-2014) and after (2016-2017) the publication of the TRIPOD statement. Authors found that 78% of studies published before (2012-2014) the TRIPOD statement adhered to the TRIPOD item 3a, whereas 84% of the post-TRIPOD studies (2016-2017) adhered to the item 3a (153). Although these studies were collected from the general medicine journals with the highest impact factor,

findings suggest that authors might be justifying a new prediction rule with a citation of existing prediction rules more often.

Ideally, the need for any new research should be identified by systematically reviewing existing evidence (44, 154) and not just with citing previous studies selected by authors. However, studies have repeatedly shown that systematic reviews are infrequently used to justify and design a new randomised controlled trial (RCT) (154-158). For example, a cross-sectional study found that only 23.4% (30 of 128) of RCTs published in high impact Orthopaedic trauma journals between 2015 and 2018 justified the need for a new RCT with citation to a systematic review (156). Another cross-sectional study showed a systematic review was used to inform the trial design in 40.6% (41 of 101) of RCT protocols published in December of 2015 (155). Only 3.5% (3 of 85) of derivation studies evaluated in this chapter cited a systematic review of existing cardiovascular CPRs (112, 120, 127), which is even substantially less than the citation rate of systematic reviews of RCTs mentioned above.

Four derivation studies by Aujesky et al. (159), Panzer et al. (160), Schnabel et al. (161), and Feringa et al. (162) declared that there was no clinical prediction rule to cite in this chapter. Aujesky et al. (159), in 2006, derived a prognostic rule for the risk of 30-day mortality in patients with pulmonary embolism because “no explicit clinical criteria exist to accurately identify patients with pulmonary embolism who are at low risk of adverse outcomes.” However, in 2005, the same authors published a prediction rule for the identical outcome

with similar predictor variables (133). Both prediction rules were found in three systematic reviews (163-165).

In 1985, Panzer et al. (160) derived a diagnostic rule for hemorrhagic stroke using data from a hospital in the USA. Although the authors stated that “a method for making such a prediction using basic clinical information has not been reported” (160), according to systematic reviews (166, 167), a diagnostic rule for hemorrhagic stroke from the UK published in 1983 by Allen et al. (168) was available.

Schnabel et al. (161) published a risk prediction rule for incident atrial fibrillation in 2009. This CPR was identified as the first of its kind in two systematic reviews (169, 170). Lastly, for the prognostic rule for long-term mortality of patients with peripheral artery disease by Feringa et al. (162), I did not find any predating prediction rule or systematic review summarizing CPRs for this outcome.

2.5.3. Strengths and Limitations

By using mixed methods, I gained in-depth understandings about why researchers might keep deriving new cardiovascular CPRs despite similar CPRs for the same problem already exist. I was able to augment the findings from the review of derivation studies that many authors do not cite any existing CPR when justifying a new cardiovascular CPR, further by thematic content analysis and author survey that revealed new cardiovascular CPRs are poorly

justified, and authors' beliefs about the importance of citing existing CPRs contradict their practice.

I used the International Register as the information source. Although the International Register was published in 2014, it only included CPRs published before 2010 (47). Also, I assessed only CPRs from the cardiovascular domain of the International Register. Therefore, results from this study may not be generalizable to the CPRs in other clinical areas or more recently derived CPRs. Furthermore, it is important to consider that all derivation studies in our review predate the publication of the TRIPOD statement (6), and many of them predate the modern expectation that a systematic review of existing evidence should be done before any new research (44, 154, 158). In fact, findings from the study by Zamanipoor Najafabadi et al. (153) suggest that authors might have been citing existing prediction rules in derivation studies more often since the publication of the TRIPOD statement. Therefore, it is only fair to interpret the findings from this chapter as a potential mechanism for having many redundant CPRs rather than an assessment of research conduct against current standards. Potentially, this limitation could have been prevented if a systematic review by Damen et al. (53) was used as the information source instead of the International Register. However, the protocol for this chapter was developed in 2014 and data collection and analysis, including the survey of authors (November 2015), were completed before the publication of the systematic review by Damen et al. (53).

The small number of the authors that I could send the survey might have made it difficult to detect the associations between the citation of existing CPR and the survey answers. The authors who cited existing CPR were slightly over-represented in the survey, which could have influenced the results. I built the survey using only three short key questions to maximize the response rate. This approach prevented me from including multiple open-ended questions in the survey questionnaire that might have generated richer survey data and led to a more in-depth analysis of the issue. Recall bias should also be considered when interpreting the survey results because two questions relied on authors' recollection of information from past CPR derivations.

Some authors indicated in their survey responses that they were unaware of an existing CPR at the time of derivation, which was at odds with the review of their derivation studies that showed the citation of existing CPRs. These discrepancies might have arisen from the different interpretations of the term "same clinical problem" in the first survey question. Some of these authors acknowledged there were existing CPRs for the problems in introduction sections but indicated existing CPRs were inapplicable to their populations or practice settings. The difference in settings or populations where existing CPRs were derived might have led some authors to determine these CPRs did not address the "same clinical problem".

One of the construct related insufficiencies identified in the content analysis, from studies published between 1983 and 2006, was difficulties in using CPRs

(e.g. requiring a special instrument). Although this insufficiency might be less important in current clinical practices where prediction rules are integrated into electronic health records (171), it might be relevant to models built using machine learning where complex codes are required because codes for machine learning models are seldom shared (172, 173). This makes it difficult to implement machine learning models in practice and impossible to validate or assess their impact (174). One of the potential solutions is to make it mandatory for researchers to share codes for machine learning models by journals and funding agencies (173). Nevertheless, existing guidelines for machine learning models already recommend sharing codes (175-180).

2.5.4. Research and clinical implications

Researchers might be inclined to create a new CPR even when CPRs for the same clinical problem already exist for the following reasons. Firstly, it has become easy to access data needed to derive a CPR due to the widespread use of electronic health records (EHRs) (181). Secondly, the process of deriving a CPR using statistical software is relatively straightforward (e.g. logistic regression), whereas researchers might be less familiar with methods needed to update an existing CPR (e.g. re-calibration). Lastly, it might be speculated from the tendencies for derivation studies to be published in journals with a higher impact factor compared with their independent external validation studies (96) that researchers and journals view deriving a new CPR as more novel work.

Researchers interested in deriving a new prediction rule should consider the following three steps recommended in clinical prediction rule literature (2, 6, 23, 74, 182, 183), which might help reduce the risk of deriving redundant prediction rules, as illustrated in Figure 2.5.1.

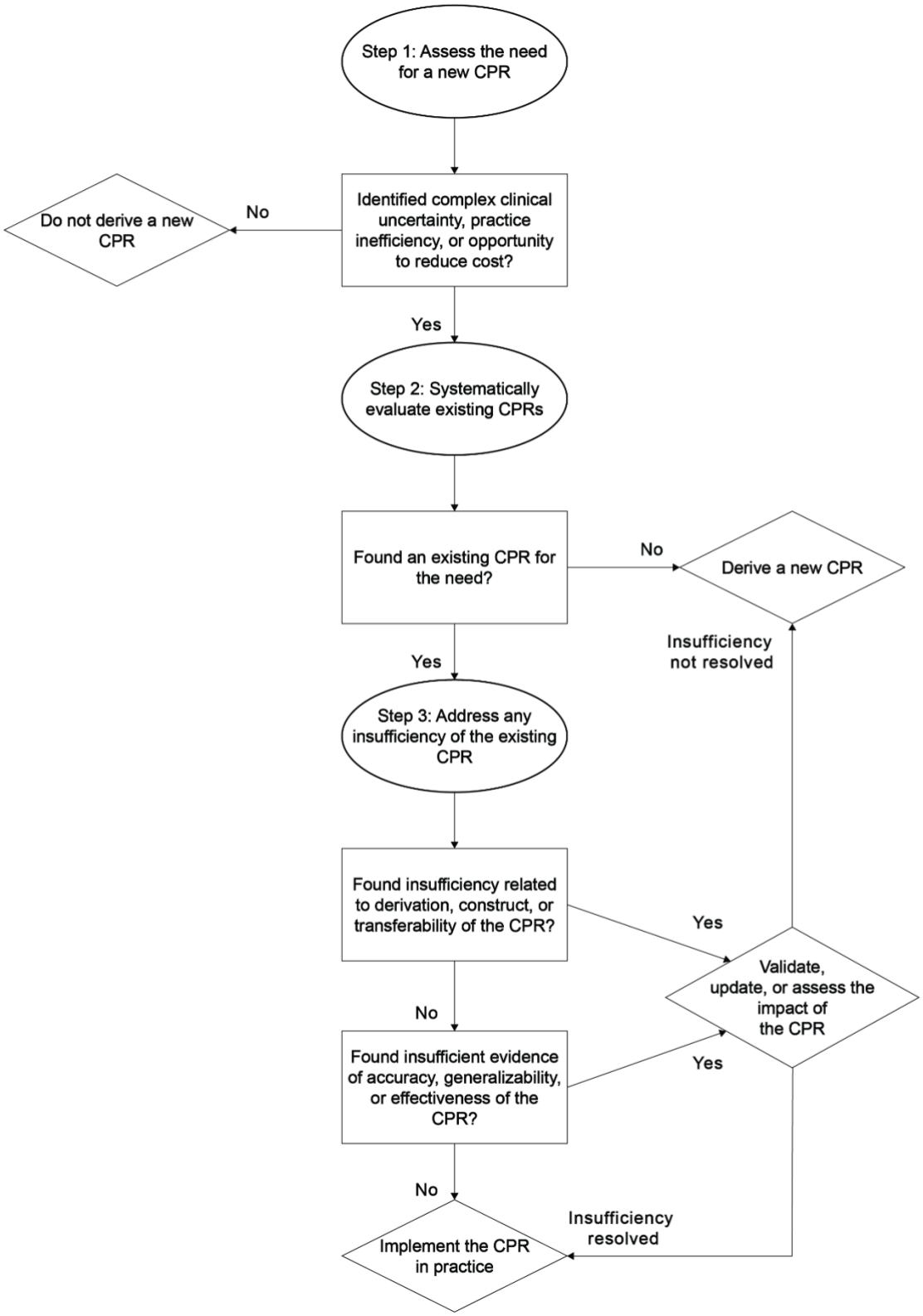


Figure 2.5.1 A framework to reduce the derivation of redundant clinical prediction rules (CPRs) and increase the value of CPR research.

(1) Researchers should establish the clinical context and need for a CPR a priori (182, 183). CPRs might help deal with complex clinical problems. For example, when deciding to start anticoagulation therapy, which is a potentially life-saving treatment for patients with atrial fibrillation but also has the risk of serious adverse effects, the CHA2DS2-VASc score (184) and HAS-BLED score (185) help clinicians and patients better understand the benefits and risks. In addition, a CPR can be useful when there is an opportunity to improve practice efficiency or reduce cost. For example, Reily et al. (186) have shown that, by using a modified version of the prediction rule by Goldman et al. (187), the efficiency of evaluating patients suspected of having acute cardiac ischemia can be improved by triaging low-risk patients to unmonitored or observation unit rather than inpatients cardiac care beds without affecting the safety of care.

(2) Once the need for a CPR is established, researchers should evaluate existing CPRs relevant to the clinical context (6, 74). This can be done by finding an up-to-date systematic review or updating an existing systematic review. When there is no existing systematic review, researchers should systematically search for and evaluate all relevant existing CPRs (188). Deriving a new CPR is well justified when no CPR addressing the need is identified.

(3) When an existing CPR is found in the previous step, researchers should examine the existing CPR for any insufficiencies and attempt to resolve

them by externally validating, updating, or assessing the impact of the existing CPR (2). For example, when an existing CPR is missing an important predictor variable, the CPR may be updated by adding the predictor variable and re-estimating the regression coefficients and intercept (2, 101, 189). When some insufficiencies fail to resolve, deriving a new CPR can be justified. When the existing CPR's insufficiencies are successfully resolved and there is sufficient evidence of accuracy, generalizability, and effectiveness, the CPR can be confidently implemented in practice (23).

The CONSORT (Consolidated Standards of Reporting Trials) 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials recommends "including a reference to a systematic review of previous similar trials or the absence of such trials" when presenting scientific background and explanation of rationale so that unnecessary trials can be avoided (190). Currently, the TRIPOD statement's recommendation goes only as far as "authors present a rationale for developing a new CPR with references to existing CPRs" (6). Elevating the recommendation for authors to include a systematic review of existing CPRs or the absence of any existing CPR when presenting a rationale for deriving a new CPR might discourage authors from creating redundant CPRs. Nevertheless, more than half of the derivation study authors who participated in the survey reported conducting either a full or limited systematic review of existing CPRs when they derived the new CPRs already.

Guideline developers and clinicians searching for a CPR to implement should beware that there might be many redundant CPRs that were created for the same purposes. Therefore, users should critically compare existing CPRs' evidence of accuracy, generalizability, and impact before choosing a CPR to implement.

2.5.5. Conclusion

In this chapter, I have explored why researchers might keep deriving many new cardiovascular CPRs despite similar CPRs for the same problem already exist. These potentially redundant CPRs that do not progress to external validation, impact assessment, and adoption stages of CPR development are a key source of waste and inefficiency in cardiovascular CPR research. I found that many authors did not cite any existing CPR when justifying a new cardiovascular CPR, although most authors agreed that it was important to do so. When previous CPRs existed, authors often explained deriving a new CPR by pointing out insufficiencies of existing CPRs related to derivation, construct, performance, transferability, and evidence. Many of these insufficiencies could have been resolved by externally validating, updating, and assessing the impact of existing CPRs. Therefore, I have proposed a framework to reduce the derivation of redundant CPRs and increase the value of CPR research. In the following two chapters, I will explore waste and inefficiency in the succeeding stage of CPR development, external validation studies.

Chapter 3. Independent external validation of prediction rules for cardiovascular disease risk

A manuscript based on this chapter was published in BMC Diagnostic and Prognostic Research:

> Ban J-W, Stevens R, Perera R. Predictors for independent external validation of cardiovascular risk clinical prediction rules: Cox proportional hazards regression analyses. Diagnostic and Prognostic Research. 2018;2(1):3. (96)

3.1. Abstract

Background: Although many cardiovascular clinical prediction rules (CPRs) exist, most have not been externally validated. Ideally, a newly derived CPR should be promptly validated in different populations and settings by independent researchers without any conflict of interest. In this chapter, I estimate the probability for a CPR for cardiovascular disease risk to have an independent external validation over time after its derivation and explore why some cardiovascular risk CPRs are externally validated by independent researchers and others are not.

Methods: Cardiovascular risk CPRs included in a systematic review were analysed. I identified independent external validation studies of these cardiovascular risk CPRs by conducting forward citation searches of their derivation studies. The outcome assessed was the time interval between the

publication of a CPR for cardiovascular disease risk and the first independent external validation study. I calculated Kaplan-Meier estimates of the probability to have an independent external validation study after a CPR for cardiovascular disease risk was derived. Using univariable Cox regression analyses, I explored whether features of derivation (study design, geographic location, sample size, number of predictors, presentation format, validation in derivation), reporting (participants, predictors, outcomes, performance measure, information for risk calculation), and publication (journal impact factor) were associated with time to the first independent external validation study.

Results: Of 125 cardiovascular risk CPRs, 29 had an independent external validation study. The median follow-up time was 118 months (95% CI, 99–130). The 25th percentile of event time was 122 months (95% CI, 91–299). The cumulative probability for a CPR for cardiovascular disease risk to have an independent external validation study at 60, 120, and 180 months after its derivation was 10.5% (95% CI, 6.2–17.4), 24.3% (95% CI, 16.7–34.6), and 32.6% (95% CI, 22.9–45.1) respectively. Cardiovascular risk CPRs from the USA were 4.15 times (95% CI, 1.89–9.13, $p = 0.0002$) more likely to have an independent external validation study. Increasing the sample size of derivation by ten times was associated with a 2.32-fold (95% CI, 1.37–3.91, $p = 0.0035$) increase in the probability of having an independent external validation study. CPRs presented with an internal validation tend to get an independent external validation study more promptly (HR = 1.73, 95% CI, 0.77–3.93, $p = 0.0257$). CPRs that reported all the information necessary for calculating a person's risk were 2.65 (95% CI, 1.01–6.96, $p = 0.0289$) times more likely to have an

independent external validation study. Publishing a CPR for cardiovascular disease risk in a journal that has one-unit higher impact factor was associated with a 6% (95% CI, 3–9, p = 0.0002) higher likelihood of an independent external validation study.

Discussion: The probability for a CPR for cardiovascular disease risk to have an independent external validation study was low even many years after it was derived. Researchers deriving a new cardiovascular CPR should use appropriate design, rigorous methods, and clear reporting because doing so might encourage other researchers to externally validate the CPR and elevate the evidence-base for its use in practice.

3.2. Introduction

When it is of interest in understanding whether a CPR performs well in different populations or settings, conducting external validation studies is the next stage in the CPR development (26, 191). Through this stage, evidence about the CPR's generalisability is generated, and clinicians will better understand whether the CPR can accurately predict an outcome when the CPR is applied to their patients. Although an internal validation using techniques such as cross-validation or bootstrapping may be included in a CPR's derivation study, it only tests the CPR's reproducibility and does not provide any information about the generalizability of the CPR (26, 191).

In some cases, understanding the generalisability of a CPR from external validation studies may be less relevant. For example, if a CPR is derived only for use in a very specific population, it may be more important to understand the CPR's reproducibility from internal validations than whether it can be applied to different populations in external validations. In these cases, validation studies that apply the CPR to different populations or subgroups of the original population will likely have a different purpose, aiming to examine variations in performance related to case-mix or prevalence of an outcome (192).

Although clinicians need to understand the generalisability of a CPR from external validation studies before they can trust and adopt it in practice (7, 49), most external validation studies are of limited value because they are poorly designed, conducted, and reported (60, 65, 79). External validation studies are sometimes published as part of a CPR's derivation study or conducted subsequently by researchers who derived the CPR. Systematic reviews have shown that CPRs tend to perform better in external validation studies conducted by researchers involved in deriving the CPRs (79, 193). These studies could have been biased because researchers might have intentionally and unintentionally led the CPRs they derived to perform more favourably (79, 194), further limiting the value of external validation studies to clinicians. Therefore, a CPR's performance should be evaluated in external validation studies conducted by independent researchers (26, 79) who ideally have no conflict of interest with the CPR's derivation study authors.

Although existing studies have evaluated whether CPRs are externally validated by independent researchers (53, 79, 195), the extent of their evaluation was primarily focused on the involvement of derivation study authors. Therefore, it is unknown whether CPRs for cardiovascular disease risks are externally validated by researchers who do not overlap with derivation study authors and do not have other traceable conflicts of interest with derivation study authors.

Furthermore, for any use of a newly derived CPR in practice to be evidence-based (1, 7), it should be supported by evidence of generalizability, ideally from timely and reliable independent external validation studies. In other words, a high-quality, independent external validation study that is conducted soon after a CPR's derivation is more useful to clinicians than the same independent external validation study that only becomes available decades later. Unfortunately, a systematic review by Siontis et al. (195) has shown that independent external validation studies often do not follow the derivation study promptly. However, it is unknown how promptly cardiovascular CPRs are externally validated by independent researchers. It is also unknown why some cardiovascular CPRs are externally validated by independent researchers and other cardiovascular CPRs are not.

3.2.1. Objectives

The objectives of this chapter are:

(1) To estimate the probability for a cardiovascular CPR to have an independent external validation study over time after the CPR is derived, and

(2) To assess whether features of derivation, reporting, and publication of cardiovascular CPR are associated with having an independent external validation study of the CPR.

3.3. Methods

3.3.1. Description of the cohort

I evaluated cardiovascular risk CPRs included in a systematic review by Damen et al. (53). I chose this systematic review as the information source for this chapter because it contained a clearly defined cohort of derivation studies of cardiovascular risk CPRs. Published in 2016, the systematic review included a greater number of as well as more recently published derivation studies of cardiovascular CPRs than the ones included in the International Register of Clinical Prediction Rules for Primary Care (47). In addition, because e the authors of the systematic review specifically searched for CPRs predicting the future risk of cardiovascular events in the general population, the CPRs included in the systematic review were relevant to primary care.

I briefly summarized below how Damen et al. (53) collected cardiovascular disease risk CPRs.

(1) Search: The authors searched Medline (Ovid) and EMBASE (Ovid) for studies published between 2004 and 2013 that derived a prognostic CPR for cardiovascular disease. Also, they identified cardiovascular risk CPRs published before 2004 by reviewing the reference list of systematic reviews found in the electronic database search.

(2) Inclusion criteria: They included articles written in English that derived a CPR for estimating the future risk of arterial disease (e.g. coronary heart disease) in the general population.

(3) Exclusion criteria: They excluded studies that updated an existing CPR by adding new predictors, derived a CPR for a venous disease (e.g., Wells' criteria for deep vein thrombosis), or derived a CPR for a population with a specific health condition (e.g. diabetes mellitus).

By counting all variations of a CPR found in each derivation study, the authors of the systematic review reported identifying 363 unique CPRs for cardiovascular disease risk (53). For this chapter, I did not draw any distinction between a CPR and its variations reported in a derivation study because external validation studies often do not specify which version of the CPR was evaluated. For example, Wilson et al. (56) published Framingham coronary heart disease risk equations as well as point scoring systems for men and women in a derivation study. An external validation study by Marrugat et al. (196) only reported that the authors applied "the original Framingham function" with a citation to the derivation study by Wilson et al. (56) without specifying

which version of the CPR was applied. For this reason, the unit of analysis for this chapter was a derivation study instead of a unique cardiovascular risk CPR.

3.3.2. Assessment of outcome

The outcome assessed in this chapter was the time interval measured in months between the publication of a derivation study and the publication of the first independent external validation study. I chose this outcome because, in addition to whether independent researchers externally validated a CPR for cardiovascular disease risk, it was equally important to understand how promptly the evidence of the CPR's generalisability from an independent external validation study became available to the potential users of the CPR.

I defined an independent external validation as an external validation study that was conducted by researchers who had no conflict of interest with authors of the derivation study. A study was classified as an independent external validation, when

(1) a CPR was applied to a new population different from the ones used in the derivation study,

(2) a performance measure of the CPR such as discrimination or calibration was reported,

(3) no author of the study overlapped with any authors of the derivation study,

(4) no author had prior co-authorship with any authors of the derivation study, and

(5) no other potential conflict of interest was identified after reviewing the authors' affiliations, funding sources, acknowledgements, and conflict of interest statements.

The history of past co-authorship was investigated by searching Scopus Author ID using the "advanced" search option, as shown in Appendix 3.1. In this example, although there was no overlap between the authors of the derivation study of Framingham 30-year cardiovascular disease risk score (197) and its external validation study (198), one of the derivation study authors and one author of the external validation study co-authored 20 publications prior to publishing the external validation study in 2014. I evaluated authors' prior collaborations with authors of the derivation study as well as other factors that could influence the researchers' independence, such as their affiliations, funding sources, acknowledgements, and conflict of interest statements because existing studies only assessed the lack of independence due to overlapping authors with derivation studies.

In external validation studies, a previously derived CPR with defined predictor variables for a defined outcome in a population and setting is applied to populations or settings different from those used in the derivation to assess whether it predicts the outcome accurately (2, 8, 30). The definitions and assessments of predictor variables and outcomes of a CPR in external

validation studies might differ from those used in the derivation, and these variations might be necessary at times due to clinical contexts (e.g. available tests) (74). I excluded studies that just deliberately added new predictor variables to a CPR or assessed outcomes unrelated to a CPR's intended outcomes because these studies might be viewed as modifying rather than externally validating the existing CPR. For example, if a study used a coronary heart disease risk score to assess the risk of atrial fibrillation, the study was excluded. For the same reason, I excluded studies that only compared the risks of an outcome estimated by one CPR with another CPR but did not assess the CPR's intended outcome. However, I acknowledge that this narrow definition of external validation used in this chapter might not be accepted universally.

In August of 2016, I conducted forward citation searches of all derivation studies of cardiovascular risk CPRs included in the systematic review by Damen et al. (53) using Scopus.

For each cardiovascular risk CPR, I screened titles and abstracts of retrieved references in chronological order, starting with the earliest, and reviewed full-text articles of potentially eligible references. This process was continued until the first independent external validation study for the cardiovascular risk CPR was identified. Cardiovascular risk CPRs that had no independent external validation by the time of the forward citation search (August of 2016) were right-censored. To complement the forward citation search, I checked the reference list of the systematic review by Damen et al. (53) to identify any additional

independent external validation studies. Using PubMed, I verified the publication dates of all derivation and independent external validation studies.

3.3.3. Predictors of an independent external validation study

Because little is known about what predicts an independent external validation, I developed a list of predictors that might be associated with having an independent external validation study of a CPR for cardiovascular disease risk by considering how CPRs are derived, reported, and published.

Firstly, I reviewed the categories of existing CPRs' insufficiencies stated by authors of derivation studies in the thematic content analysis of chapter 2 and looked for features of CPR derivation that might be important to researchers who are planning to conduct an external validation study. I selected the following six features of a derivation study as potential predictors of an independent external validation: study design, geographic location, sample size, number of predictors, presentation format, and validation in derivation.

(1) Study design: Ideally, a derivation study should be conducted using a cohort study design (24, 199, 200). Researchers planning an external validation study might disfavour a CPR derived using a case-control study design because the absolute risk of an outcome cannot be calculated (24), and the estimated performance measure, especially calibration, can be incorrect (201).

(2) Geographic location: I used the United Nation's standard country or area codes for statistical use (M49) (202) to define geographical regions where CPRs were derived. Researchers might prefer to externally validate CPRs derived in a particular geographic location that has similar characteristics with the site of external validation because they might view these CPRs as more transferable.

(3) Sample size: Traditionally, it has been suggested that at least ten events per predictor variable are needed when deriving a CPR (203, 204). However, it has been shown that this rule of thumb of ten events per variable lacks an evidence base (205). Instead, determining whether a derivation study has an adequate sample size should be done by considering multiple factors such as the number of individuals in the sample, outcome events, and predictor variables considered (206-208). A sufficient number of participants should be included in a derivation study to reduce the risk of obtaining a biased or imprecise predictive performance measure of CPR (206-208). Researchers might distrust CPRs derived using an insufficient number of participants and might not choose these CPRs for external validations.

(4) Number of predictors: When a CPR has many predictors, authors of an external validation study might find it challenging to use. On the other hand, when a CPR is built with only a few predictors, the CPR might be viewed as missing essential predictor variables.

(5) Presentation format: Researchers might prefer to externally validate a CPR presented with a simplified format for risk calculation such as

scoring system, chart, or online calculator rather than a CPR presented as a mathematical equation.

(6) Validation in derivation: A derivation study may include internal validation, external validation, or both. Authors of an external validation study might prefer CPRs presented with internal and external validation showing the evidence of the CPR's reproducibility and generalizability.

Secondly, I reviewed the TRIPOD statement (6) to identify derivation studies' reporting features that might be essential for conducting an external validation study.

(1) Description of participants: I assessed whether authors fully described eligibility criteria, study settings, and key characteristics because all of these items are important to authors of external validation studies when determining the transferability of a CPR. For example, for researchers who want to externally validate a diagnostic CPR for pulmonary embolism in pregnant patients without a provoking event (e.g. history of recent hospitalization or immobilization) or cancer in a primary care setting, clear descriptions of eligibility criteria, study settings, and key characteristics in derivation studies will be needed to judge whether the CPR is suitable.

(2) Description of predictors: It would be difficult for researchers to externally validate a CPR when its derivation study provided insufficient details

about the predictors and failed to describe how and when they were measured.

(3) Description of outcomes: Similarly, it would be hard to externally validate a CPR when the outcomes were ambiguously described without including how and when they were measured.

(4) Description of performance measure: I assessed whether a performance measure (discrimination or calibration) was reported because authors who want to externally validate a CPR need to know whether the CPR performed well in the derivation study.

(5) Information needed for risk calculation: To externally validate a CPR, researchers must be able to calculate the risk using the CPR. Therefore, I assessed whether sufficient details about a CPR needed for risk calculation were provided. This is typically done by presenting an intercept and coefficients for regression models (6, 74). In addition, how continuous variables were measured (e.g. kg) and how categorical variables were coded (e.g. non-smoker = 0, smoker = 1) should be clearly described. For survival models, regression coefficients should be presented with baseline survival at relevant time points (6, 74). Alternatively, a prediction rule presented using a simplified score (e.g. CHA2DS2-VASc score), online calculator, or statistical code (e.g. QRISK3) would also allow researchers to calculate the risk of a person (74).

(6) Lastly, I hypothesized that CPRs published in journals with a higher impact factor were more likely to be externally validated by independent

researchers. I used the impact factor reported in 2015 Thompson Reuters Journal Citation Index. A list of potential predictors of an independent external validation study and their definitions are presented in Table 3.3.1.

Table 3.3.1 Potential predictors for a prediction rule for cardiovascular disease risk to have an independent external validation.

Predictor variable	Definition
A. CPR derivation	
1. Study design	
Cohort	A prediction was made before an outcome is verified.
Case-control	An outcome was verified before a prediction is made.
2. Geographic location	
USA	A clinical prediction rule (CPR) was derived in the USA.
Other	A CPR was derived outside the USA
3. Sample size	The number of participants used to derive a CPR.
4. Number of predictors	The number of predictors included in the largest model.
5. Presentation format	
Friendly	A simplified format for a risk calculation (e.g. scoring system, chart, and online calculator) was provided.
Unfriendly	A simplified format for a risk calculation was not provided.
6. Validation in derivation	
External	CPR's performance was assessed in a new population different from that of derivation study.
Internal	CPR's performance was assessed using a split sample, cross-validation, or bootstrapping.
None	No internal or external validation was conducted.
B. Reporting and publication	
1. Description of participants	
Clear	Eligibility criteria, settings and key characteristics of participants were clearly described.
Unclear	Participants were unclearly described.
2. Description of predictors	
Clear	All predictors were clearly described, including how and when they were measured.
Unclear	Predictors were unclearly described.
3. Description of outcomes	

Predictor variable	Definition
Clear	Outcomes were clearly described, including how and when they were assessed.
Unclear	Outcomes were unclearly described.
4. Performance measure	
Reported	A CPR's performance measure, such as discrimination or calibration, was reported.
Not reported	No performance measure was reported.
5. Information for risk calculation	
Reported	All the information (e.g. a constant and all regression coefficients, a scoring system with probabilities of an outcome) needed for calculating individual risks was provided.
Not reported	Information provided was insufficient for risk calculation.
6. Journal impact factor	The Impact factor reported in 2015 Thompson Reuters Journal Citation Index.

3.3.4. Statistical analysis

A CPR for cardiovascular disease risk's probability to have an independent external validation study was estimated using the Kaplan-Meier method. I estimated cumulative probability for a CPR for cardiovascular disease risk to have an independent external validation study at 60, 120, and 180 months. Also, I reported the 25th percentile of event time because the cumulative probability of the event, having an independent external validation, never reached 50%. To summarise the length of follow up in my study, I used the reverse Kaplan-Meier method (209). Although there are other methods for estimating a median follow-up time such as time to end-of-study or Korn's follow-up time, these methods tend to either produce an inaccurate estimate of a follow-up or be difficult to calculate (210). Therefore, based on the reverse Kaplan-Meier method, I estimated a median time to censoring after switching

outcomes between cardiovascular risk CPRs that had an independent external validation study and ones that were right-censored.

A logarithmic transformation was applied to derivation studies' sample size because this variable had a very skewed distribution, Figure 3.3.2. Also, because only a small number of cardiovascular risk CPRs from continental Europe, the UK, Asia, and other geographic areas had an independent validation, they were grouped into “other” location. Therefore, there were “USA” and “other” level for the geographic location variable. The derivation studies with missing information about predictor variables were excluded from each corresponding analysis.

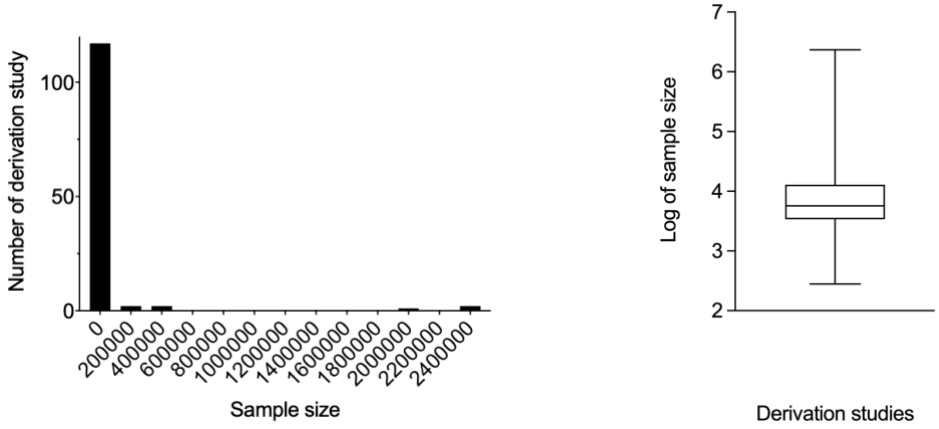


Figure 3.3.1 Distribution of sample size variable before and after logarithmic transformation.

I used Cox proportional hazards regression to evaluate the association between potential predictors and the time interval between the publication of a cardiovascular CPR's derivation study and the first independent external validation study. I focused on estimating HR and 95% CI in univariable Cox proportional hazards regression analyses because the small number of outcome events prohibited evaluating predictor variables using a properly conducted multivariable analysis.

In addition to univariable Cox proportional hazards regression analyses, I graphically compared exposure groups by plotting Kaplan-Meier estimates of a cardiovascular CPR's probability of having an independent external validation study for each categorical variable level. For continuous variables, I generated three levels using tertiles as cutpoints and plotted Kaplan-Meier estimates of a cardiovascular CPR's probability to have an independent external validation study. Logrank tests were used to compare the probabilities of independent external validation study between these categories (or levels).

The proportional hazards assumption was tested using scaled Schoenfeld residuals (211), as shown in Figure 3.3.3 and Figure 3.3.4, and no clear violation was detected. Stata (Release 14. College Station, TX: StataCorp LP) was used for all analyses.

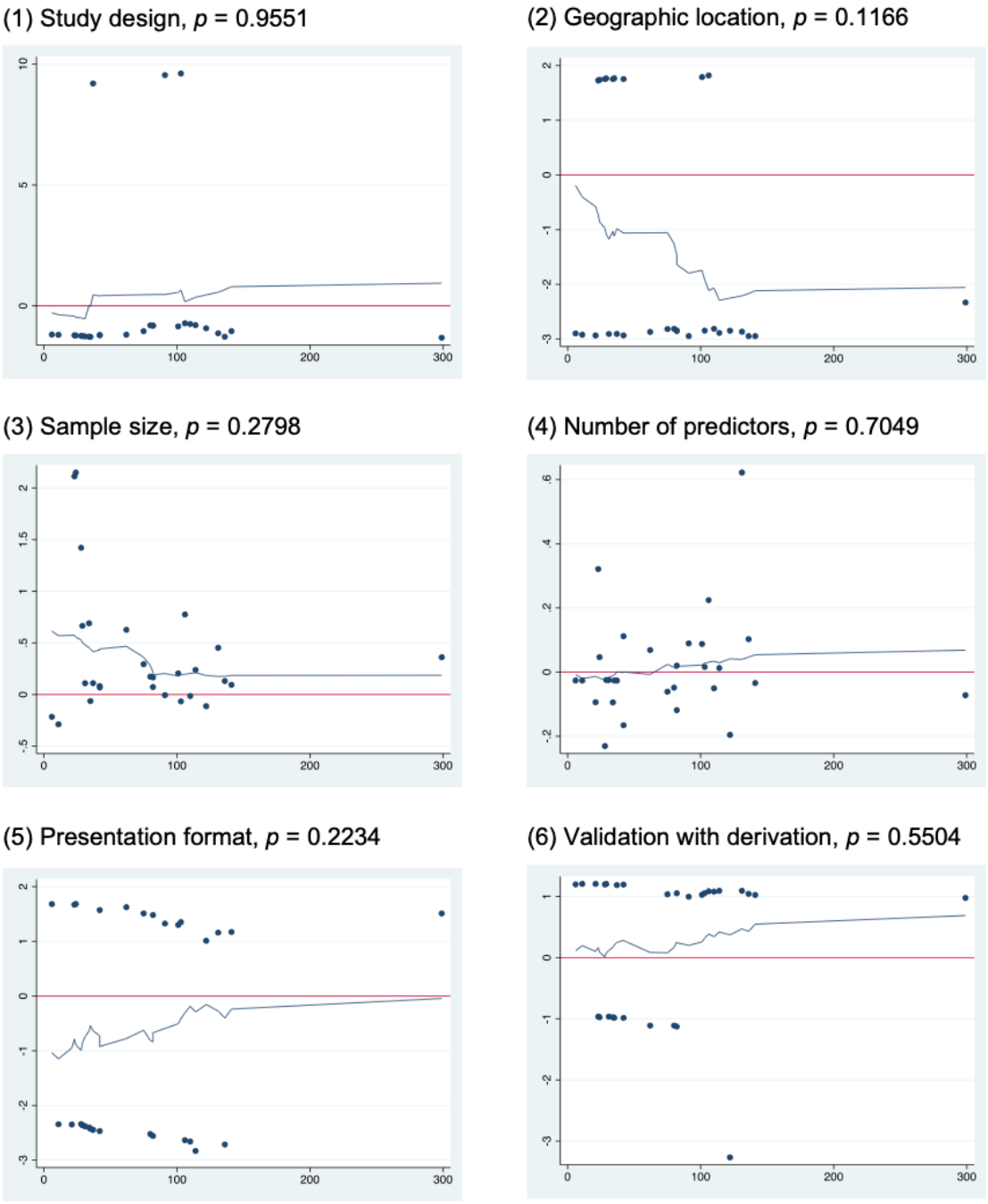


Figure 3.3.2 Assessment of proportional-hazard assumption for univariable Cox regression models with derivation-related predictor variables using scaled Schoenfeld residuals.

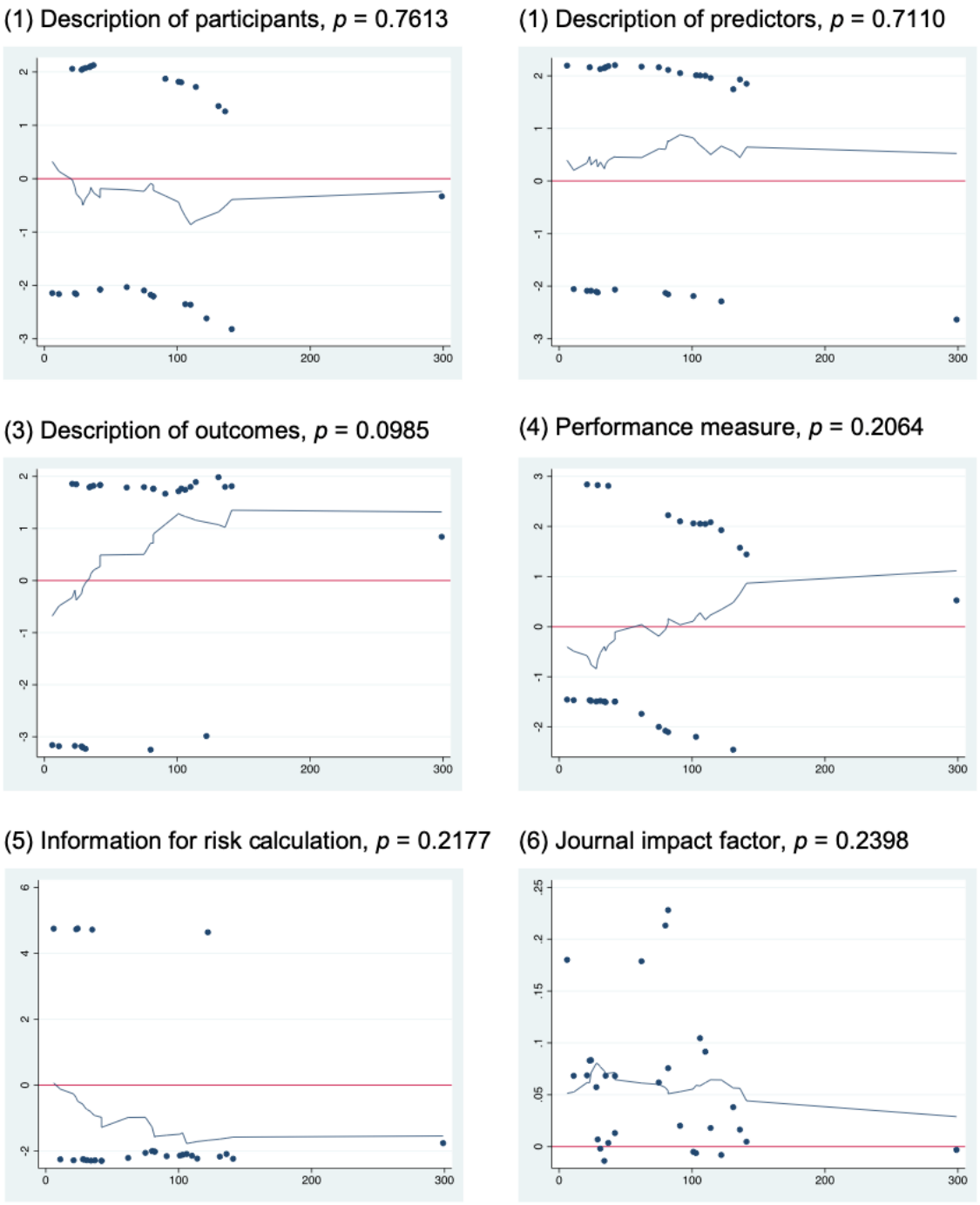


Figure 3.3.3 Assessment of proportional-hazard assumption for univariable Cox regression models with reporting and publication-related predictor variables using scaled Schoenfeld residuals.

3.4. Results

3.4.1. Cardiovascular CPRs with an independent external validation study

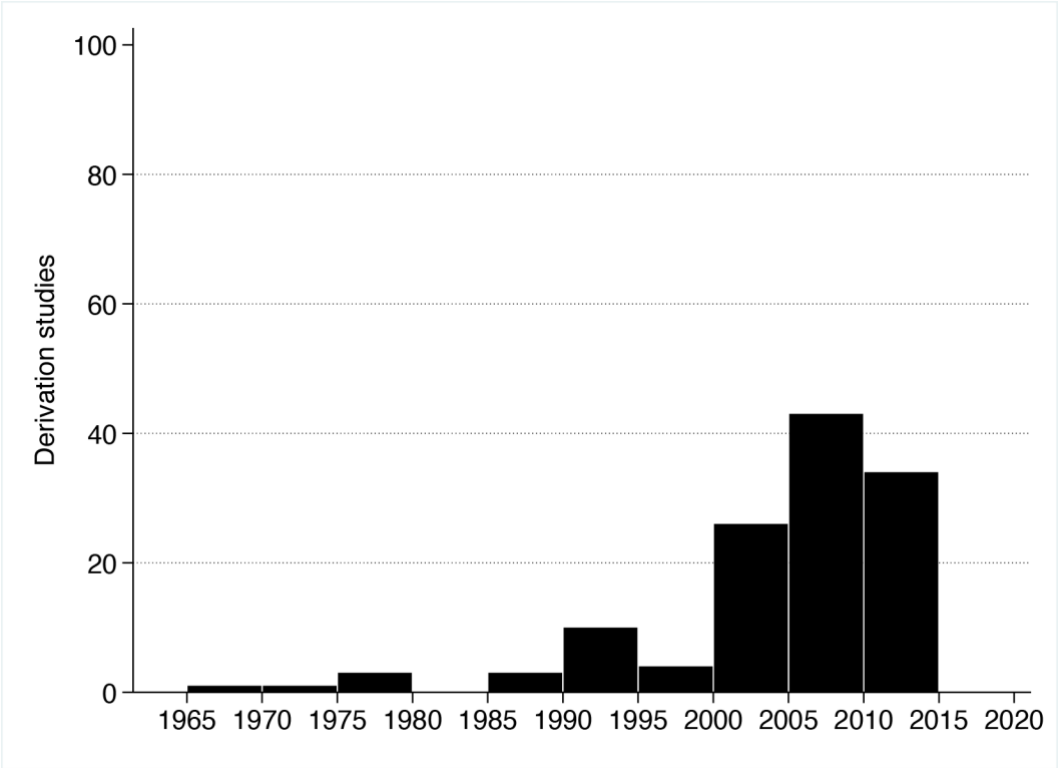
I assessed 125 cardiovascular risk CPRs included in the systematic review by Damen et al. (53). The characteristics of these cardiovascular risk CPRs and their derivation studies included in the analyses are summarized in Table 3.4.1. There was one derivation study (212) for which study design and sample size could not be determined. For another study (213), the number of predictors could not be determined. I summarised publication years of derivation studies and independent external validation studies in Figure 3.4.1.

Table 3.4.1 Characteristics of prediction rules for cardiovascular disease risk and their derivation studies.

	CPRs with an independent external validation, n = 29	CPRs with no independent external validation, n = 96	All cardiovascular risk CPRs, n = 125
A. CPR derivation			
1. Study design [a]			
Cohort	25 (89.3)	85 (88.5)	110 (88.7)
Case-control	3 (10.7)	11 (11.5)	14 (11.3)
2. Geographic location			
USA	20 (69.0)	23 (24.0)	43 (34.4)
Continental Europe	4 (13.8)	25 (26.0)	29 (23.2)
UK	2 (6.9)	10 (10.4)	12 (9.6)
Asia	0 (0.0)	19 (19.8)	19 (15.2)
Other	0 (0.0)	10 (10.4)	10 (8.0)
Multiple countries	3 (10.3)	9 (9.4)	12 (9.6)
3. Sample size of derivation, median (IQR) [a]	6,032 (5,277-19,306)	5,206 (2,721-12,299.5)	5,722.5 (3,396-12,711.5)
4. Number of predictors, median (IQR) [b]	8 (7-9)	8 (6-9)	8 (6-9)

5. Presentation format			
User-friendly	15 (51.7)	37 (38.5)	52 (41.6)
Not user-friendly	14 (48.3)	59 (61.5)	73 (58.4)
6. Validation in derivation			
External	1 (3.5)	18 (18.8)	19 (15.2)
Internal	9 (31.0)	23 (24.0)	32 (25.6)
None	19 (65.5)	55 (57.3)	74 (59.2)
B. Reporting and publication			
1. Description of participants			
Clear	15 (51.7)	46 (47.9)	61 (48.8)
Unclear	14 (42.3)	50 (52.1)	64 (51.2)
2. Description of predictors			
Clear	11 (37.9)	54 (56.3)	65 (52.0)
Unclear	18 (62.1)	42 (43.8)	60 (48.0)
3. Description of outcomes			
Clear	8 (27.6)	40 (41.7)	48 (38.4)
Unclear	21 (72.4)	56 (58.3)	77 (61.6)
4. Performance measure			
Reported	16 (55.2)	67 (69.8)	83 (66.4)
Not reported	13 (44.8)	29 (30.2)	42 (33.6)
5. Information for risk calculation			
Reported	24 (82.8)	54 (56.3)	78 (62.4)
Not reported	5 (17.2)	42 (43.8)	47 (37.6)
6. Impact factor, median (IQR)	15.1 (4.3-17.2)	3.9 (3.1-7.1)	4.3 (3.2-15.1)

Values are the number (%) of CPRs unless indicated otherwise.
[a] Could not be determined for Adult Treatment Panel III rule. (212)
[b] Could not be determined for CHD prevention rule by McNeil et al. (213)



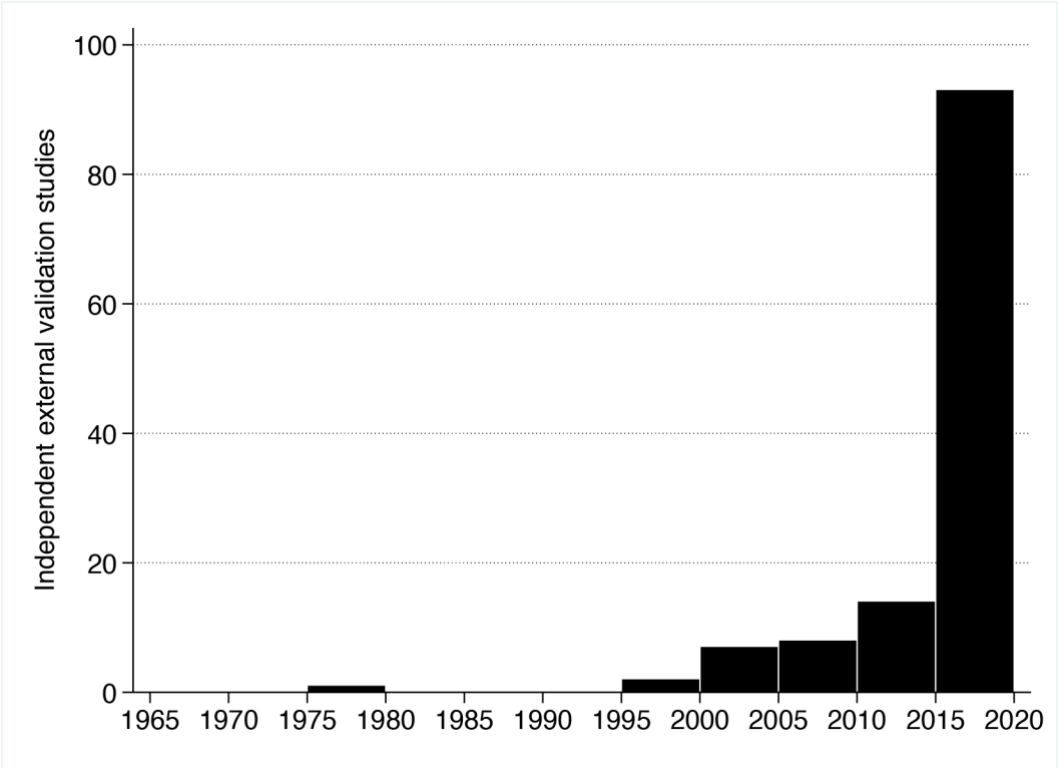


Figure 3.4.1 Publication year of derivation studies and independent external validation studies.

The median time from the publication of a CPR for cardiovascular disease risk to that date of my forward citation search (median follow-up time) was 118 months (95% CI, 99-130). Published in 1975, the 5-year coronary heart disease risk score for middle-aged men by Keys et al. (214) had the longest follow-up time of 532 months.

A flow diagram in Figure 3.4.2 summarizes how cardiovascular risk CPRs with an independent external validation study were identified. For 33 of 125 (26.4%) cardiovascular CPRs, I found external validation studies that had no overlapping author with the derivation study. However, external validation studies of four cardiovascular risk CPRs had been conducted by at least one author who had prior co-authorship with one of the derivation study authors. Therefore, of 125 cardiovascular CPRs assessed, 29 (23.2%) CPRs had an independent external validation study and 96 (76.8%) CPRs had no independent external validation study.

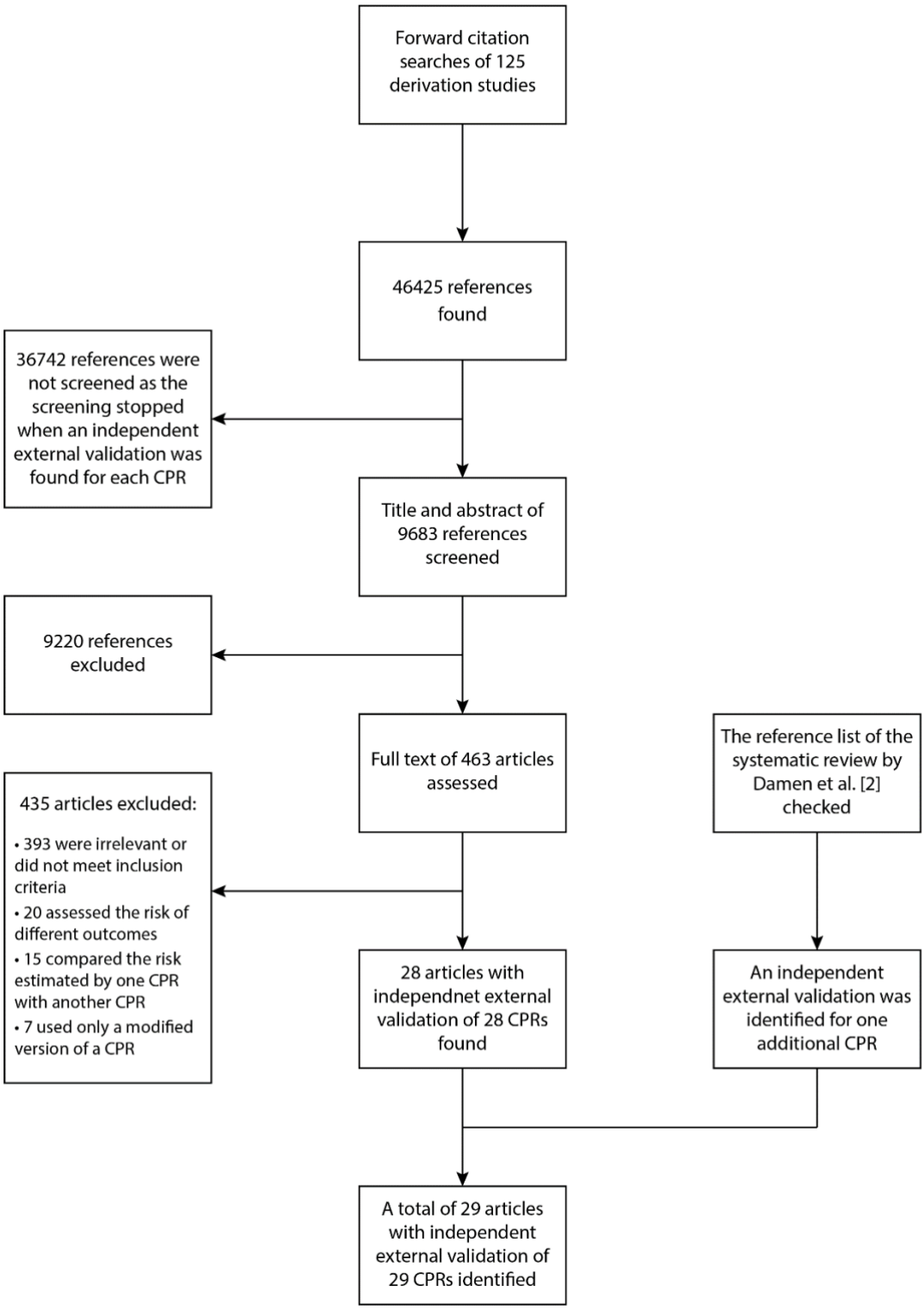


Figure 3.4.2 A flow diagram. Identifying prediction rules for cardiovascular disease risk with an independent external validation study.

The list of journals that published derivation studies of cardiovascular risk CPRs and their independent external validation studies is provided in Table 3.4.2. Derivation studies were most frequently published in Circulation (n = 14) followed by the BMJ (n = 7). The American Heart Journal, the American Journal of Cardiology, and the European Journal of Cardiovascular Prevention and Rehabilitation each published five derivation studies. Independent external validation studies were most frequently published in the BMJ (n = 3) and the American Journal of Cardiology (n = 3). The median impact factor of journals that published the derivation studies of cardiovascular risk CPRs and their independent external validation studies was 15.1 (4.3-17.2) and 5.1 (IQR, 3.4–8.9), respectively.

Table 3.4.2 Number of derivation and independent external validation studies of cardiovascular risk prediction rules published by journals.

Derivation study	Journal
14	Circulation
7	The BMJ
5	The American Heart Journal, the American Journal of Cardiology, the European Journal of Cardiovascular Prevention and Rehabilitation
4	The Journal of Clinical Epidemiology, Preventive Medicine, Stroke
3	Atherosclerosis, Circulation Journal, the European Heart Journal, the European Journal of Epidemiology, Heart, the International Journal of Epidemiology, the JAMA, the Journal of Cardiovascular Risk, the Journal of Epidemiology and Community Health
2	Annals of Internal Medicine, Diabetes Care, the European Journal of Preventive Cardiology, Hypertension Research, Italian Heart Journal, the Journal of the American College of Cardiology, Journal of Chronic Diseases, the Journal of Epidemiology, the Lancet, Nutrition, Metabolism & Cardiovascular Diseases, PLOS ONE
1	Acta Cardiologica, the American Journal of Hypertension, the American Journal of Preventive Medicine, Australian and New Zealand Journal of Public Health, BMC Research Notes, Cardiovascular Diabetology, Cerebrovascular Diseases, Clinical Chemistry, Clinical Endocrinology, Current Medical Research & Opinion, Diabetologia, the European Journal of Clinical Investigation, Heart International, the Hellenic Journal of

	Cardiology, Hypertension, the International Journal of Cardiology, the International Journal of Obesity, International Journal of Stroke, the Journal of the American Society of Hypertension, Journal of Atherosclerosis and Thrombosis, the Journal of Electrocardiology, Journal of Hypertension, the Journal of Nutrition, the Laryngoscope, the Medical Journal of Australia, Neurological Sciences, Public Health Nursing, Statistics in Medicine
Independent external validation study	Journal
3	The American Journal of Cardiology, the BMJ
2	Diabetes Care, the European Journal of Cardiovascular Prevention and Rehabilitation, the Journal of the American College of Cardiology, the Journal of Epidemiology and Community Health, Stroke
1	The American Journal of Medicine, Annals of Internal Medicine, Atención Primaria, Circulation, Heart, Hypertension, Hypertension Research, Indian Journal of Community Health, the International Journal of Cardiology, Journal of hypertension, the Medical journal of Australia, Neuroepidemiology, Nutrition, Metabolism & Cardiovascular Diseases

3.4.2. Probability for an independent external validation study

Kaplan-Meier estimates of the probability for a CPR for cardiovascular disease risk to have an independent external validation study are illustrated in Figure 3.4.3.

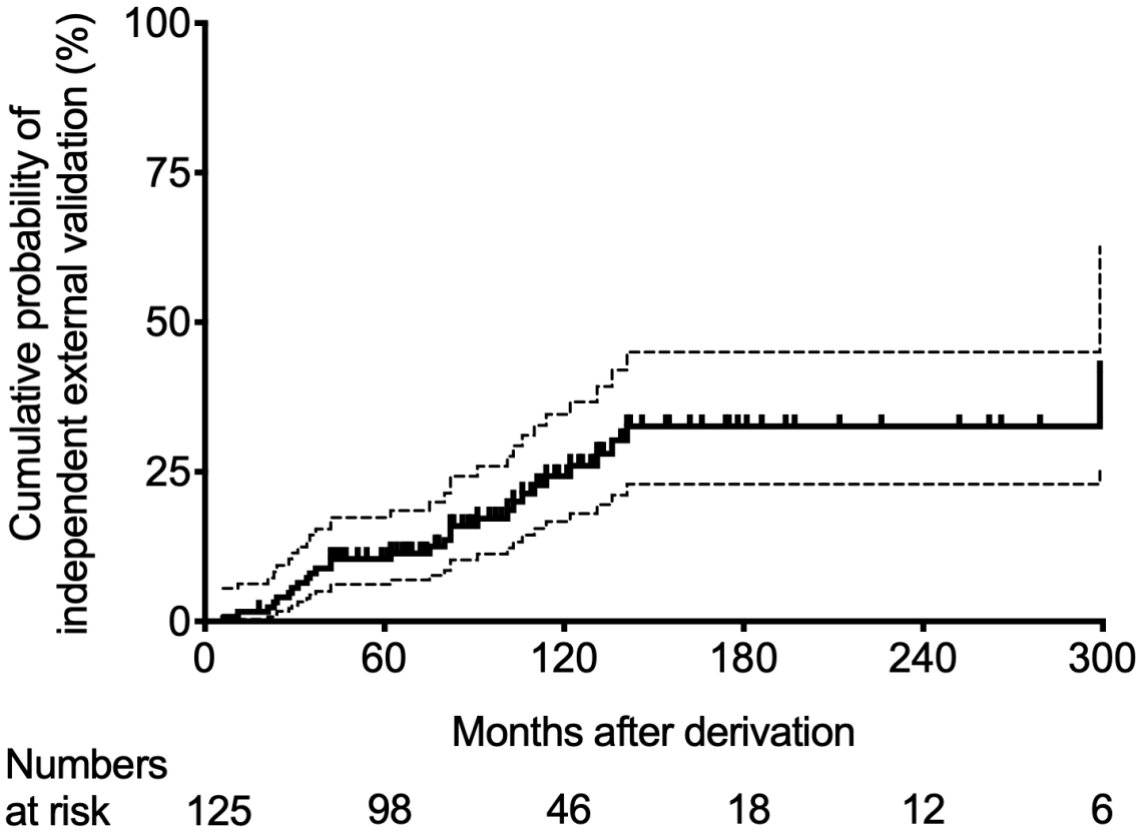


Figure 3.4.3 Kaplan-Meier plot. Probability for a prediction rule for cardiovascular disease risk to have an independent external validation study and 95% confidence interval.

It took 122 months (95% CI, 91-299) before the probability for a CPR for cardiovascular disease risk to have an independent external validation study reached 25%. The cumulative probability for a CPR for cardiovascular disease risk to have an independent external validation study at 60, 120, and 180 months (or 5.0, 10.0. and 15.0 years) after its derivation was 10.5% (95% CI, 6.2–17.4), 24.3% (95% CI, 16.7–34.6), and 32.6% (95% CI, 22.9–45.1) respectively.

A coronary heart disease risk score by Polonsky et al. (215) had the shortest interval of 6 months until the first independent external validation study, followed

by the Framingham general cardiovascular disease risk score by D'Agostino et al. (216) and Framingham hard coronary heart disease risk scores from the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) (217) with corresponding time to the first independent validation study of 11 and 21 months. All independent external validation studies were published within 142 months after the publication of its derivation study, except for the Framingham coronary heart disease risk score by Wilson et al. (218) which took 299 months until the first independent external validation was published.

3.4.3. Univariable analysis

The Kaplan-Meier estimates of a cardiovascular CPR's probability to have an independent external validation study for exposure groups are graphically compared in Figure 3.4.4 and Figure 3.4.5.

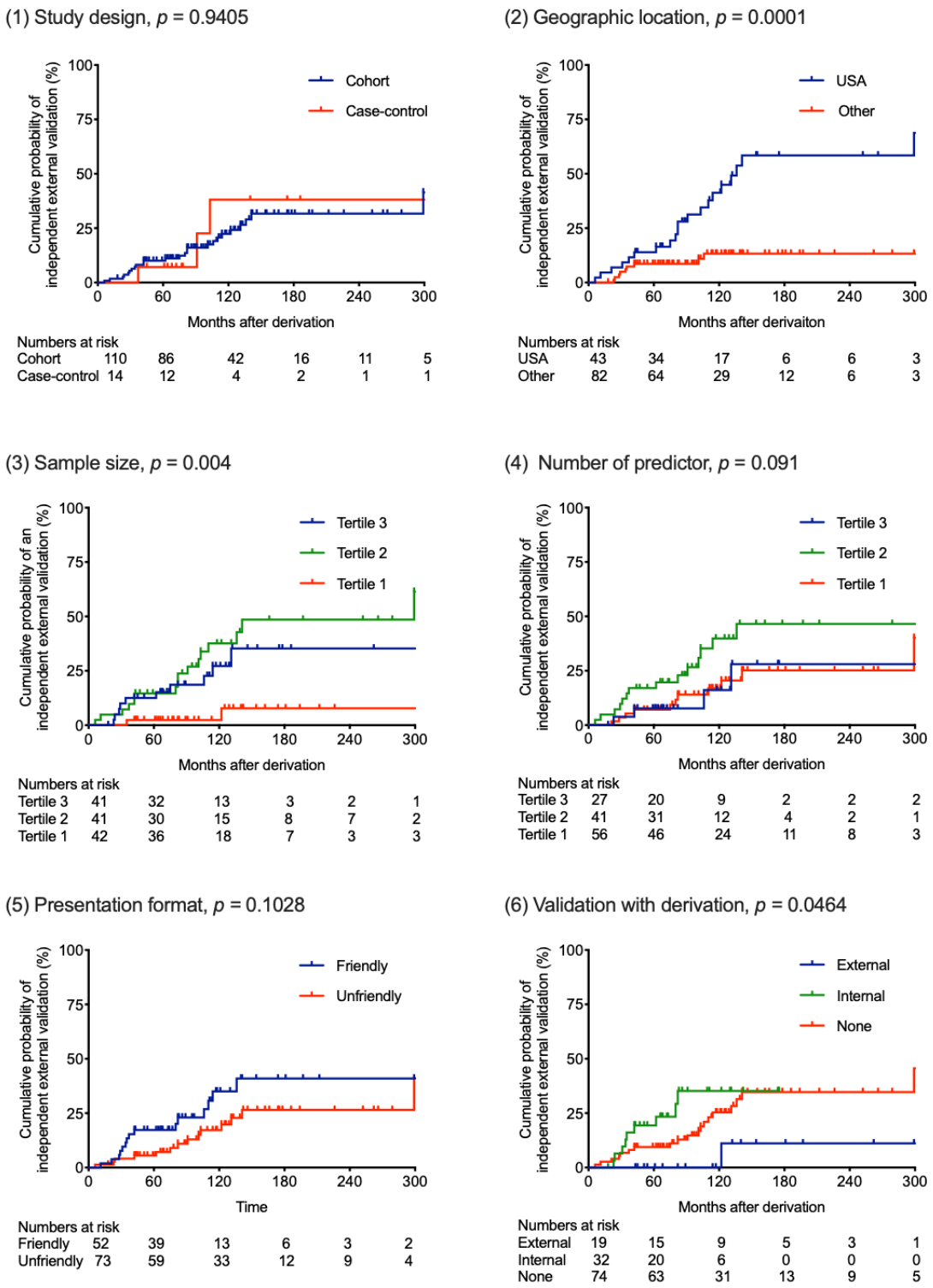


Figure 3.4.4 Kaplan-Meier plot. Influence of derivation-related variables on the probability for a cardiovascular clinical prediction rule to have an independent external validation study.

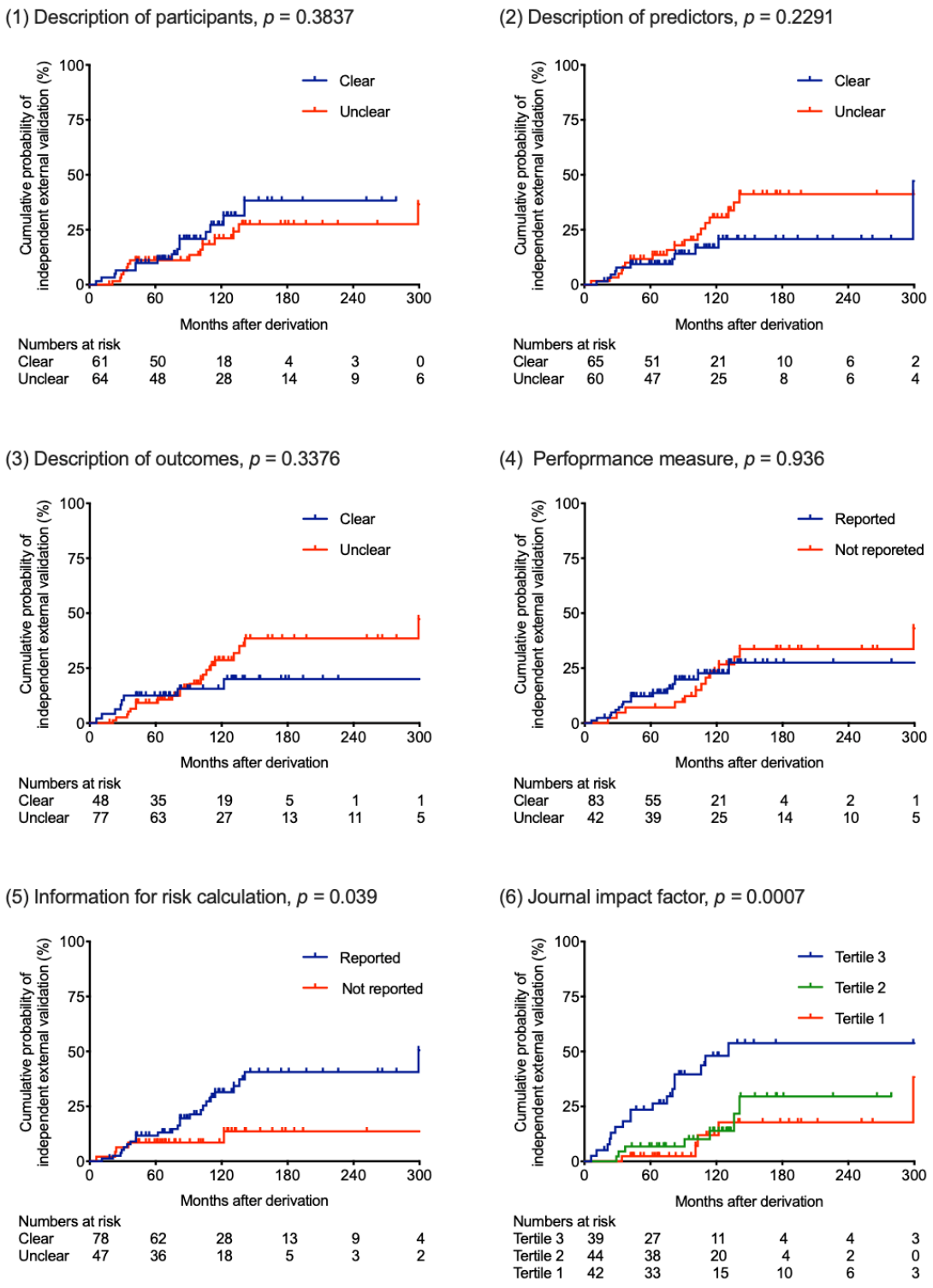


Figure 3.4.5 Kaplan-Meier plot. Influence of reporting and publication-related variables on the probability for a prediction rule for cardiovascular disease risk to have an independent external validation study.

The results of univariable analyses from Cox proportional hazards regressions are presented in Table 3.4.3. Three of the six features of CPR derivation studies assessed might be associated with having an independent external validation study of a CPR for cardiovascular disease risk. For the geographic location, a CPR for cardiovascular disease risk derived in the USA was 4.15 times more likely to have an independent external validation study (95% CI 1.89-9.13) than cardiovascular risk CPRs derived elsewhere. In a post hoc sensitivity analysis that excluded 17 cardiovascular risk CPRs derived by Framingham Heart Study researchers, the HR for cardiovascular risk CPRs derived in the USA relative to cardiovascular risk CPRs derived elsewhere was 2.46 (95% CI, 0.92-6.61, p = 0.0842).

The sample size of derivation studies was associated with having an independent external validation with a HR of 2.32 (95% CI 1.37-3.91) for the Log_{10}(Sample size). The observed proportion of CPRs with an independent external validation study was 0.0% when the derivation sample size was 100-1,000, 23.4% when it was 1,001-10,000, 22.6% when it was 10,001-100,000, 25.0% when it was 100,001-1,000,000, and 66.7% when it was 1,000,001-10,000,000.

The categorical variable, “validation in derivation,” seems to be associated with having an independent external validation study (p = 0.0257). However, neither providing internal validation (HR = 1.73, 95% CI 0.77-3.90) nor external validation (HR = 0.19, 95% CI 0.03-1.42) was associated with a significant

increase in the probability of having an independent external validation study compared with not providing any validation in a derivation study of cardiovascular risk CPR. On the other hand, presenting internal validation might increase the probability of having an independent external validation study by 9.10 times (p = 0.037, 95% CI 1.14-72.5) compared with including external validation in a derivation study of a CPR for cardiovascular disease risk.

Of six reporting and publication-related features analysed, reporting information for risk calculation (HR = 2.65, 95% CI 1.01-6.96) and publishing the derivation study in a journal with higher impact factor (HR = 1.06, 95% CI 1.03-1.09) were associated with having an independent external validation study of a CPR for cardiovascular disease risk.

I also conducted a post hoc sensitivity analysis to evaluate the influence of CPRs from Framingham Heart Study and QResearch that are often recommended in guidelines and published in high impact journals. Even when 20 CPRs derived by the Framingham Heart Study and QResearch researchers were excluded, publishing the derivation study in a journal with a higher impact factor was still associated with having an independent external validation study with a HR of 1.07 (95% CI, 1.03-1.11, p < 0.001).

Table 3.4.3 Univariable Cox proportional hazards regression analyses.

	Hazard ratio	(95% CI)	*p*-value
A. CPR derivation			
1. Study design			0.9409
Cohort	0.96	(0.29-3.17)	

	Hazard ratio	(95% CI)	*p*-value
Case-control	1	-	
2. Geographic location			0.0002
USA	4.15	(1.89-9.13)	
Other	1	-	
3. Log_{10}(Sample size)	2.32	(1.37-3.91)	0.0035
4. Number of predictors	1.01	(0.92-1.12)	0.7833
5. Presentation format			0.1095
User-friendly	1.82	(0.88-3.79)	
Not user-friendly	1	-	
6. Validation in derivation			0.0257
External	0.19	(0.03-1.42)	
Internal	1.73	(0.77-3.90)	
None	1	-	
B. Reporting and publication			
1. Description of participants			0.3853
Clear	1.39	(0.66-2.94)	
Unclear	1	-	
2. Description of predictors			0.2269
Clear	0.63	(0.30-1.34)	
Unclear	1		
3. Description of outcomes			0.3282
Clear	0.67	(0.30-1.52)	
Unclear	1		
4. Performance measure			0.9360
Reported	1.03	(0.48-2.20)	
Not reported	1		
5. Information for risk calculation			0.0289
Reported	2.65	(1.01-6.96)	
Not reported	1		
6. Impact factor	1.06	(1.03-1.09)	0.0002

3.5. Discussion

In this chapter, I assessed the probability for a newly derived cardiovascular risk CPR to have an independent external validation study over time and explored whether 12 characteristics of derivation, reporting, and publication of cardiovascular risk CPRs were associated with having an independent external validation study.

3.5.1. Summary of findings

I found that the probability for a CPR for cardiovascular disease risk to have a timely independent external validation study was low. This greatly limits the value of studies deriving new CPRs, because, without strong evidence of generalizability, CPRs cannot make an evidence-based contribution to clinical practice. I found that CPRs derived in the USA were four times more likely to be externally validated by independent researchers, although this is heavily influenced by multiple CPRs from the Framingham Heart Study. Besides the geographic location where a cardiovascular CPR was derived, using a larger sample size and providing the results of an internal validation were features of CPR derivation that were associated with having an independent validation study of a CPR for cardiovascular disease risk. Publishing the derivation study in a journal with a higher impact factor and providing sufficient information for risk calculation were characteristics of reporting and publication that were associated with having an independent external validation study of a CPR for cardiovascular disease risk. Providing the results of an internal validation and sufficient information for risk calculation are at least within the control of the

derivation study authors, and might also provide a route for authors to increase the likelihood that their published CPRs will progress further along the pathway to evidence-based practice.

3.5.2. Comparison with existing literature

The term, independent external validation, has been used in CPR literature to indicate that a CPR was externally validated by researchers who were not involved in deriving it (193, 219-221) and when comparing CPRs in a dataset unrelated to any datasets used in CPRs' derivation studies (222-225). On the other hand, my definition of independent external validation was stricter in the way I defined "independent researchers" as researchers who were not involved in deriving the CPRs and also did not have any other potential conflicts of interest such as previous research collaboration with derivation authors.

A systematic review by Collins et al. (79) that summarized the methods and reporting of external validation studies of CPRs found that authors of 51 of 78 (65.4%) external validation studies included researchers who were involved in deriving the CPR: 33 (42.3%) were published as a part of a derivation study, and 18 (23.1%) included at least one author who was involved in deriving the CPR. Because I found some external validation studies of cardiovascular risk CPRs that had no overlapping author with their derivation studies had been conducted by authors who had prior co-authorships with derivation study authors, the proportion of external validation studies conducted by researchers with conflict of interest might be higher than suggested by Collins et al. (79).

A systematic review by Siontis et al. (195) found that the probability for a prognostic CPR published before 2011 to have an external validation study by completely different researchers from derivation study authors was 16% at five years after its derivation. The cumulative probability for a CPR for cardiovascular disease risk to have an independent external validation study at 60 months (or five years) after its derivation found in this chapter was 10.5% which is lower than the probability for a prognostic CPR to have an independent external validation study reported by Siontis et al. (195). It could be speculated that this difference might be due to the relative abundance of cardiovascular risk CPRs and longer time required to externally validate them compared with prognostic CPRs for other clinical conditions such as cancer.

3.5.3. Strengths and Limitations

One of the strengths of this study was conducting survival analysis by choosing the time to an independent external validation study of CPR as the outcome, whereas most previous studies evaluated the presence of an external validation study conducted by authors who were not involved in deriving the CPR (53, 79, 226). For clinicians interested in using a new cardiovascular CPR that showed promising accuracy in a derivation study, it is important for them to promptly know that the CPR also performs well in independent external validation studies. Having the first independent external validation study decades later will be much less meaningful to these clinicians. I started with a clearly defined cohort of derivation studies of cardiovascular risk CPRs and ascertained the outcome for all included cardiovascular risk CPRs up to the date of the forward

citation search. In addition, complete data about predictor variables were extracted for almost all derivation studies: of 125 derivation studies, one had two missing predictor variables, and another had one missing predictor variable. I analysed cardiovascular risk CPRs included in a systematic review by Damen et al. (53), which reported that 19% of cardiovascular risk CPRs had an independent external validation. Despite applying a stricter definition of independent external validation and using derivation and validation study as units of analysis, I found that 23.2% of cardiovascular risk CPRs had an independent external validation which is higher than reported in the systematic review by Damen et al. (53). I believe this shows the robustness of the forward citation search used in this study as a method to identify external validation studies of a CPR.

This study also had the following limitations. I dichotomised external validation studies to the ones conducted by independent researchers and others conducted by non-independent researchers. However, externally validating prediction rules by researchers who derived them can be a good first step (79) when external validation studies are properly designed, well-conducted, and clearly reported. Therefore, these studies should not be viewed as research waste just because some authors were involved in deriving the CPRs. On the other hand, it is important to make a clear distinction between external validation studies conducted by independent researchers and the ones done by researchers who derived them (79) because systematic reviews have consistently shown those conducted by researchers derived them tend to

produce inflated performance (79, 193, 227), which might have been caused by spin or selective reporting (79).

In addition, my approach does not necessarily capture complexities and practical challenges in conducting external validation studies. For example, there might be much fewer incentives for researchers to conduct external validation studies than deriving a new CPR (e.g. availability of funding, publishing research in higher impact journals). Further, this approach solely relied on what was reported in derivation studies to explain why some prediction rules were externally validated by independent researchers. Therefore, I could not account for more contextual reasons for conducting external validation studies. For example, externally validating a prediction rule might be prompted to meet the necessity of external bodies such as guideline developers or policymakers.

I defined independent external validation as an external validation study conducted by researchers who have no conflict of interest with the derivation study authors. I evaluated authors' prior collaborations with authors of the derivation study but also other sources for conflict of interest such as the authors' affiliations, funding sources, acknowledgements, and conflict of interest statements. Although I attempted to identify all pragmatically searchable conflicts of interest, some forms of collaboration between authors of derivation and external validation study may not have been traceable. Further, given limited incentives and complexities in motives for researchers to conduct

external validation studies, this definition of independence might have been too strict. Nevertheless, I did not find any external validation study that would have been otherwise categorised as “independent” because of these other sources of conflict of interest.

Because authors of external validation studies often did not provide sufficient details of CPRs they are validating, I relied on external validation study authors’ citation of a derivation study to confirm which CPR is validated. Unfortunately, authors of some external validation studies do not always cite the CPR they are validating accurately (79) and it is possible that some independent external validation studies were not identified in my forward citation searches.

The number of available cardiovascular risk CPRs and independent external validation studies precluded assessing the predictor variables in a multivariable analysis. Although it might have been possible to conduct a multivariable Cox regression by using advanced methods such as Firth’s correction to reduce bias from the small sample size and outcome events (228, 229), I was unaware of such methods when I developed the study protocol and published the paper based on this chapter (96). There are several implications of only conducting univariable analyses. Firstly, I could not control for potential confounding even though some predictor variables evaluated in the study were correlated. For example, the studies published in higher impact journals generally had larger sample sizes. Further, potential interactions between independent variables could not be explored or incorporated into the model. Secondly, because I

conducted multiple univariable analyses, the risk of type 1 error might have been increased. Therefore, findings from this chapter should be considered only as exploratory and hypothesis-generating. Any positive results should be interpreted cautiously.

I evaluated CPRs for cardiovascular disease risk included in the systematic review by Damen et al. (53) published in May 2016. I conducted forward citation searches of these CPRs in Aug 2016 to assess the outcome, collected and analysed data through July 2017, and drafted a manuscript and submitted it to the BMC diagnostic and prognostic research in October 2017. Therefore, although the findings from this chapter might have been contemporary when the manuscript based on this chapter was published in February of 2018 (96), they might not be directly applicable to more recently derived CPRs.

Lastly, I did not assess the risk of bias of included external validation studies. Because the data collection and analysis for this chapter were conducted between 2015 and 2017, a paper based on this chapter was published in Jan 2018 (96), and the final version of the Prediction model Risk Of Bias Assessment Tool (PROBAST) (230) was not published until 2019, I was not able to carry out the risk of bias assessment. Although I considered using preliminary versions of PROBAST (231, 232), the lack of clear guidance on the proper application of the tool precluded assessing the risk of bias using the earlier versions.

3.5.4. Research and clinical implications

Although hundreds of cardiovascular CPRs already exist, the number of studies deriving a new cardiovascular CPR continues to increase (47, 51, 52). The findings from this study, along with previous systematic reviews (47, 53, 79, 195, 226), highlights that too much focus is on creating new CPRs. Following the framework presented in Figure 2.5.1 of chapter 2, it might be possible to change the focus of CPR research to externally validating, updating, assessing the impact of existing CPRs. In addition to researchers' affinity for deriving a CPR, academic journals also seem to favour publishing studies deriving a new CPR over studies externally validating existing CPRs suggested by the median impact factor of journals that published derivation studies of cardiovascular risk CPRs and their independent external validation in this study of 15.1 (4.3-17.2) and 5.1 (IQR, 3.4-8.9) respectively. It is necessary for journal editors to start viewing well-conducted independent external validation studies as studies generating a higher level of evidence than studies deriving a new CPR, especially when there are already many similar CPRs. Without this paradigm shift in journals' view on CPR research, redundant CPRs will continue to be published, and independent external validations and impact assessments of existing high-quality CPRs will likely be delayed.

When a new CPR is well justified, it should be created using proper design and rigorous methods to avoid adding more redundant CPRs. Researchers should especially consider using appropriate sample size and conducting an internal validation as these might increase the chance of having an external validation

by independent researchers. Systematic reviews have pointed out that the quality of reporting in CPR research is poor (59, 60, 71, 73, 79, 233). Published in 2015, the TRIPOD statement (6) provided much-needed guidance to authors deriving and validating CPRs. I chose five reporting features from the TRIPOD statement that might be particularly important to researchers externally validating CPRs and found providing sufficient details of a CPR needed for a risk calculation was associated with a higher chance of having an independent external validation. Although this study found such association in only one of the five reporting features assessed, I firmly believe that all important information should be unambiguously described according to the TRIPOD statement's recommendations, so that others can validate, update, implement, and use the CPR.

Clinicians who want to use a CPR need to understand the evidence of generalizability is established by appraising the performance of the CPR presented in external validation studies. Clinicians should recognize that cardiovascular CPRs externally validated by independent researchers are uncommon. When a potential conflict of interest is discovered between the authors of derivation study and the authors of external validation studies, the CPR's performance should be interpreted with caution because it could have been overestimated.

3.5.5. Conclusion

In this chapter, I inquired into one of the problems in the second stage of cardiovascular CPR development, the lack of independent external validation studies. I have shown that for most newly derived cardiovascular CPRs, evidence of generalizability evaluated by independent researchers is not promptly available to potential users. This lack of high-quality evidence of generalizability might be one of the barriers for clinicians who want to use CPRs in practice. I also found that certain features of CPR derivation, reporting, and publishing might be associated with a greater chance of having an independent external validation. These features include using a larger sample size in a derivation study, including internal validation in a derivation study, providing sufficient details of the CPR for risk calculation, derived in the USA, and publishing in a journal with a higher impact factor. It can be hypothesized from this finding that deriving a CPR using appropriate design, rigorous methods, and clear reporting can increase the chance of other independent researchers to externally validate the CPR quickly. In the next chapter, I will explore another problem in the second stage of cardiovascular CPR development, the value of having many external validation studies of a CPR when most cardiovascular CPRs do not have an external validation study.

Chapter 4. Inefficiencies in external validation studies of Framingham Wilson coronary heart disease risk rule

4.1. Abstract

Background: External validation studies can generate 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 by contributing to estimating its overall performance. While most cardiovascular CPRs do not have any external validation study, some CPRs have been externally validated repeatedly. In this chapter, I examine the inefficiencies of external validation studies of the Framingham Wilson CHD risk rule.

Methods: I conducted a forward citation search of the Framingham Wilson CHD risk rule 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, I examined whether authors updated the Framingham Wilson CHD risk rule when they concluded that the CPR performed poorly. I also assessed the proportion of external validation studies that contribute to understanding the Predicted/Observed (P/O) event ratio and *c* statistic of the Framingham Wilson CHD risk rule to their full potential.

Results: A total of 39 external validation studies of the Framingham Wilson CHD risk rule were identified. Of 39 external validation studies, 26 (66.7%) concluded the Framingham Wilson CHD risk rule performed poorly but did not update it. Twenty-two external validation studies were conducted with data that could be included in the meta-analyses. Of these, 13 (59.1%) failed to fully contribute to meta-analyses of P/O ratio and *c* statistic because the performance measures were not reported and could not be calculated from provided data.

Discussion: The majority of external validation studies did not generate meaningful evidence about the Framingham Wilson CHD risk rule's generalizability. Researchers might increase the value of external validation studies by update the CPR when it performs poorly and presenting all relevant performance measures.

4.2. Introduction

External validation is a stage of CPR development where a CPR derived in a particular population is evaluated in different populations with a comparable health condition to the health condition of the derivation population (e.g. in another country) (191, 234). An external validation study can generate evidence about a CPR's generalizability in two different ways, as illustrated in Figure 4.2.1. This figure is an adaptation of Fig. 17.1 from the book titled "Clinical prediction models: a practical approach to development, validation, and updating" (72) with permission from the author as well as the publisher.

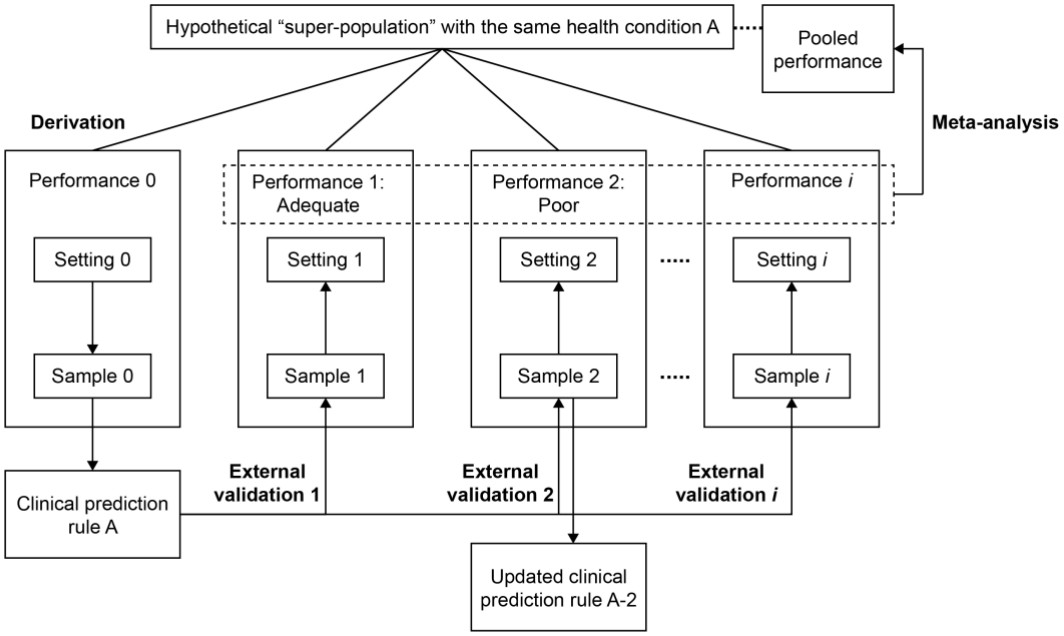


Figure 4.2.1 A conceptual framework of efficient evidence generation from external validation studies of a clinical prediction rule. The figure is an adaptation of Fig 17.1 from “Clinical prediction models: a practical approach to development, validation, and updating” by Steyerberg (72).

On a narrow scope, an external validation study evaluates how well a CPR predicts an outcome when it is applied to a new population and setting (26, 30). But more broadly, the results of external validation studies contribute to understanding the CPR's overall performance within a hypothetical "super-population" with the same health condition when it is meta-analysed in a systematic review (72). The results from meta-analyses of a CPR's external validation studies can help clinicians understand the CPR's average performance and inform how they use the CPR in practice. For example, the American College of Cardiology and American Heart Association (ACC/AHA) recommended in 2014 using the pooled cohort equation derived from a diverse population instead of the third Adult Treatment Panel (ATP III) version of the Framingham 10 year CHD risk score because it was derived from an entirely

white population (235). However, a systematic review and meta-analyses of external validation studies showed that the ACC/AHA pooled cohort equation overestimates cardiovascular disease (CVD) risk similarly to the Framingham ATP III CHD risk score (236). Based on these results, clinicians could use caution when interpreting a person's CVD risk estimated using the ACC/AHA pooled cohort equation.

For the results of an external validation study to contribute to a meta-analysis, it should provide relevant performance measures of a prediction rule. According to the TRIPOD statement, calibration and discrimination "should be reported in all prediction model papers" (74). However, systematic reviews have shown that external validation studies often do not report performance measures, especially calibration (53, 73, 79). When external validation studies do not report relevant performance measures, they might not be able to contribute to understanding a CPR's overall performance in a meta-analysis to their full potential, limiting the value of external validation studies.

When a CPR performs adequately in an external validation study, it generates evidence about the CPR's generalizability to populations used in the external validation study. Thus, clinicians can apply the CPR to populations and settings similar to those used in the external validation study with confidence in its predictive performance (1, 7).

However, a CPR that accurately predicts an outcome in a derivation study often performs unsatisfactorily in external validation studies. This performance gap may be due to the inflated accuracy from suboptimal design and methods used to derive the CPR or differences in populations and settings between derivation and external validation studies (2, 30). When a CPR performs poorly in an external validation study, researchers sometimes declare that the CPR is not generalisable and discourage using the CPR. But this approach is an inefficient use of research and evidence for the following reasons. Firstly, by dismissing the CPR that performed poorly in the external validation study, we lose the knowledge generated in the derivation study (2, 8). Moreover, with this approach, either derivation studies of new CPRs or more external validation studies of existing CPRs need to be conducted to identify a CPR applicable to the population. Secondly, this approach might lead to generating many redundant CPRs for the same problem (2) because, as I found in the thematic content analysis of chapter 2, the poor performance of CPR is one of the most commonly stated justifications by authors who derived a new cardiovascular CPR.

Alternatively, researchers should attempt to update the CPR that performs poorly in an external validation study and evaluate whether the performance could reach clinically meaningful accuracy (2, 8, 29, 30). A well-designed CPR derived in a large dataset might still perform poorly in external validation studies. For example, when the likelihood of an outcome in external validation datasets is very different from the ones used in a CPR's derivation, the CPR might perform poorly in external validation studies (2, 192). Researchers could

update the CPR simply by adjusting the calibration intercept and might be able to improve the calibration of the CPR (2, 31, 32). A CPR might also perform poorly when the strengths of predictor variables differ in external validation datasets (2, 192). In this case, the CPR's calibration and discrimination might improve when coefficients of predictor variables are re-estimated using external validation datasets (2, 31).

A systematic review has shown that updating a CPR using a simple technique could substantially improve the CPR's performance (236). Furthermore, there are examples of newly derived CPRs that did not perform better than the updated versions of existing CPRs (237, 238). Therefore, updating a CPR that performs poorly in external validation is a more efficient use of research and evidence. Also, updating the CPR can generate evidence that might be more useful to clinical practice than simply rejecting the CPR.

While most cardiovascular CPRs do not have a timely conducted independent external validation study, as I have shown in chapter 3, many external validation studies have been performed repeatedly for some cardiovascular CPRs. Framingham Wilson CHD risk rule was identified as the most frequently validated cardiovascular CPR in systematic reviews (47, 53). In this chapter, I quantify inefficiencies among studies externally validating the Framingham Wilson CHD risk rule.

4.2.1. Objectives

The objectives of this chapter are:

(1) to examine whether authors of external validation studies update the Framingham Wilson CHD risk rule when it performs poorly in new populations or settings, and

(2) to evaluate whether external validation studies contribute to understanding the overall performance of the Framingham Wilson CHD risk rule.

4.3. Methods

4.3.1. Information source and search for external validation studies

The cardiovascular CPR evaluated in this chapter is the Framingham Wilson CHD risk rule. In 1998, Wilson et al. (56) published a gender-specific CPR that 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). The predictors included in the Framingham Wilson CHD risk rule are age, cigarette use, diabetes mellitus status, total and high-density lipoprotein cholesterol categories according to the second report of the National Cholesterol Education Program expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (NCEP ATP II) (239), and blood pressure categories according to the fifth report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (240).

Although I considered other cardiovascular CPRs with multiple external validation studies such as Framingham Anderson CVD risk rule (241, 242), $CHADS_2$ and CHA_2DS_2-VASc scores (184, 243), ABCD and $ABCD^2$ scores (134, 244), and Wells rules for DVT and PE (245, 246), I chose to evaluate the Framingham Wilson CHD risk rule because existing systematic reviews (47, 53), as well as my preliminary searches, suggested that the Framingham Wilson CHD risk rule had more external validation studies than any other cardiovascular CPR.

In November 2015, I conducted a forward citation search of the derivation study of the Framingham Wilson CHD risk rule (56) in Scopus. The initial forward citation search was updated in June 2020. I did not limit the search to any specific language, date range, access type, or document type. Instead of a systematic electronic database search, I chose to use the forward citation search because the robustness of a forward citation search as a method to identify external validation studies has been demonstrated in chapter 3.

4.3.2. Study selection

The references found in the forward citation search were imported into Sente, a reference management software. I screened the title and abstract of all imported references and reviewed the full-text article of relevant references. I included a study, 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 (56) or presented sufficient details about the CPR including baseline survival function, regression coefficient, hazard ratio, and scoring systems (6) for the verification.

The Framingham Heart Study explicitly defines each component of CHD's composite outcome: angina pectoris, acute myocardial infarction, coronary insufficiency, and CHD death (247). For example, angina pectoris is defined as

"brief recurrent chest discomfort of up to 15 minutes duration, precipitated by exertion or emotion and relieved by rest or by nitroglycerin" (247). Since the Framingham Heart Study defined these components of CHD, standard practices for evaluating patients for CHD have evolved (248, 249). When a study evaluated CHD as an outcome, but CHD or its component was assessed differently than the Framingham Heart Study definition, I used my judgement as an Internist who routinely diagnoses and manages cardiovascular disease to determine whether it was sensible. For example, a study by Herrera et al. (250) compared the performance of the Framingham Wilson CHD risk rule and REGICOR for predicting the development of ischemic heart events in patients with HIV infection. One component of ischemic heart event was angina pectoris which was defined as "compatible chest pain history plus a positive exercise stress test, scintigraphy, or coronary arteriogram." Although different from the Framingham Heart Study's original definition, I determined that the outcome definition of angina pectoris by Herrera et al. (250) was sensible because it was consistent with modern standards for evaluating stable patients with chest pain (248, 249).

A systematic review reported that prediction horizons in 311 of 314 (99.0%) CPRs for cardiovascular disease risk were five years or longer (53); there were 72 (22.9%) CPRs with a prediction horizon of five years or longer but shorter than ten years, 209 (66.6%) CPRs with a prediction horizon of 10 years, and 30 (9.6%) CPRs with a prediction horizon longer than ten years (53). Because there is no clearly defined sufficient follow-up duration needed to assess the development of coronary heart disease appropriately, I chose to include studies

with an average follow-up of five years or longer because this was the intended timeframe for predicting cardiovascular disease outcomes in the vast majority of CPRs for cardiovascular disease risk according to the systematic review (53).

I excluded studies that evaluated any previously published, modified versions of the Framingham Wilson CHD risk rule because they were new CPRs different from the Framingham Wilson CHD risk rule.

4.3.3. Data collection

From each study meeting selection criteria, I collected bibliographic information such as the first author's name and the journal that published the study. The publication date was verified in PubMed, and the impact factor of the journal was obtained from the 2019 Journal Citation Reports by Web of Science. In addition, I recorded the name of the cohort, geographic location of the study, number of participants, duration of follow-up, and performance measures. I also assessed whether the study's authors evaluated Framingham Wilson CHD risk equations or scores and how outcomes were defined and ascertained.

I categorised the included studies into two groups:

(1) studies in which one of the stated aims was externally validating Framingham Wilson CHD risk rule, 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 data provided in the study.

For studies in which the aim was externally validating Framingham Wilson CHD risk rule, I assessed whether authors concluded Framingham Wilson CHD risk rule performed adequately or inadequately. I interpreted the authors' conclusions 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 example, I considered that Herrera et al. (250) concluded the performance of the Framingham Wilson CHR risk rule was inadequate based on their statement that "the Framingham risk function significantly overestimated ischemic heart event in HIV infected patients" in the discussion section. For external validation studies where authors concluded Framingham Wilson CHD risk rule performed inadequately, I also examined whether the authors updated the Framingham Wilson CHD risk rule.

I 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. I conducted the forward citation searches, screened the references found, assessed the eligibility, and collected data from included

studies. A second reviewer (Lucy Abel) verified the eligibility of included articles and collected data independently. Any disagreements were adjudicated through discussions.

4.3.4. Data analysis

I calculated the proportion of external validation studies in which authors concluded 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.

I meta-analysed Predicted/Observed (P/O) event ratio and *c* statistic of Framingham Wilson CHD risk rule from studies conducted in the USA and European geographic regions with similar cardiovascular disease burden, the prevalence of risk factors, and health care environment (251): the UK, Northern Europe, Western Europe, and Southern Europe. I chose to evaluate the P/O ratio and *c* statistic because these were the most frequently reported performance measures in external validation studies of cardiovascular CPRs (53). First, P/O ratios were converted to Observed /Expected ratios (O/E). Then, I transformed O/E ratios to natural logarithmic scale and *c* statistics to logit scale. The standard errors for ln (O/E ratio) and logit (*c* statistic) were calculated using methods described by Debray et al. (252). After meta-analysing, the summary estimate of the O/E ratio was converted back to the P/O ratio. Finally, I presented the summary estimate of the P/O ratio and *c* statistic and their 95% prediction intervals. When multiple studies were conducted using data from the

same cohort, only the study with the largest dataset was included in the meta-analyses.

I measured the proportion of eligible external validation studies that did not contribute to understanding the performance of the Framingham Wilson CHD risk rule in meta-analyses to their full potential because either P/O ratio or *c* statistic was not reported or could not be obtained from provided data. An article may contain more than one external validation study. The unit of analysis was an external validation study rather than an article for all analyses in this chapter.

4.3.5. Additional analyses

I carried out a post hoc citation analysis to examine the impact of external validation studies that concluded the Framingham Wilson CHD risk rule performed inadequately but did not update it. In Scopus, I conducted forward citation searches of these external validation studies that concluded the Framingham Wilson CHD risk rule performed poorly but did not update it. Then, I identified studies that derived a new cardiovascular CPR using the Framingham Wilson CHD risk rule's poor performance as one of the justifications and cited these external validation studies. These methods are identical to those I used in chapter 2 to determine whether authors of derivation studies justified creating a new cardiovascular CPR based on existing CPR' insufficiencies with a citation to existing CPR. I excluded studies that evaluated a risk factor of an outcome in a multivariable analysis because authors did not

necessarily intend to develop a new cardiovascular CPR, although the results may contain a CPR.

4.4. Results

4.4.1. External validation studies of the Framingham Wilson CHD risk rule

A total of 97 studies that evaluated the Framingham Wilson CHD risk rule from 83 publications were included, as illustrated in Figure 4.4.1.

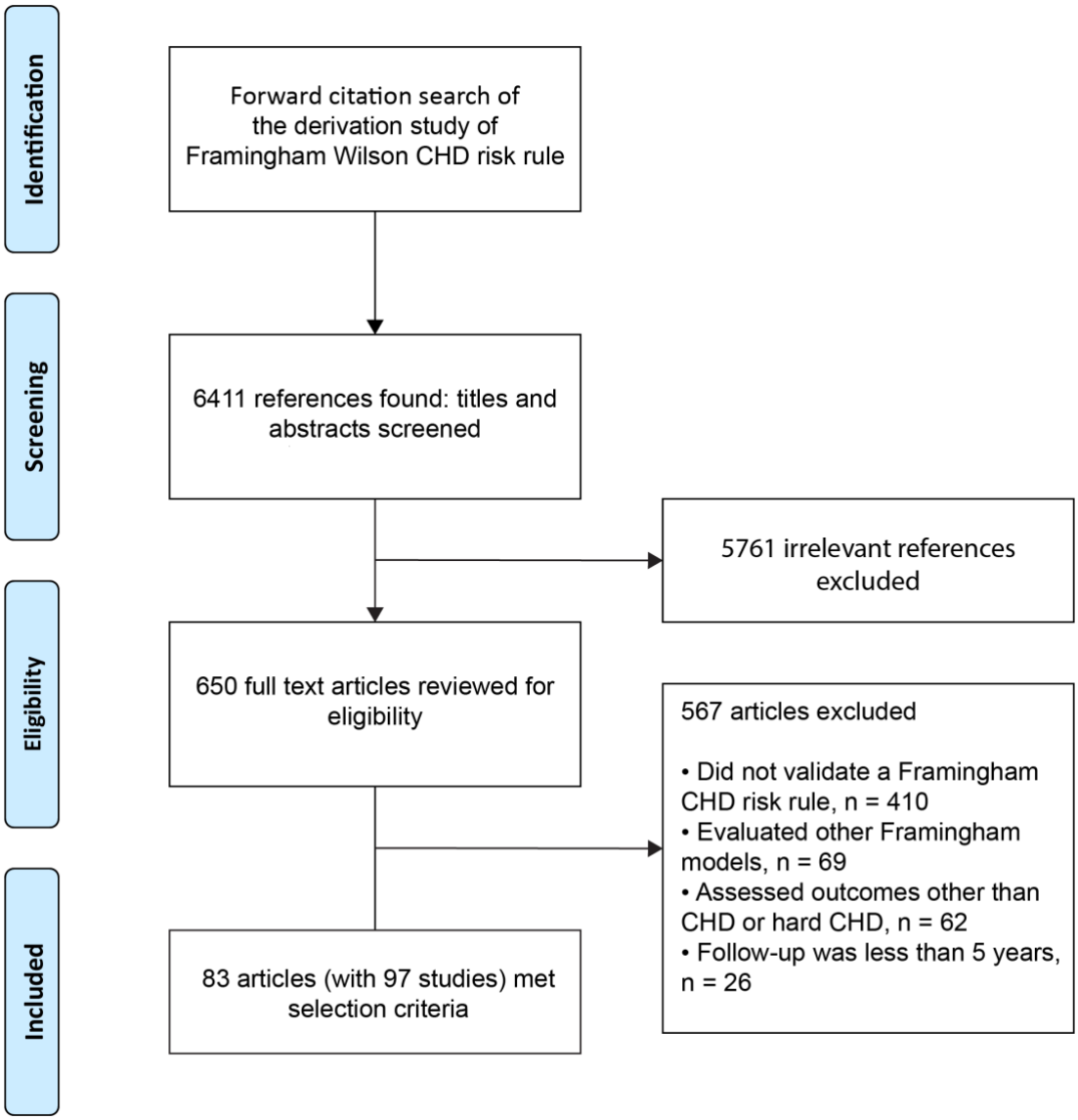


Figure 4.4.1 Search and selection of studies that evaluated the Framingham Wilson CHD risk rule.

Characteristics of the included studies are summarised in Table 4.4.1. For 39 (40.2%) of 97 included studies, one of the aims was externally validating the Framingham Wilson CHD risk rule. For 58 (59.8%) 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.

Table 4.4.1 Characteristics of included studies.

Characteristic	External validation studies of Framingham Wilson CHD risk rule, n = 39	Other studies with a performance measure of Framingham Wilson CHD risk rule, n = 58	All included studies, n = 97
Publication year, median (IQR)	2008 (2006-2012)	2010 (2008-2014)	2009 (2007-2014)
Journal impact factor, median (IQR)	2.941 (1.985-5.864)	8.945 (4.153-23.603)	5.213 (2.740-21.317)
Location			
In the USA	7 (18.0)	21 (36.2)	28 (28.9)
Outside the USA	32 (82.0)	37 (63.8)	69 (71.1)
Sample size, median (IQR)	1,482 (608-5,611)	3,124.5 (1,392-6,604)	2,550 (899-5,732)
Number of events, median (IQR)	133 (61-325)	204 (117-499)	187.5 (92-460.5)
Duration of follow-up in years, median (IQR)	10.0 (7.2-10.0)	8.6 (6.4-10.9)	9.5 (6.4-10.0)
Description of CPR being validated			
Risk equation	7 (17.9)	17 (29.3)	24 (24.7)
Risk score	4 (10.3)	9 (15.5)	13 (13.4)
Unclear	28 (71.8)	32 (55.2)	60 (61.9)
Handling of missing data			
Complete case analysis	30 (76.9)	27 (46.6)	57 (58.8)
Imputation	0 (0.0)	16 (27.6)	16 (16.5)
Unclear	9 (23.1)	15 (25.9)	24 (24.7)

Values are numbers (%) unless indicated otherwise.

External validation studies of Framingham Wilson CHD risk rule had a smaller median sample size and outcome events of 1,482 (IQR, 608-5,611) and 133 (IQR, 61-325) compared with those from studies in which a performance measure of Framingham Wilson CHD risk rule could be obtained, 3,124.5 (IQR,

1,392-6,604) and 204 (IQR, 117-499), respectively. The median impact factor of journals that published external validation studies of Framingham Wilson CHD risk rule was 2.941 (IQR, 1.985-5.864), whereas the median impact factor of journals that published studies with a performance of measure of Framingham Wilson CHD risk rule was 8.945 (IQR, 4.153-23.603). The list of journals that published included studies is provided in Appendix 4.1.

Performance measures of the Framingham Wilson CHD risk rule that were reported by included studies are summarised in Table 4.4.2. Of 39 studies that aimed to externally validate the Framingham Wilson CHD risk rule, only 13 (33.3%) presented a calibration plot, and 6 (15.4%) provided a calibration table as recommended in the TRIPOD statement. For discrimination, *c* statistic was reported with a 95% confidence interval in 16 of 39 (41.0%) external validation studies, adhering to the recommendation in the TRIPOD statement.

Table 4.4.2 Performance measures of the Framingham Wilson CHD risk rule reported by included studies.

Performance measure	External validation studies of Framingham Wilson CHD risk rule, n = 39	Other studies with a performance measure of Framingham Wilson CHD risk rule, n = 58	All included studies, n = 97
Any calibration	22 (56.4)	9 (15.5)	31 (32.0)
P/O ratio with 95% CI	3 (7.7)	0 (0.0)	3 (3.1)
P/O ratio without 95% CI	4 (10.3)	0 (0.0)	4 (4.1)
Calibration plot	13 (33.3)	1 (1.7)	14 (14.4)
Calibration table	6 (15.4)	0 (0.0)	6 (6.2)
Calibration intercept	2 (5.13)	0 (0.0)	2 (2.1)
Calibration slope	1 (3.0)	1 (1.7)	2 (2.1)
Goodness of fit test	9 (23.08)	7 (12.1)	16 (16.5)
Any discrimination	28 (71.8)	36 (62.1)	64 (66.0)

Performance measure	External validation studies of Framingham Wilson CHD risk rule, n = 39	Other studies with a performance measure of Framingham Wilson CHD risk rule, n = 58	All included studies, n = 97
C statistic with 95% CI	16 (41.0)	9 (15.5)	25 (25.8)
C statistic without 95% CI	11 (28.2)	12 (20.7)	23 (23.7)
ROC curve	16 (41.0)	5 (8.6)	21 (21.6)
Discrimination slope	1 (2.6)	1 (1.7)	2 (2.1)
Any classification	9 (23.1)	0 (0.0)	9 (9.3)
Sensitivity or specificity	9 (23.1)	0 (0.0)	9 (9.3)
Positive or negative predictive value	8 (20.5)	0 (0.0)	8 (8.2)
Positive or negative likelihood ratio	5 (12.8)	0 (0.0)	5 (5.2)
Diagnostic odds ratio	4 (10.3)	0 (0.0)	4 (4.1)
Accuracy	4 (10.3)	0 (0.0)	4 (4.1)
Other*	1 (2.6)	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.

4.4.2. Inefficiency from failure to update the Framingham Wilson CHD risk rule

Authors' interpretations of Framingham Wilson CHD risk rule's performance in external validation study and their consequent actions are summarised in Figure 4.4.2. Of 39 external validation studies of the Framingham Wilson CHD risk rule, authors of 26 (66.7%) concluded that the performance was inadequate but did not update the CPR.

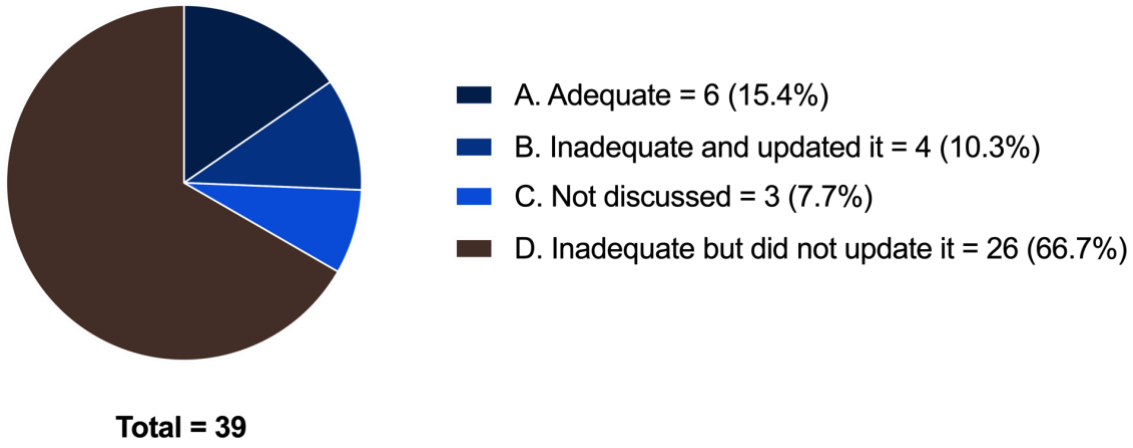


Figure 4.4.2 Authors' interpretations of Framingham Wilson CHR risk rule's performance in external validation study and consequent actions.

In the post hoc citation analysis, I found a total of 1,177 references by conducting forward citation searches of these 26 external validation studies that concluded the performance Framingham Wilson CHD risk rule was inadequate but did not update it. From these references, I identified 18 studies that derived a new cardiovascular CPR using the poor performance of 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 18 new cardiovascular CPRs, three were derived by overlapping authors with external validation studies and 15 were derived by different authors, as presented in Appendix 4.2.

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

A total of 43 studies were included in the meta-analysis of the P/O ratio: 14 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. In 30

studies, the ln (O/E ratio) and standard error were calculated using the ten-year predicted probability of CHD from the Framingham Wilson CHD risk rule and ten-year event rate extrapolated from a different follow-up duration. Only three of these were external validation studies of the Framingham Wilson CHD risk rule. In seven studies (five external validation studies), the ln (O/E ratio) and standard error were obtained using the ten-year predicted and observed event rate. For four external validation studies that presented the predicted and observed event rate at a follow-up duration other than ten years, the ln (O/E ratio) was calculated from extrapolated ten-year predicted and observed event rate. Finally, in two external validation studies, the ln (O/E ratio) and standard error were calculated from the predicted and observed number of CHD events at ten years.

The meta-analysis results of the P/O ratio are presented in Figure 4.4.3. The summary P/O ratio from the meta-analysis of 17 studies conducted in the USA was 1.065 with a 95% prediction interval of 0.231-4.918, as illustrated in Figure 4.4.3-(a). The summary P/O ratios (95% prediction intervals) from the meta-analyses of studies conducted in Europe were 1.632 (0.779-3.418) for the UK, 1.846 (0.393-8.654) for Northern Europe, 1.399 (0.287-6.835) for Western Europe, and 2.232 (1.279-3.896) for Southern Europe, as illustrated in Figure 4.4.3-(b), (c), (d), and (e).

Twenty studies aimed to externally validate the Framingham Wilson CHD risk rule were conducted using data that could potentially contribute to the meta-

analysis of the P/O ratio. Of these, six studies could not be included because they neither reported a P/O ratio nor provided sufficient information about the data to estimate it, as presented in Table 4.4.3.

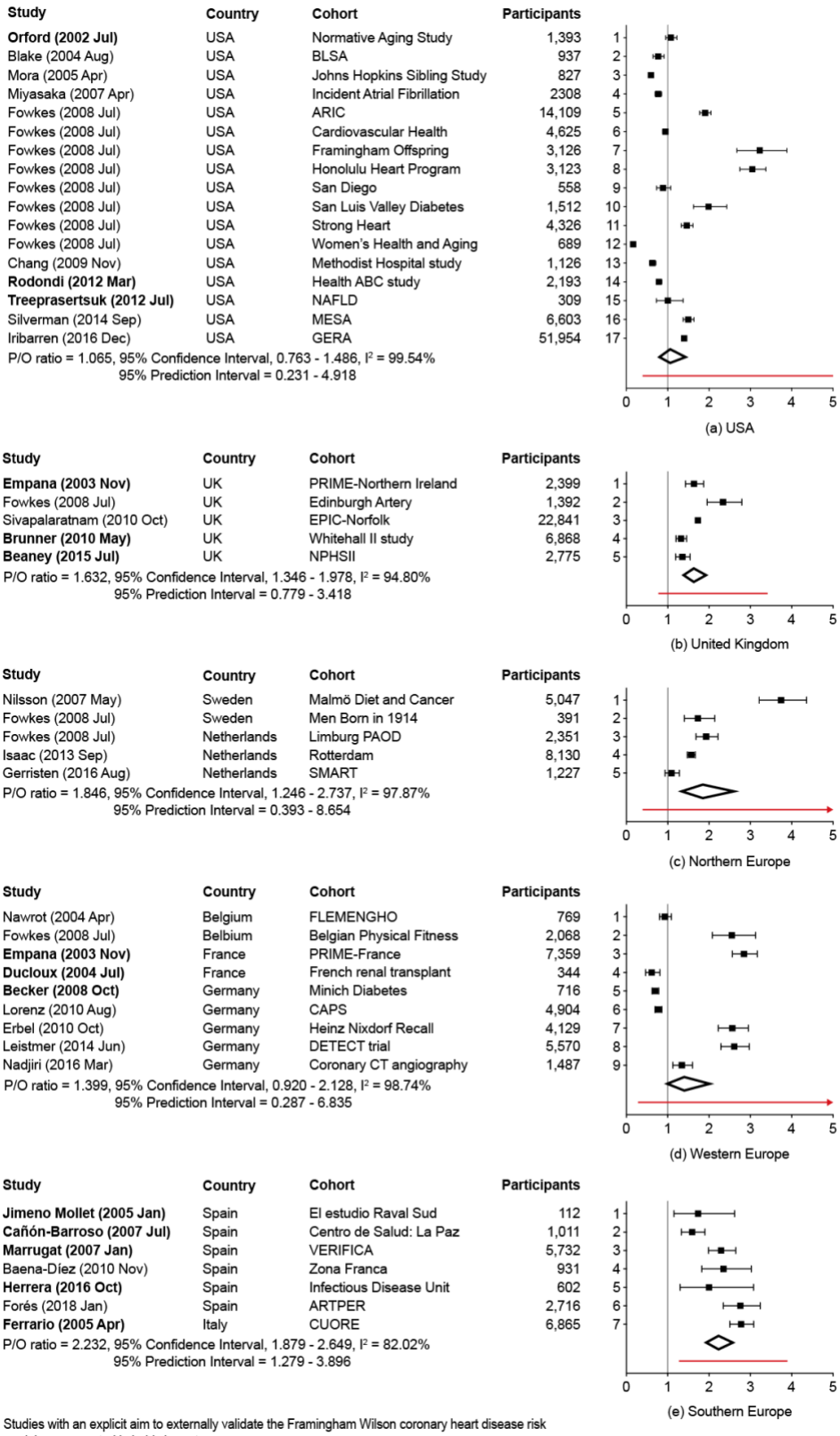


Figure 4.4.3 Meta-analyses of Predicted/Observed (P/O) ratio from studies conducted in the USA and Europe.

Table 4.4.3 Performance measures reported or calculated from data provided in external validation studies of the Framingham Wilson coronary heart disease risk rule.

Study	Location	Cohort	Performance measure reported or calculated from data	
			P/O ratio	*C* statistic
Orford 2002 (253)	USA	Normative Aging Study	Calculated from data	Reported**
Mainous 2007 (254)	USA	ARIC	**Not available***	Reported
Vaidya 2007 (255)	USA	Johns Hopkins Sibling study	Ineligible***	**Not available***
Rodondi 2012 (256)	USA	Health ABC study	Calculated from data	Reported**
Treeprasertsuk 2012 (257)	USA	NAFLD	Calculated from data	**Not available***
Gander 2014 (258)	USA	ACLS	**Not available***	Reported
Emapana 2003 (259)	UK	PRIME-UK	Reported	Reported**
Simmons 2008 (260)	UK	EPIC-Norfolk	Ineligible***	Reported
Brunner 2010 (261)	UK	Whitehall II study	Calculated from data	Reported
Beaney 2015 (262)	UK	NPHSII	Calculated from data	**Not available***
Van Der Heijden 2009 (263)	Northern Europe	Hoorn study	**Not available***	Reported
Merry 2012 (264)	Northern Europe	CAREMA	**Not available***	Reported
Empana 2003 (265)	Western Europe	PRIME- France	Reported	Reported**
Ducloux 2004 (266)	Western Europe	French renal transplant cohort	Calculated from data	**Not available***
Guckelberger 2006 (267)	Western Europe	Charité-Universitätsmedizin Berlin liver transplant	**Not available***	Reported**
Becker 2008 (268)	Western Europe	Munich Diabetes	Calculated from data	Reported
Jimeno-Mollet 2005 (269)	Southern Europe	El estudio Raval Sud	Calculated from data	**Not available***
Ferrario 2005 (270)	Southern Europe	CUORE	Calculated from data	**Not available***
Marrugat 2007 (196)	Southern Europe	VERIFICA	Calculated from data	Reported**
Cañón-Barroso 2007 (271)	Southern Europe	Centro de Salud: La Paz	Calculated from data	Reported
Herrera 2016 (250)	Southern Europe	Infectious Disease Unit at Hospital del Mar	Calculated from data	**Not available***
Protopsaltis 2004 (272)	Southern Europe	Tzaneio General Hospital	**Not available***	Reported**

P/O ratio: Predicted/Observed event ratio.

Study	Location	Cohort	Performance measure reported or calculated from data	
			P/O ratio	*C* statistic

* Not reported by the authors of external validation studies and could not be calculated by me using data provided in the external validation study despite being conducted using data that could contribute to the meta-analyses.
** Reported with a 95% confidence interval.
*** There were other external validation studies conducted using a bigger dataset from the same cohort.

The meta-analysis results of the *c* statistic of the Framingham Wilson CHD risk rule are presented in Figure 4.4.4. The summary *c* statistic and 95% prediction interval from the meta-analysis of seven studies conducted in the USA were 0.695 and 0.459-0.841, Figure 4.4.4-(a). The summary *c* statistic (95% prediction intervals) from the meta-analyses of studies conducted in Europe were 0.694 (0.367-0.899) for the UK, 0.703 (0.303-0.928) for Northern Europe, 0.691 (0.602-0.768) for Western Europe, and 0.686 (0.631-0.737) for Southern Europe, as presented in Figure 4.4.4-(b), (c), (d), and (e).

Twenty-two studies aimed to externally validate the Framingham Wilson CHD risk rule were conducted with data that could potentially contribute to the meta-analysis of *c* statistic. Of these, seven studies could not be included because they neither reported a P/O ratio nor provided data to estimate it, as presented in Table 4.4.3.

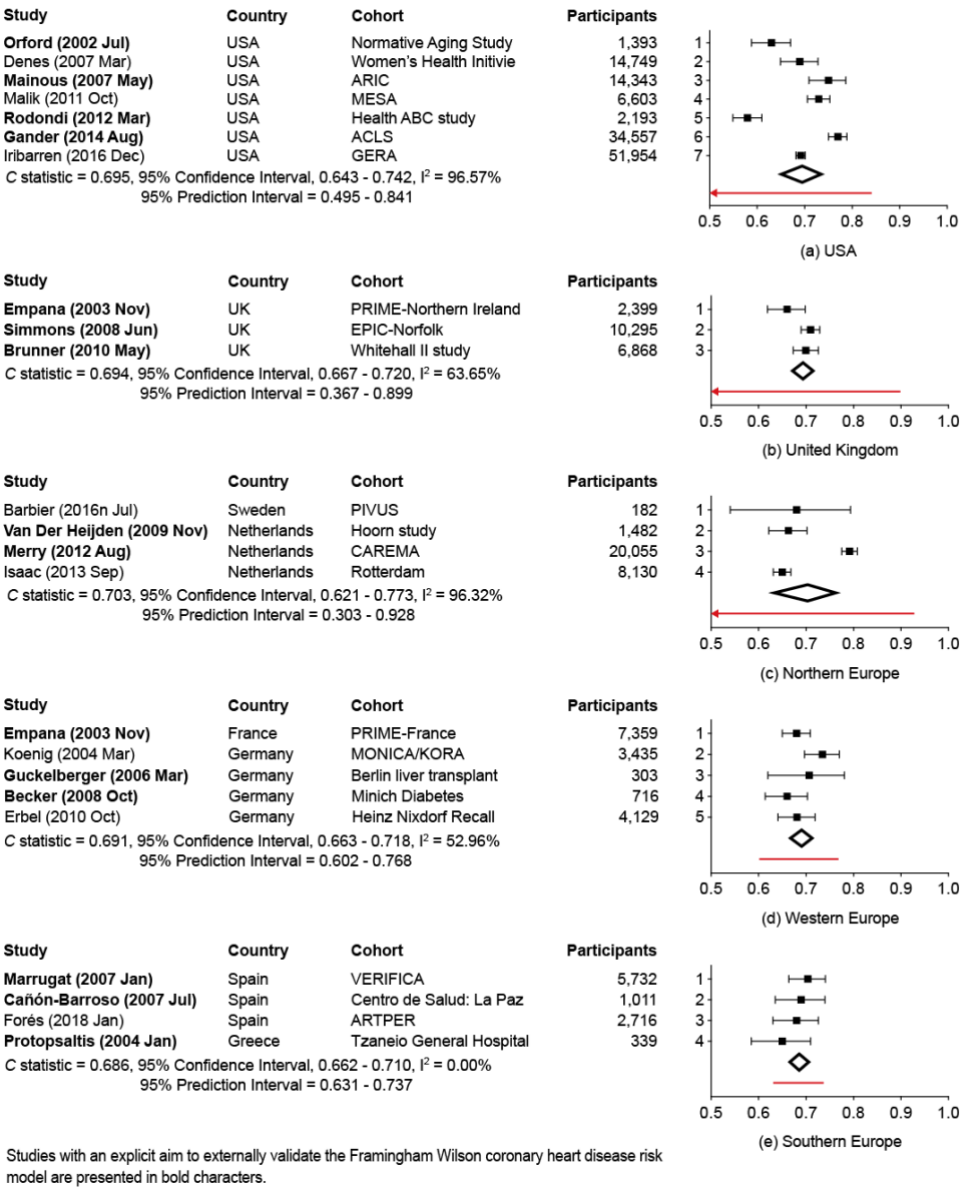


Figure 4.4.4 Meta-analyses of *c* statistic from studies conducted in the USA and Europe.

In summary, 30 of 39 studies that aimed to externally validate the Framingham Wilson CHD risk rule were conducted in the USA and Europe. Eight studies were excluded because there were other studies conducted using a bigger dataset from the same cohort. Therefore, of 22 studies aimed to externally validate the Framingham Wilson CHD risk rule, only nine (40.9%) contributed to

understanding the summary P/O ratio and *c* statistic in eligible meta-analyses. On the other hand, 13 (59.1%) studies aimed to externally validate the Framingham Wilson CHD risk rule did not contribute to eligible meta-analyses to their full potential: six studies for the P/O ratio and seven studies for *c* statistic because they neither reported relevant performance measures nor provided sufficient information about the data to estimate them.

4.5. Discussion

In this chapter, I assessed 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 an overarching population.

4.5.1. Summary of findings

Many (66.7%) external validation studies found that the Framingham Wilson CHD risk rule performed inadequately but did not update the CPR. Subsequently, 18 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.

Most external validation studies of the Framingham Wilson CHD risk rule did not report calibration and discrimination measures as recommended by the

TRIPOD statement. The majority (59.1%) of studies that aimed to externally validate the Framingham Wilson CHD risk rule did not contribute to understanding its performance to full potential because they neither reported relevant performance measures nor provided sufficient information about the data to estimate them.

4.5.2. Comparison with existing literature

Although systematic reviews have pointed out that efforts to externally validate CPRs are disproportionately focused on a small number of CPRs (47, 53), this was the first study to quantify the value of these external validation studies conducted repeatedly for a CPR.

I 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 that showed external validation studies often do not report recommended performance measures, especially calibration (53, 59, 62, 73, 79). In addition, I 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" (6). The results from this chapter that showed performance measures are often presented without confidence intervals are also in line with findings from existing systematic reviews (62, 79, 193).

A systematic review by Damen et al. (236) that summarised the performance of widely used cardiovascular CPRs reported that the Observed/Expected event (O/E) ratio of the Framingham Wilson CHD risk rule was 0.44 (95% CI, 0.31-0.62) for men and 0.43 (95% CI, 0.26-0.72) for women in Europe, which correspond to P/O ratio of 2.27 (95% CI, 1.61-3.23) and 2.33 (95% CI, 1.39-3.85) respectively. These P/O ratios of the Framingham Wilson CHD risk rule are higher than the P/O ratio estimated in European regions from this chapter. The discrepancy might be due to the following reasons. Firstly, summary estimates from random-effects meta-analyses represent the average predictive performance specified by assumed distributions of predictive performance measures across included external validation studies (273). Although the majority of 16 external validation studies included in the systematic review (236) overlapped with external validation studies included in this chapter, the summary O/E ratio obtained in the systematic review only represents an average predictive performance across those external validation studies included. Secondly, the systematic review authors seem to have misinterpreted data from an external validation study by Reissigová and Zvárová (237). The systematic review reported that the O/E ratio from this external validation study was 0.22 (95% CI, 0.18-0.26) (236), which corresponds to a P/O ratio of 4.55 (95% CI, 3.85-5.56), suggesting severe overestimation of CHD risk by the Framingham Wilson CHD risk rule. Contrarily, the external validation study by Reissigová and Zvárová (237) reported that "observed risk of CHD (16.4%) was higher than predicted risk (12.8%) by Framingham Wilson CHD risk rule", indicating the underestimation of CHD risk.

4.5.3. Strengths and Limitations

I found 83 articles with studies that evaluated the Framingham Wilson CHD risk in populations different from those used in the derivation study by conducting a forward citation search. The previously discussed systematic review by Damen et al. (236) used the results of another systematic review by the same research group (53) and searched electronic databases using a strategy developed by an information specialist experienced in Cochrane review to find articles with an external validation study of Framingham Wilson CHD risk rule. In addition, the authors of the systematic review conducted a forward citation search of the derivation study in Scopus to supplement the main search, although limited to articles published after 2013. Despite all these searches, the systematic review authors only identified 16 articles with an external validation study of the Framingham Wilson CHD risk rule. This demonstrates the efficiency and robustness of the forward citation search used in this chapter.

This study has the following limitations. I showed that failing to update the Framingham Wilson CHD risk rule led to deriving potentially redundant cardiovascular CPRs. However, the initial study protocol did not include searching for studies that derived a new cardiovascular CPR using the Framingham Wilson CHD risk rule's poor performance in the external validation study as one of the justifications. I decided to conduct this post hoc citation analysis after discovering that the majority of external validation studies simply dismissed the Framingham Wilson CHD risk rule when it performed poorly, and it was necessary to understand the impact of these studies. This was the first

study that demonstrated research waste and inefficiency in external validation studies of a cardiovascular CPR. However, it is uncertain that findings are generalisable to other CPRs that have many external validation studies.

Some external validation studies of the Framingham Wilson CHD risk rule were conducted using data from the same cohorts. For example, nine external validation studies were produced using data from Centro de Salud: La Paz (271, 274-279), which had a total sample size of 1,011. Although it seemed wasteful to repeatedly publish external validation studies of a prediction rule using a small dataset, it was difficult to assess whether these studies constitute research inefficiency or waste using methods used in this chapter.

4.5.4. Research and clinical implications

Researchers planning to externally validate a CPR should systematically evaluate existing evidence and determine whether a new external validation study is needed. When the need is established, researchers should specify meaningful predictive performance for the clinical context in advance so that the conclusions from the study are not data-driven (30). Researchers should avoid simply dismissing the CPR when it performs less well than the pre-defined threshold in an external validation study. Instead, researchers should attempt to update the CPR and evaluate whether the CPR can reach clinically meaningful predictive performance so that the evidence generated is useful to clinical practice. Also, journal editors should demand that authors of external validation studies provide a clear justification for the study based on a systematic

examination of existing evidence and update the CPR that performs poorly. In addition, authors should present all relevant performance measures in external validation studies so that other researchers and clinicians could use the results of external validation studies to evaluate further and implement the prediction rules.

Conducting a systematic review or using evidence from systematic reviews to determine whether a new external validation study is justified can be challenging (44, 280). Creating a living evidence system for a CPR that continuously updates meta-analyses with the latest evidence of generalizability about the CPR (281) might be useful for researchers considering an external validation study as well as clinicians and other stakeholders of CPR research.

Clinicians should understand a CPR that performed poorly in an external validation study might have an updated version that performs well in similar populations (196, 250, 282).

4.5.5. Conclusion

In this chapter, I have shown that conducting more external validation studies of a CPR does not always make the CPR's evidence of generalizability stronger. Further, I demonstrated that dismissing a CPR when it performed poorly can lead to creating many new, potentially redundant CPRs, which is consistent with the findings from chapter 2 that authors often justify deriving a new CPR by

citing the poor performance of existing CPR. As I argued previously in chapters 2 and 3, cardiovascular CPR research's focus should shift from deriving new CPRs to externally validating existing CPRs and updating them when needed. At the same time, efforts should be taken to ensure evidence that can inform patients, clinicians, researchers, and decision-makers about the generalizability of CPRs is generated in the external validation stage of CPR development. In the next chapter, I evaluate the inefficiency and waste in the impact assessment stage related to design, methods, and reporting.

Chapter 5. Design, methods, and reporting of impact studies of cardiovascular clinical prediction rules

A manuscript based on this chapter was accepted in the Journal of Clinical Epidemiology:

> Ban J-W, 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. Journal of clinical epidemiology. 2021. (283)

5.1. Abstract

Background: Impact assessment is a stage of a clinical prediction rule (CPR) development where the CPR's effectiveness is evaluated. Although existing systematic reviews have demonstrated research waste and inefficiency due to poor design, methods, and reporting among derivation and external validation studies of various CPRs, no systematic review has evaluated these shortcomings for impact studies. Therefore, I systematically evaluate design, methods, and reporting of impact studies of cardiovascular CPRs in this chapter.

Methods: I examined the impact studies of cardiovascular CPRs found in the International Register of Clinical Prediction Rules for Primary Care. I searched for impact studies of these cardiovascular CPRs by conducting forward citation and electronic database searches. Design of impact study was considered

appropriate when randomised or non-randomised experimental study design was used, except for uncontrolled before-after study. For impact studies with appropriate study design, I evaluated the quality of methods and reporting of impact studies using relevant risk-of-bias assessment tools and reporting guidelines. I also compared the quality of methods and reporting between impact and matched control studies.

Results: I found 110 impact studies of cardiovascular CPRs. Of these, 65 (59.1%) used designs classified as inappropriate. Of 45 impact studies with appropriate design, 31 (68.9%) had substantial risk of bias and 26 (57.8%) complied with less than half of reporting domains. Only six (5.5%) were conducted using appropriate study design, avoided substantial risk of bias, and complied with more than half of domains in reporting guidelines. The quality of methods and reporting was not clearly different between impact and matched control studies.

Discussion: Poor design, methods, and reporting were common among impact studies of cardiovascular CPRs. This appears to be a feature of complex interventions rather than specific to CPRs. Nevertheless, users of CPRs should critically evaluate evidence showing the effectiveness of CPRs.

5.2. Introduction

Many clinical prediction rules (CPRs) exist for various cardiovascular conditions, including coronary heart disease, stroke, venous thromboembolism, and congestive heart failure (47, 51, 53). These CPRs may go through derivation, external validation, and impact assessment stage, where distinctive evidence about the CPR is generated (7, 8, 23, 24). In previous chapters, I explored some of the research waste and inefficiency in the derivation and external validation stage. I showed that researchers often do not justify a new cardiovascular CPR by evaluating existing CPRs (94), which could have led to the development of many similar redundant cardiovascular CPRs (53, 95). I also established that most of these cardiovascular CPRs do not have a timely-conducted independent external validation study (96). Furthermore, I demonstrated that many external validation studies do not always generate useful evidence about a CPR's generalizability. Following the usual order of evidence accumulation in CPR development (7, 23), this chapter will focus on impact studies which is a crucial last step before CPRs are implemented in practice because accurate and generalizable CPRs do not always improve care.

Choosing a study design inadequate to address a research question, conducting a study using suboptimal methods, or reporting insufficient information in a publication (36) can lead to waste in healthcare research because findings from studies with these flaws contribute little to clinical decision making or future research (34).

To understand the scope of existing evidence about the design, methods, and reporting of CPR studies, I conducted a simple search in Medline (Ovid) using a strategy presented in Appendix 5.1. I reviewed 2290 references found in the search and identified 18 systematic reviews whose primary aims included evaluating design, methods, or reporting of CPR studies. Table 5.2.1. summarises these systematic reviews

Table 5.2.1 Systematic reviews that evaluated design, methods, or reporting of clinical prediction rule (CPR) studies.

Study	Year	Included in review scope			Study type	Clinical context
		Design	Methods	Reporting		
Wasson et al. (50)	1985	No	Yes	Yes	Derivation	CPRs published in four general medicine journals
Laupacis et al. (4)	1997	No	Yes	Yes	Derivation	CPRs published in four general medicine journals
Mushkudiani et al. (57)	2008	No	Yes	No	Derivation	CPRs for traumatic brain injury
Altman (58)	2009	Yes	Yes	Yes	Derivation and external validation	CPRs for breast cancer
Mallett et al. (73)	2010	No	No	Yes	Derivation	CPRs for cancer
Mallett et al. (71)	2010	Yes	Yes	Yes	Derivation	CPRs for cancer
Collins et al. (59)	2011	No	Yes	Yes	Derivation	CPRs for type 2 diabetes mellitus
Bouwmeester et al. (60)	2012	Yes	Yes	Yes	Derivation, external validation, and impact	CPRs published in six general medicine journals
Hussain et al. (61)	2013	No	Yes	Yes	Derivation	CPRs for mortality after thermal injury

Study	Year	Included in review scope			Study type	Clinical context
		Design	Methods	Reporting		
Collins et al. (62)	2013	No	Yes	Yes	Derivation	CPRs for chronic kidney disease
Collins et al. (79)	2014	No	Yes	Yes	External validation	CPRs published in core clinical journals in 2010
Wen et al. (63)	2016	No	Yes	Yes	Derivation	CPRs for perioperative pancreatic fistula
Ban et al. (233)	2016	Yes	Yes	Yes	External validation	CPRs in systematic reviews
Sahle et al. (64)	2017	No	Yes	Yes	Derivation	CPRs for congestive heart failure
Whittle et al. (284)	2018	No	No	Yes	Derivation	CPRs published in 2015
Heus et al. (65)	2018	No	No	Yes	Derivation and external validation	CPRs published in ten journals with high impact factor
Tan et al. (66)	2019	Yes	Yes	Yes	Derivation and external validation	CPRs for obstetric care
Bradley et al. (67)	2019	Yes	Yes	Yes	Derivation	CPRs for pancreatic cancer

These systematic reviews showed that flaws in design, methods, and reporting were common among derivation studies (4, 50, 57-67, 71, 73, 284) and external validation studies (58, 60, 65, 66, 79, 233) of CPRs. Although a systematic review by Bouwmeester et al. (60) evaluated impact studies, the authors included only three impact studies and examined limited features of study design and methods. Therefore, I systematically evaluate the design, methods, and reporting of studies assessing the impact of cardiovascular CPRs in this chapter.

5.2.1. Objectives

The objectives of this chapter are:

(1) to assess whether impact studies of cardiovascular CPRs used appropriate study design,

(2) to summarise the quality of methods and reporting for impact studies that were conducted using appropriate study design, and

(3) to compare the quality of methods and reporting between impact studies of cardiovascular CPRs with appropriate study design and matched control studies evaluating other types of non-pharmacologic intervention.

5.3. Methods

5.3.1. Information source and search

I evaluated cardiovascular CPRs included in the International Register of Clinical Prediction Rules for Primary Care (47). I chose the International Register as the information source for the following reasons. Firstly, although the systematic review by Damen et al. (53) also contained a large number of cardiovascular CPRs, it only included a narrow scope of CPRs restricted to cardiovascular disease risk in the general population. Therefore, I expected efforts to assess the impact to be limited to a few CPRs such as cardiovascular CPRs derived from Framingham Heart Study (56, 212, 216, 241). Secondly, authors of the International Register updated their search in 2015, and the

updated International Register contained a larger number of unique cardiovascular CPRs than the systematic review by Damen et al. (53). Although the updated version of the International Register was not published, I was granted access to the CPRs whose derivation, validation, or impact studies were included in the cardiovascular domain. Lastly, the updated International Register had CPRs for a wide range of cardiovascular problems in 17 clinical domains such as “K01: Heart pain,” “K22: Risk factor for cardiovascular disease,” K77: Heart failure,” “K90: Stroke/cerebrovascular disease,” and “K93: Pulmonary embolism” (47). Therefore, it was likely that there were many impact studies that evaluated CPRs for various cardiovascular problems.

Forward citation search was the primary method used to identify impact studies of cardiovascular CPRs included in the updated International Register. I used this method because it efficiently identified external validation studies in Chapters 3 and 4. In both chapters, forward citation searches identified a greater number of external validation studies than existing systematic reviews that sought external validation studies using comprehensive electronic database searches (53, 285). Therefore, I expected the forward citation searches to identify impact studies efficiently. However, the forward citation search was used under the assumption that authors of impact studies identified the CPRs by correctly citing their derivation studies.

For cardiovascular CPRs for which the derivation study was included in the International Register, I searched for their impact studies by conducting forward

citation searches of the derivation studies using Scopus (286). For cardiovascular CPRs for which only a validation or impact study was included in the International Register, I first identified their derivation studies and then conducted forward citation searches. It was not possible to conduct forward citation searches of CPRs that were introduced by guidelines and never published in a journal. For these CPRs, I searched for their impact studies in Medline (Ovid) and EMBASE (Ovid) using strategies provided in Appendix 5.2.

All searches were completed on 3 September 2018 without any language restrictions. I searched for impact studies that were published after the most recent reporting guidelines relevant to impact studies with experimental study designs became available. Of those reporting guidelines used in this chapter, the extension of the CONSORT Statement for cluster randomised trials (287) was published last in 2012. Therefore, searches were limited to publication years from 2013.

5.3.2. Eligibility criteria and study selection

I aimed to include studies that sought to assess the impact of using a cardiovascular CPR. A study was eligible for inclusion if:

(1) the aim stated in the title, abstract, or introduction section was to evaluate the impact of an intervention that included using a cardiovascular CPR or promoting the use of it,

(2) the CPR was one of the cardiovascular CPRs included in the updated version of the International Register, and

(3) it assessed the CPR's impact on clinical outcome, care process, practice efficiency, or healthcare provider outcome (23, 91, 288).

I exclude a study if:

(1) it could not be unambiguously determined that a CPR was the main component of the intervention,

(2) it was a validation study that only estimated a hypothetical impact by reporting a predictive performance such as negative predictive value,

(3) it was a secondary analysis or duplicate publication of previously published primary results, or

(4) it was a review article, decision modelling study, or simulation modelling study.

Impact studies often evaluate cardiovascular CPR as a part of a complex intervention. Because determining whether a CPR was the main components of a complex intervention was not always straightforward, a second reviewer (Mei Sum Chan, Tonny Muthee, or Arsenio Paez) independently assessed the eligibility of all impact studies.

5.3.3. Definition of an impact study

Using a modified version of the framework for impact analyses by Wallace et al. (23), I described different types of impact analysis in a broad sense and their aims in the pre-experimental, experimental, and post-experimental phases.

In the pre-experimental phase, studies can explore a prediction rule's potential impact. Acceptability, the likelihood of using a CPR when implemented, may be assessed in a survey or qualitative study (10, 23, 182). For example, a survey by Brehaut et al. (289) showed that emergency physicians viewed the new Canadian C-spine rule as less easy to use than the Ottawa ankle rule, and only 74.2% indicated they would consider using it. In addition, decision analysis can explore the potential effectiveness of using a prediction rule in practice (8, 290). For example, a decision analysis by Zommer et al. (291) showed that a cardiovascular prediction rule for patients with severe mental illness had potential benefits of 15 additional quality-adjusted life years and a cost saving of £53,000 per 1,000 patients, over ten years.

The primary research question in the experimental phase is whether a using clinical prediction rule improves patient outcome, care process, or cost (10, 23). Although it is desirable to use an experimental design in this phase (1, 8, 23), various designs have been used in studies that aimed to evaluate the effectiveness of clinical prediction rules (292). For example, Hemming et al. (293), in a stepped wedge cluster randomised trial, found a targeted case-finding programme using a modified Framingham rule for cardiovascular

disease risk by Anderson et al. (241, 294) increased the prescription of antihypertensives and statins in previously untreated high-risk patients. Similarly, Sheibani et al. (295) used an interrupted time series design to show that the CHA_2DS_2-VASc and HAS-BLED scores (184, 185), incorporated into a computerized decision support system, improved guideline-recommended anticoagulation care.

Once the effectiveness of a prediction rule is proven in the experimental phase, the implementation of the prediction rule and long-term sustained practice improvement can be evaluated in the post-experimental phase. Implementing a prediction rule involves translating the prediction rule into routine clinical practice using strategies such as a clinical decision support system in electronic health record software (2, 23, 182). For example, Bessen et al. (296) showed that implementing the Ottawa ankle rules using a multifaceted strategy, which consisted of education, radiography request form, reminders, audit, and gatekeepers, increased the use of the Ottawa ankle rules and decreased the radiography.

For this chapter, I used a narrow definition of impact study and focused on studies in the experimental phase that aimed to evaluate the effectiveness of using a prediction rule.

5.3.4. Assessment of study design

I assessed whether the design of studies aimed to evaluate the effectiveness of cardiovascular CPRs was appropriate, primarily using definitions consistent with acceptable study designs for the Cochrane Effective Practice and Organisation of Care (EPOC) review (297) and the framework for the impact analysis of clinical prediction rules (23). I also considered discussions from other CPR literature about impact studies (1, 2, 7, 8, 33, 290) to define appropriate and inappropriate study designs for impact studies of CPR. The appropriate and inappropriate study designs for impact studies evaluating the effectiveness of a prediction rule is summarised in Table 5.3.1.

Table 5.3.1 Appropriate and inappropriate study design for studies evaluating the effectiveness of a clinical prediction rule adopted from the Cochrane Effective Practice and Organization of Care review guideline (297).

Study design	Definition
Appropriate design	
A. Randomised experiment	
1. Cluster randomised controlled trial	Researchers randomly allocate groups of individuals to different interventions (8, 273, 297).
2. Stepped wedge cluster randomised trial	Researchers cross over groups of individuals from the usual care to intervention in random order (8, 273, 297).
3. Randomised controlled trial	Researchers randomly allocate individual participants to different interventions (273, 297).
B. Non-randomised experiment	
1. Non-randomised trial	Researchers assign individual participants to interventions using a nonrandom method (297).
2. Interrupted time-series study	Researchers assess an outcome in multiple time points before and after implementing an intervention in a group of individuals (297-299).
3. Repeated measure study	Researchers assess an outcome before and after implementing interventions within the same individuals (8, 297)
4. Controlled before and after study	Researchers compare an outcome assessed in a group of individuals before and after implementing intervention with the outcome assessed in a control group of individuals without implementing the intervention (297, 298).
Inappropriate study design	

1.	Uncontrolled before-after study	Researchers assess an outcome in a group of individuals before and after implementing an intervention without a control group (8, 297, 298)
2.	Observational study	Researchers do not assign interventions but assess exposures and outcomes in a group of individuals (297, 300).
3.	Non-comparative study	Researchers describe what happens to an individual or a group of individuals without a comparator (297, 300).

The ideal study design for evaluating a CPR's impact is a randomised experiment where researchers deliberately and randomly allocate participants to interventions (1, 8, 10, 23, 33, 297). It may involve randomly assigning individual participants (randomised controlled trial) or groups of individuals (cluster randomised trial) to interventions. Randomised control trial is a standard design in studies evaluating the effectiveness of health interventions because properly conducted random allocation to intervention groups can protect against selection bias and confounding (298, 300). For studies evaluating CPRs, randomising individual participants to interventions may not always be the best approach (301). This is because clinicians delivering the interventions cannot be blinded and are susceptible to contamination. Therefore, cluster randomised trials, where groups of individuals (e.g. patients in primary care clinic) are randomised (8, 273, 297), are more desirable than individually randomised controlled trials when evaluating the impact of a CPR (8, 290).

Although the risk of contamination might be reduced in cluster randomised trials (23, 290), the main drawback of this design is inefficiency due to the large sample size required and related costs (290, 301). Alternatively, a stepped-wedge cluster randomised trial, where researchers cross over groups of

individuals from the usual care to intervention in random order, can be used to evaluate the impact of a CPR (8, 273, 297). This design is more efficient than cluster randomised trials because it requires fewer clusters (290, 302). On the other hand, because groups of individuals crossover from the usual care to intervention, a stepped-wedge cluster randomised trial takes longer to complete and can be susceptible to contamination and time trend (273, 302).

Conducting randomised experiments might not always be practical for various reasons: they can be expensive, take too long, or be unsuitable for some clinical contexts (303-305). Certain non-randomised experimental study designs are described in the prediction rule literature as acceptable alternatives to randomised experiments (1, 8, 10, 23, 91, 297) when a randomised experiment is not feasible. These study designs include non-randomised trial, interrupted time series study, repeated measures study, and controlled before-after study.

In non-randomised trials, researchers assign participants to interventions using a nonrandom method (e.g. alternating assignment) (297). Although there is a greater risk of bias (e.g. selection bias) due to the lack of randomisation than randomised experiments (273, 300), having a concurrent control group and uniform outcome assessment makes this design more rigorous than observational and non-comparative study design (300, 306, 307). In interrupted time-series studies, researchers assess the impact of an intervention by measuring the outcome in multiple time points before and after implementing the intervention (297-299). The primary advantage of this design is that it allows

preexisting trends in the outcome to be accounted for by analysing the outcome in multiple time points as the implementation of the intervention happens in routine clinical practice (298, 299, 308). However, some of the methods require advanced statistical knowledge (e.g. multistage regression), which is the main drawback of this design (299).

Repeated measures studies involve multiple assessments of an outcome within the same individuals before and after different interventions are applied (8, 297, 309, 310). This design is simple and less expensive to implement than other non-randomised experimental designs (8), but the outcome assessment might be influenced by the order of interventions applied (309, 310). In controlled before-after studies, an outcome is assessed in a group of individuals before and after receiving an intervention and is compared with the outcome assessed in control group individuals who did not receive the intervention (23, 297, 298). Having a control group can allow this design to account for existing trends in outcome over time (298).

In summary, non-randomised trial, interrupted time series study, repeated measures study, and controlled before-after study are experimental study designs with a control group. These study designs can be used to evaluate the impact of prediction rules when a randomised experiment is not feasible (297, 298), understanding they might be susceptible to bias due to confounding (305, 311).

Other study designs such as uncontrolled before-after study, observational study, and non-comparative study are inappropriate for evaluating a CPR's impact (10, 33, 297). When a study assesses an outcome in a group of participants before and after receiving an intervention without a control group, it is referred to as an uncontrolled before-after study. This design is discouraged because it is difficult to determine whether an improvement in the outcome is attributable to the preexisting trends, participants' awareness of taking part in the study, or intervention (8, 290, 297, 298, 312, 313). Some journals even explicitly recommend against using uncontrolled before-after study design (297, 314). Furthermore, nonexperimental designs, such as observational and non-comparative studies should be avoided because of inevitable bias and confounding (91, 297, 298, 306, 307). Meta-epidemiological studies have shown that uncontrolled before-after studies and observational studies may overestimate the effects of interventions (315-317).

5.3.5. Assessment of methods

For impact studies of cardiovascular CPR with appropriate study design, I further evaluated whether the impact studies were conducted using adequate methods and avoided the substantial risk of bias in their results. To assess the quality of methods in impact studies, I used the Cochrane Collaboration's instruments for evaluating the risk of bias: the revised Cochrane risk-of-bias tool for randomised trials (RoB 2) (318) and additional considerations for cluster-randomised trials (319) for randomised experiments, and the Risk Of Bias In

Non-randomised Studies of Interventions (ROBINS-I) tool (320) for non-randomised experiments.

I used the Cochrane Collaboration's instruments to assess the impact studies' risk of bias for the following reasons. Firstly, all three Cochrane Collaboration's tools for assessing the risk of bias are comparably structured because they were all developed based on the 2011 version of the Cochrane risk-of-bias tool for randomised trials (321), as compared in Table 5.3.2. Although these tools address different bias domains relevant to study designs in the "Pre- or at-intervention stage", they all address parallel bias domains in the "Post-intervention stage". Secondly, these tools share common features in the process of risk-of-bias assessment such as relating risk-of-bias assessment to a specific intervention effect (e.g. intention-to-treat effect or per-protocol effect), using signalling questions for the domain-level risk-of-bias judgement, methods for reaching an overall risk-of-bias decision, and the presentation of the risk-of-bias assessment. Lastly, although ROBINS-I assesses risk-of-bias domains only relevant to non-randomised studies of interventions in "Pre- or at-intervention stage", risk of bias assessed using ROBINS-I is comparable to the risk of bias assessed using other Cochrane risk-of-bias tools because it assesses the risk of bias in reference to a hypothetical, ideally-conducted, target randomised experiment (320).

Table 5.3.2 Bias domains in the Cochrane Collaboration's risk of bias assessment tools.

Stage of study	The revised Cochrane risk-of-bias tool		The Risk Of Bias In Non-randomised Studies of Interventions (ROBINS-I) tool (320)
	The revised Cochrane risk-of-bias tool for randomised trials (RoB 2) (318)	RoB 2: additional considerations for cluster-randomised trials (319)	
Pre- or at-intervention stage	• Bias arising from the randomisation process	• Bias arising from the randomisation process • Bias arising from the timing of identification and recruitment of individual participants in relation to timing of randomization	• Bias due to confounding • Bias in selection of participants into the study • Bias in classification of interventions
Post-intervention stage	• Bias due to deviations from intended interventions • Bias due to missing outcome data • Bias in measurement of outcome • Bias in selection of the reported result	• Bias due to deviations from intended interventions • Bias due to missing outcome data • Bias in measurement of outcome • Bias in selection of the reported result	• Bias due to deviations from intended interventions • Bias due to missing data • Bias in measurement of outcome • Bias in selection of the reported result

For the bias due to deviations from the intended interventions domain, I assessed the risk-of-bias for intention-to-treat effect because impact studies of CPR typically aim to evaluate the effect of assigning participants to intervention rather than the effect of assigning participants to intervention and adhering to it

(319). When several outcomes were reported in an impact study, I assessed the risk of bias for the primary outcome's first reported result. Risk of bias is categorized by three levels (low, some concern, high) in the revised Cochrane risk-of-bias tools, and four levels (low, moderate, serious, critical) in ROBINS-I. I considered the risk of bias was substantial and the results were untrustworthy, when the assessment using the revised Cochrane risk-of-bias tools was high and when the assessment using ROBINS-I was serious or critical.

5.3.6. Assessment of reporting

I also examined whether impact studies with appropriate study design complied with reporting guidelines that were available before 2013. For each study design, I constructed an instrument for evaluating the quality of reporting using criteria from all relevant reporting guidelines because multiple reporting guidelines were applicable to certain study design. Criteria that were not always applicable were excluded. Each instrument consists of 21 reporting domains, as presented in Appendix 5.3. The instrument for assessing the reporting of randomised controlled trials was constructed using criteria from the CONSORT 2010 Statement for reporting parallel group randomised trials (322) and the extension of the CONSORT Statement for randomised trials of non-pharmacologic treatments (323). The instrument for assessing the reporting of cluster randomised trials was constructed using criteria from the CONSORT 2010 Statement for reporting parallel group randomised trials (322), the extension of the CONSORT Statement for randomised trials of non-pharmacologic treatments (323), and the extension of the CONSORT

Statement for cluster randomised trials (287). For the instrument for assessing the reporting of non-randomised intervention studies, criteria from the Transparent Reporting of Evaluations with Non-randomised Designs (TREND) statement (324) were used.

For criteria in the reporting guidelines that have multiple elements, I evaluated whether authors reported all elements. For example, there are two elements in the reporting of blinding: “who was blinded after assignment to interventions” and “how blinding was implemented.” Therefore, I assessed whether impact studies reported both elements of blinding. The reporting of blinding was considered incomplete when impact studies did not clearly report either who was blinded or how blinding was achieved.

For each impact study, I also checked the following additional items that were mentioned in the CONSORT 2010 Statement for reporting parallel group randomised trials but were not one of the checklist items (322): whether the justification for the new study was provided with a citation to a systematic review of similar studies (or notion that such study was absent), whether results were interpreted with a systematic examination of existing evidence, and whether participant flow was presented using a flow diagram. Further, I checked items from “Other information” section of the CONSORT 2010 Statement for reporting parallel group randomised trials (322): whether registration number and name of the trial registry were provided, whether it was explained how the

study protocol could be accessed, and whether the sources of funding and funders' role were described.

Although these were not key items included in the checklist, I chose to evaluate them for all impact studies with appropriate study design for the following reasons. Firstly, they were relevant to both randomised and non-randomised experimental study designs. Secondly, some of these were advocated as measures to reduce waste in research. For example, it has been argued that using a systematic review to justify a new study and interpreting the results of the new study in the context of existing evidence could prevent unnecessary or potentially unethical new studies (38, 44, 154, 158, 325). Thirdly, research papers infrequently include a flow diagram or sufficient information within the flow diagram (326) even though it has been shown that providing a flow diagram improves the reporting quality of participant flow (327). Lastly, information about study registration, protocol, and funding is essential when assessing the risk of bias (320, 321, 328, 329).

5.3.7. Comparison with matched controls

For each eligible impact study with appropriate study design, I searched for a control study matched by intervention type, study design, journal, and publication date. I identified a control study that evaluated the effectiveness of a non-pharmacologic intervention, used an identical design as the relevant impact study, and was published in the same issue of the journal where the impact study was published. If a control study meeting these criteria was not found in

the same issue of the journal, preceding and succeeding issues were searched. The study with the closest publication date to the publication date of the impact study was selected. I excluded studies that were published before 2013, evaluated an intervention in non-human subjects, or assessed the impact of an intervention that included using a CPR or promoting its use.

Using the procedures identical to those used to evaluate the quality of methods and reporting of impact studies as outlined in sections 5.3.4 and 5.4.5, I assessed the methods and reporting of matched control studies. All impact and matched control studies were randomly assigned to one of three reviewers (Mei Sum Chan, Tonny Muthee, and Arsenio Paez) who then assessed the risk of bias and reporting independently. I attempted to resolve disagreements in the risk of bias and reporting assessments with the independent reviewer through discussions. A third reviewer adjudicated any disagreements that could not be resolved through discussion.

5.3.8. Analysis

I reported the number and proportion of impact studies using appropriate and inappropriate study designs. For impact studies that were conducted using appropriate study designs, I presented the proportion of studies with a substantial risk of bias and the mean number of reporting domains that studies complied with. I compared the proportion of impact and control studies with a substantial overall risk of bias using McNemar's test for paired proportions. I

used paired *t*-test to compare the mean number of reporting domains that impact and control studies complied with.

5.4. Results

5.4.1. Impact studies of cardiovascular CPR

The search and selection of impact studies are summarised in Figure 5.4.1. The updated International Register included derivation, validation, or impact studies of 512 unique variations of 202 cardiovascular CPRs. I screened a total of 42,769 references, 40,644 from forward citation searches of 194 derivation studies of cardiovascular CPRs and 2,125 from electric database searches of four cardiovascular CPRs introduced by guidelines that were never published in a journal. I found a total of 110 impact studies of cardiovascular CPRs after reviewing 788 full-text articles.

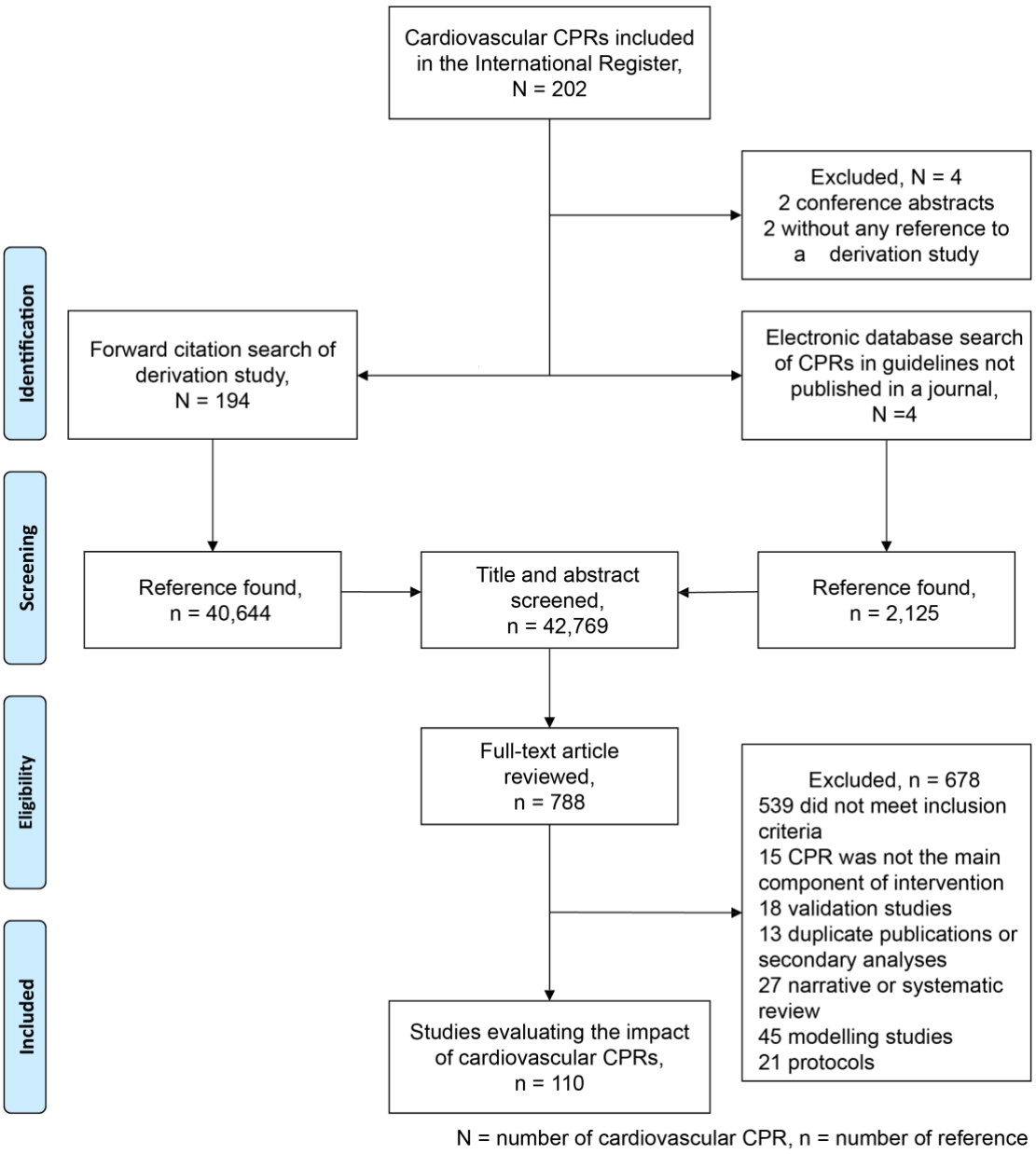


Figure 5.4.1 A flow diagram summarizing the search and selection of impact studies of cardiovascular clinical prediction rule (CPRs).

The cardiovascular CPR most commonly evaluated in impact studies was the Framingham general cardiovascular risk score by D'Agostino (216) with 20 impact studies, followed by CHA_2DS_2-VASc score (184) with 13 studies, Wells rule for pulmonary embolism (123, 246, 330) with 11 studies, and $CHADS_2$

score (243) with 10 studies, and HAS-BLED score (185) with 9 studies, as presented in Table 5.4.1.

Table 5.4.1 Cardiovascular clinical prediction rules evaluated in impact studies.

Subdomain of International Register (47)	Clinical prediction rule	Number of impact study
K22: Risk factor cardiovascular disease	Framingham cardiovascular risk score by D'Agostino (216)	20
	Other Framingham risk scores (56, 212, 241, 279)	8
	American College of Cardiology/American Heart Association pooled cohort equations (235)	5
	World Health Organization/International Society of Hypertension risk prediction charts (331)	4
	Prospective Cardiovascular Münster (PROCAM) score (332)	3
	Systematic Coronary Risk Evaluation (SCORE) charts (129)	3
	UK Prospective Diabetes Study (UKPDS) risk engine (125)	3
	QRISK (112) and QRISK2 (20) scores	3
	New Zealand cardiovascular risk charts (333)	2
	ASSIGN score (21)	1
	Copenhagen risk score [a] (334)	1
	Cambridge risk score [a] (335)	1
	Finnish diabetes risk score (FINDRISC) (336)	1
K90: Stroke / cerebrovascular accident	CHA_2DS_2-VASc score (184)	13
	$CHADS_2$ score (243)	10
	HAS-BLED score (185)	9
	$ABCD^2$ score (244)	7
	HINTS examination (337)	1
	SAMe-TT_2R_2 score (338)	1
K93: Pulmonary embolism	Wells rule for pulmonary embolism (123, 246, 330)	11
	Pulmonary embolism rule-out criteria (PERC) (141)	3
	Christopher study rule for pulmonary embolism (339)	1
	Pulmonary Embolism Severity Index (PESI) (133)	1
	Attribute-matching system for pulmonary embolism by Kline (340)	1
	Wells rule for deep vein thrombosis (245, 341)	2

Subdomain of International Register (47)	Clinical prediction rule	Number of impact study
K94: Phlebitis / thrombophlebitis	Upper extremity deep venous thrombosis score by Constans (342)	1
	Diagnostic score for deep vein thrombosis by Oudega (147)	1
K74: Ischemic heart disease with angina	Attribute-matching system for acute coronary syndrome by Kline (343)	2
	Duke treadmill score (136)	1
	Chest pain rule for coronary artery disease by Bösner (344)	1
K75: Acute myocardial infarction	TIMI risk score (345)	5
K01: Heart pain	Vancouver chest pain rule (135)	1
K84: Heart disease other	Termination of resuscitation guideline (346)	1

5.4.2. Study design of impact study

Of 110 impact studies of cardiovascular CPRs included in the systematic review, study designs classified as appropriate were used in 45 (40.9%) impact studies: 35 randomised experiments and 10 non-randomised experiments, whereas inappropriate study designs were used in 65 (59.1%) impact studies. Characteristics of these impact studies are summarised in Table 5.4.2. Statistically significant results were reported in 71 of 90 (78.9%) impact studies that presented a primary outcome with a statistical hypothesis test using a *p*-value or confidence interval (CI): 25 of 35 (71.4%) randomised experiments, 7 of 9 (77.8%) non-randomised experiments, and 39 of 46 (84.8%) studies with an inappropriate design that presented a primary outcome with a statistical hypothesis test.

Table 5.4.2 Characteristics of impact studies of cardiovascular clinical prediction rule with appropriate and inappropriate study design.

Characteristic	Appropriate study design, n = 45		Inappropriate study design, n = 65	Total, n = 110
	Randomized experiment, n = 35 • 23 randomized controlled trials • 12 cluster randomized trials	Non-randomized experiment, n = 10 • 2 non-randomized trials • 4 interrupted time series studies • 4 repeated measures studies	• 40 uncontrolled before-after studies • 7 cohort studies • 18 non-comparative studies	
Publication date				
Median, IQR	September 2015, April 2014 – December 2016	November 2016, July 2015 – February 2018	August 2016, June 2014 – September 2017	March 2016, May 2014 – August 2017
Journal impact factor				
Median, IQR	4.5, 2.7 – 11.7	2.1, 1.3 – 3.3	1.8, 0.8 – 4.0	2.6, 1.3 – 4.5
Location				
North America	13 (37.1)	4 (40.0)	26 (40.0)	43 (39.1)
Europe	12 (34.3)	4 (40.0)	20 (30.8)	36 (32.7)
Oceania	7 (20.0)	1 (10.0)	7 (10.8)	15 (13.6)
Asia	2 (5.7)	1 (10.0)	7 (10.8)	10 (9.1)
Central and South America	0 (0.0)	0 (0.0)	3 (4.6)	3 (2.7)
Africa	1 (2.9)	0 (0.0)	1 (1.5)	2 (1.8)
International	0 (0.0)	0 (0.0)	1 (1.5)	1 (0.9)
Setting [a]				
Outpatient	17 (48.6)	7 (70.0)	21 (32.3)	45 (40.9)
ED	6 (17.1)	1 (10.0)	19 (29.2)	26 (23.6)
Inpatient	1 (2.9)	1 (10.0)	8 (12.3)	10 (9.1)
Community	6 (17.1)	0 (0.0)	4 (6.1)	10 (9.1)
Pharmacy	1 (2.9)	0 (0.0)	4 (6.1)	5 (4.5)
Multiple	3 (8.6)	0 (0.0)	4 (6.1)	7 (6.4)
Other	1 (2.9)	1 (10.0)	5 (7.7)	7 (6.4)
Sample size [b]				
Median, IQR	570, 240 – 1723	373, 144 – 832	353.5, 200 – 1556	499.5, 208.5 – 1611.5
Primary outcome				
Statistical test	35 (100.0)	9 (90.0)	46 (70.8)	90 (81.8)
No statistical test	0 (0.0)	1 (10.0)	19 (29.2)	20 (18.2)

Values are numbers (%) of impact studies unless indicated otherwise.

[a] Other settings were emergency medical technician service, health insurance, medical conference, military, and online.

[b] One interrupted time series study and one uncontrolled before-and-after study did not report sample size.

5.4.3. Risk of bias in impact studies with appropriate study design

The results of the risk-of-bias assessment for impact studies with appropriate study design are summarised in Table 5.4.3. The overall risk of bias was substantial in 31 of 45 (68.9%) impact studies with appropriate study design: 14 of 23 (60.9%) randomised controlled trials, 8 of 12 (66.7%) cluster randomised trials, and 10 of 10 (100.0%) non-randomised experiments. The distribution of risk-of-bias judgement within each bias domain is illustrated using unweighted bar plots in Figure 5.4.2.

Table 5.4.3 Impact studies with appropriate study design but with substantial risk of bias.

Stage of study	Risk of bias domain	Randomized experiment		Non-randomized experiment [c], n = 10	Total, n = 45
		Randomized controlled trial [a], n = 23	Cluster randomized trial [b], n = 12		
Pre-intervention or at-intervention	1. Randomization process	0 (0.0)	0 (0.0)	N/A	N/A
	2. Timing of identification and recruitment of individual participants	N/A	4 (33.3)	N/A	N/A
	3. Confounding	N/A	N/A	10 (100.0)	N/A
	4. Selection of participants	N/A	N/A	1 (10.0)	N/A
	5. Classification of intervention	N/A	N/A	0 (0.0)	N/A
	Stage	0 (0.0)	4 (33.3)	10 (100.0)	14 (31.1)
Post-intervention	6. Deviation from the intended intervention	9 (39.1)	2 (16.7)	1 (10.0)	12 (26.7)
	7. Missing data	4 (17.4)	1 (8.3)	0 (0.0)	5 (11.1)
	8. Measurement of outcome	4 (17.4)	1 (8.3)	7 (70.0)	12 (26.7)
	9. Selection of reported result	2 (8.7)	2 (16.7)	0 (0.0)	4 (8.9)
	Stage	14 (60.9)	6 (50.0)	7 (70.0)	27 (60.0)
	Overall risk of bias	14 (60.9)	8 (66.7)	10 (100.0)	31 (68.9)

Values are numbers (%) of impact studies.
[a] Studies with high risk of bias assessed by the revised Cochrane risk-of-bias tool for randomized trials (RoB 2.0).
[b] Studies with high risk of bias assessed by additional considerations for cluster-randomized trials in the revised Cochrane risk-of-bias tool for randomized trials (RoB 2.0).
[c] Studies with serious or critical risk of bias assessed by the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool.

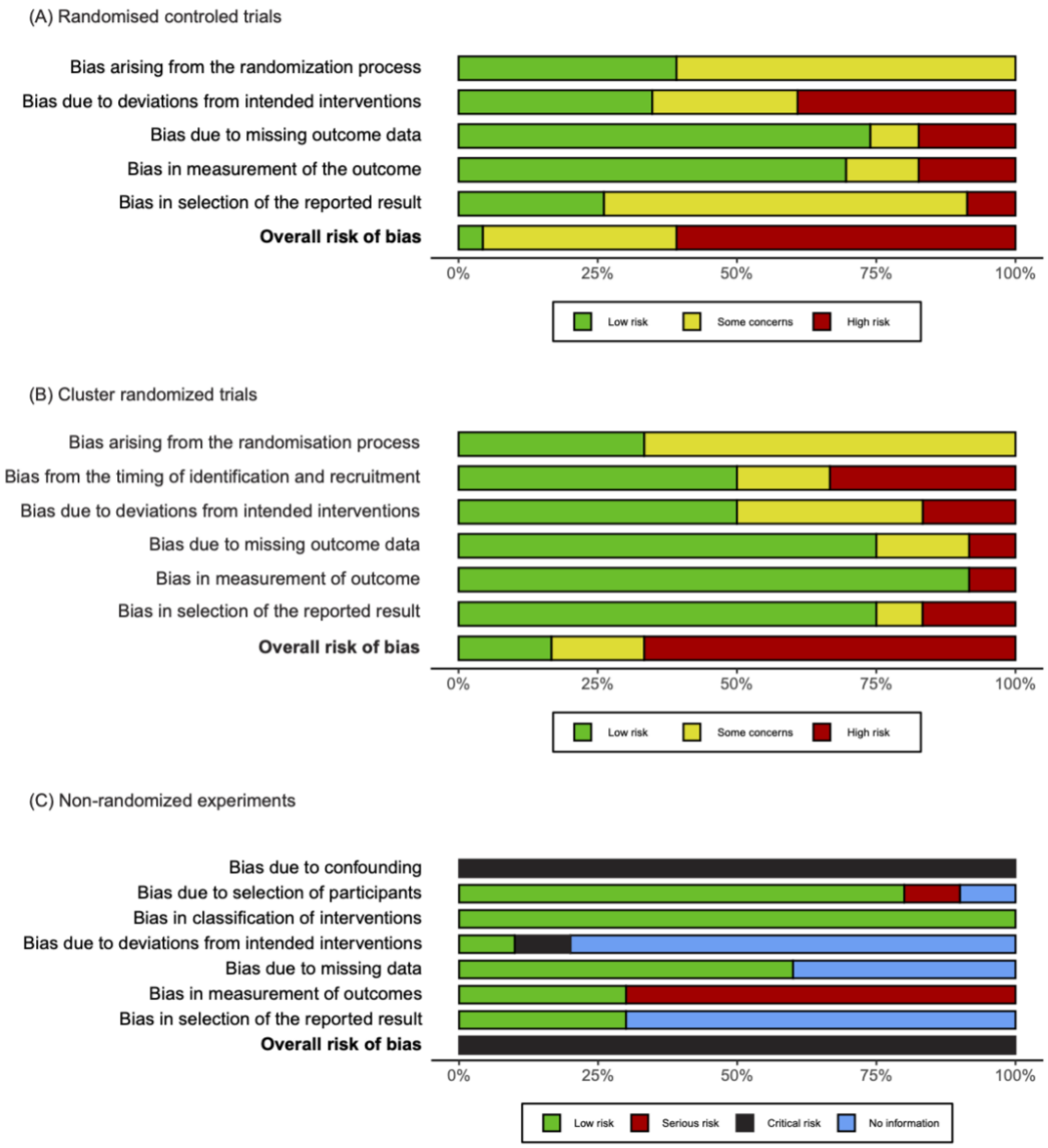


Figure 5.4.2 Distribution of risk-of-bias judgement in bias domains for (A) randomised controlled trials, (B) cluster randomised trials, and (C) non-randomised experiments.

5.4.4. Reporting of impact studies with appropriate study design

Of 45 impact studies with appropriate study design, 26 (57.8%) complied with less than half of reporting domains. Figure 5.4.3 illustrates the distribution of the number of reporting domains that impact studies of cardiovascular CPRs complied with. The mean number of reporting domains that impact studies with

appropriate study design complied with was 10.2 of 21 domains (95% CI, 9.3 and 11.1): 11.4 (95% CI, 10.1 and 12.7) for randomised controlled trials, 8.7 (95% CI, 7.0 and 10.4) for cluster randomised trials, and 9.3 (95% CI, 8.4 and 10.2) for non-randomised experiments.

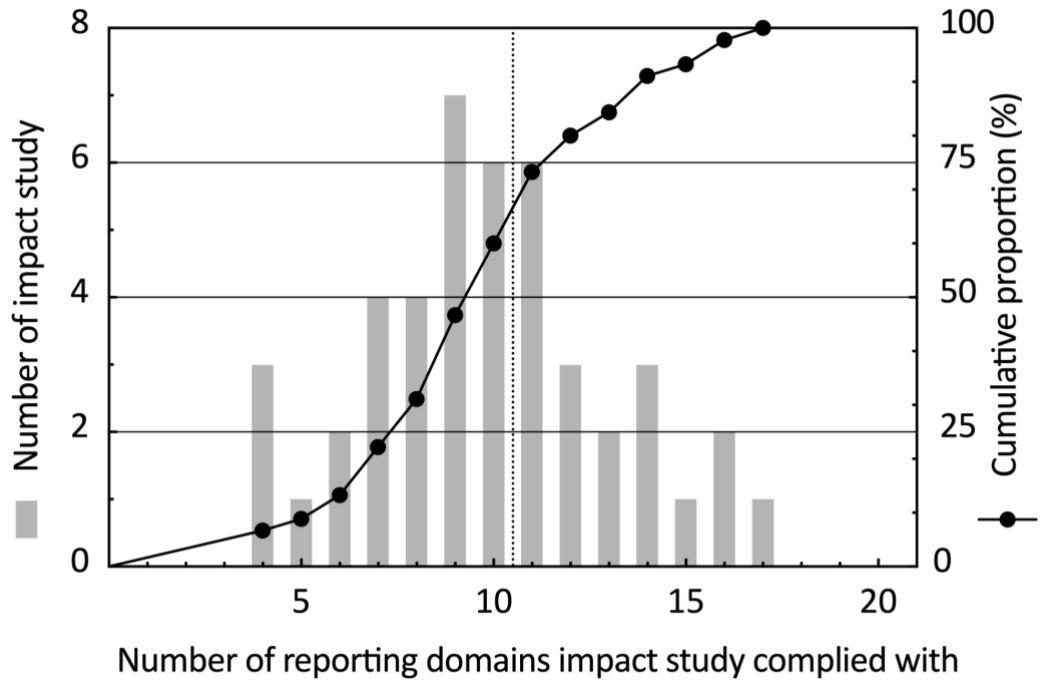


Figure 5.4.3 Frequency and cumulative distribution of the number of reporting domains that impact studies of cardiovascular clinical prediction rules complied with.

Only 24 (53.3%) impact studies justified the new study with a citation to a systematic review of similar studies (or noted that such study was absent) and 17 (37.8%) impact studies interpreted results with a systematic examination of existing evidence. Participant flow was presented using a flow diagram in 33 (73.3%) impact studies. Registration number and name of the trial registry were provided in 24 (53.3%) impact studies, how the study protocol can be accessed

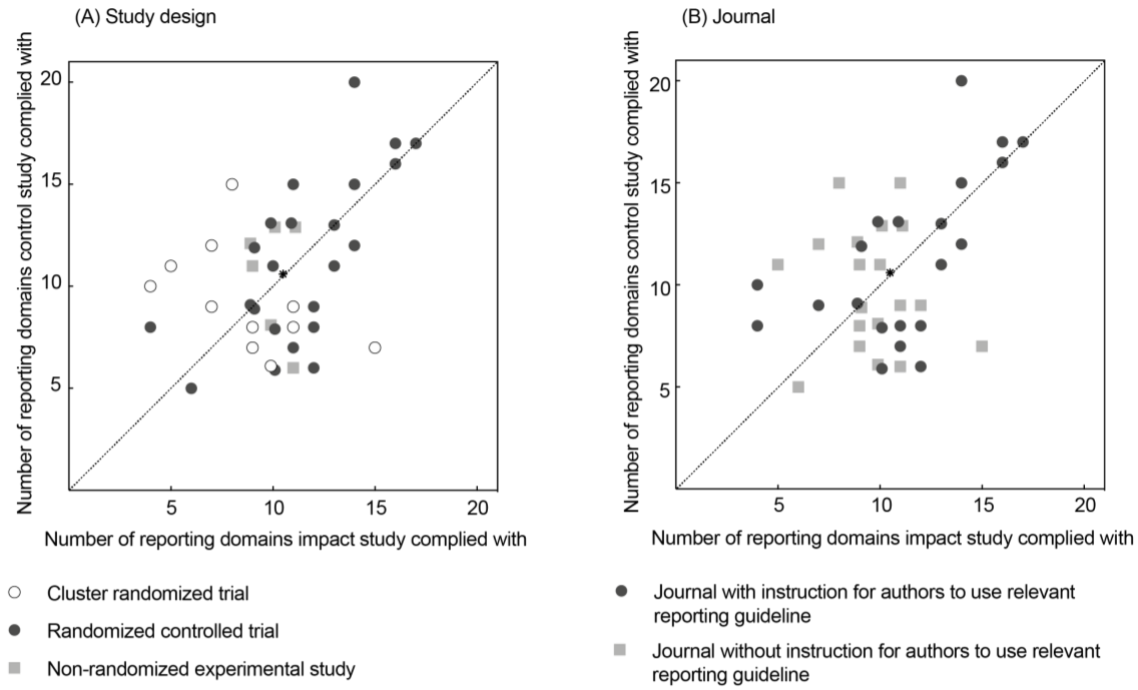
was explained in 17 (37.8%) impact studies, and the sources of funding and funders' role were described in 29 (64.4%) impact studies.

Of 110 impact studies of cardiovascular CPRs, only six (5.5%) were conducted using appropriate study design, avoided substantial risk of bias, and complied with more than half of domains in reporting guidelines (347-352).

5.4.5. Comparison of risk of bias and reporting with matched control studies

For 40 of 45 impact studies with appropriate study design, a matched control study was found. For one cluster randomised trial (353) and four repeated measures studies (354-357), no matched control study was found. The median interval between the publication of an impact study and the matched control study was 0 month (Interquartile range, -3 and 1).

In the matched analysis of 40 impact and 40 control studies, substantial risk of bias was detected in 27 (67.5%) impact studies and 28 (70.0%) control studies. There was no clear difference in the proportion of impact and control studies with substantial risk of bias (difference = -2.5%, 95% CI, -27.4 and 22.4, p = 0.83). Similarly, the mean number of reporting domains that impact and control studies complied with was not clearly different (difference = -0.2, 95% CI, -1.3 and 1.0, p = 0.79), Figure 5.4.4. The full results of risk-of-bias judgement and reporting assessment are provided in Appendix 5.4 and Appendix 5.5.



The asterisk (*) represents the mean number of reporting domains impact and control studies complied with, 10.5 and 10.6 (difference = -0.2, 95% confidence interval, -1.3 and 1.0, p = 0.79).

Figure 5.4.4 Number of reporting domains that impact and control studies complied with according to (A) study design and (B) journal where studies were published.

Two CPRs were originally derived to estimate the risk of developing diabetes (335, 336). They were included in the International Register because a validation study used these CPRs to estimate cardiovascular disease risk. The conclusions of this chapter did not change in sensitivity analyses that excluded impact studies of these two CPRs.

5.5. Discussion

5.5.1. Summary of findings

I presented the first study that systematically summarised design, methods, and reporting of impact studies of CPRs. The majority of impact studies of cardiovascular CPRs included in this systematic review were conducted using study designs inappropriate for evaluating the effectiveness of a CPR. Uncontrolled before-after study was the most frequently used study design. Even when impact studies used appropriate study designs, most impact studies were unusable in clinical decisions or future research because findings had substantial risk of bias or publications included insufficient details. Few impact studies of cardiovascular CPRs included in this review (5.5%) were conducted using appropriate study design, applied methods adequate enough to avoid substantial risk of bias, and complied with more than half of domains in reporting guidelines. The quality of methods and reporting was not different between impact studies of cardiovascular CPRs and studies evaluating other types of non-pharmacologic interventions.

5.5.2. Comparison with existing literature

As highlighted in the introduction section of this chapter, existing systematic reviews have consistently shown that poor design, methods, and reporting were prevalent in derivation and validation studies of CPRs (4, 50, 57-67, 71, 73, 79, 233, 284). This chapter demonstrated that poor design, methods, and reporting were also common among impact studies of cardiovascular CPRs. These

findings filled the evidence gap and completed understanding of persistent and troublesome trends throughout the CPR development stages.

I found that uncontrolled before-after study was by far the most frequently used study design among impact studies of cardiovascular CPRs. Existing systematic reviews of impact studies have typically focused on randomised experiments (91, 93). Because uncontrolled before-after studies were often excluded in these systematic reviews, it is possible that before-after study design was underreported. Interestingly, a recently published systematic review found cohort study was the most common study design for non-randomised studies of interventions funded by the UK National Institute for Health Research Biomedical Research Centres (358). Therefore, it might be hypothesized that the common use of uncontrolled before-after study is a feature of impact studies of CPRs. The same systematic review found the risk of bias assessed using the ROBINS-I tool was either serious or critical in all non-randomised trials (358), which is consistent with this chapter's findings. The main reason for the serious or critical risk of bias in both this chapter and the systematic review was leaving confounding bias from lack of randomisation unaddressed.

The compliance with reporting guidelines of impact studies in this chapter was worse than the ones described in recently conducted systematic reviews of other types of intervention (359, 360). However, these reviews only assessed the presence of reporting using criteria from reporting guidelines, rather than the completeness of reporting. On the other hand, the compliance rate of impact

studies to reporting guidelines in this chapter was comparable to a study that evaluated the completeness of reporting for domains from all relevant extensions in addition to the CONSORT statement [65].

QRISK scores were used by almost all participants who took part in a recent survey of general practitioners (GPs) in the UK (361). However, I did not find any impact study with an appropriate study design that demonstrated the effectiveness of QRISK scores. This might suggest GPs' use of QRISK scores was influenced by other factors such as recommendations from guideline or integration of CPRs in electronic health records (361), rather than evidence. Although I found that most impact studies included in my systematic review reported the results favouring the use of cardiovascular CPRs, which is consistent with previous systematic reviews of cardiovascular CPR impact studies (93, 362-365), the presence of publication bias could not be ruled out.

Systematic reviews have shown that journal endorsement of reporting guidelines (359, 366-368) and impact factor (369, 370) are associated with reporting quality, which might explain the correlation of reporting quality between impact and matched control studies selected from the same journal in Figure 5.4.4.

5.5.3. Strengths and Limitations

In Chapter 2, I found that researchers deriving new CPRs often do not cite existing evidence (94). Correspondingly, it is possible that authors of some

impact studies did not identify the CPRs they evaluated by citing the CPRs' derivation studies, and therefore, these studies were not found in the forward citation searches. Despite this limitation, I found a larger number of impact studies of cardiovascular CPRs than existing systematic reviews (91, 93, 365).

Although I conducted forward citation searches of derivation studies and identified a larger number of impact studies than existing systematic reviews of impact studies of cardiovascular CPRs that used comprehensive search strategies (93, 362-365), I only evaluated impact studies of cardiovascular CPRs published since 2013. Therefore, findings from this systematic review may not be generalizable to impact studies of CPRs for other clinical areas or impact studies published before 2013.

Although I identified control studies that were matched by study design, journal, and publication date, I was only able to limit the control studies to ones that evaluated non-pharmacologic interventions rather than non-pharmacologic interventions for cardiovascular disease. In addition, the availability of matched control studies and constraints in time and resources precluded me from including more than one control study for each impact study, which could have affected the power to detect differences in matched analyses (371).

I used criteria from reporting guidelines to assess the quality of reporting. These guidelines are intended to help authors prepare clear research reports but not necessarily meant to be used as tools for determining the reporting quality.

Although there are other studies that assessed reporting quality using reporting guidelines, it is often difficult to compare the findings due to heterogeneity in assessment methods between studies.

5.5.4. Research and clinical implications

Evaluating the effectiveness of a nonpharmacologic intervention such as cardiovascular CPRs is very challenging (9, 23, 372), which might explain why only a few cardiovascular CPRs have been evaluated in impact studies (93, 365). Firstly, unlike pharmacologic interventions where evaluations in routine clinical practice are not typically undertaken, a CPR needs to be evaluated in routine clinical practice as a component of a complex intervention (23). For example, before conducting an impact study, Karlsson et al. (373) had to develop a clinical decision support system integrated into electronic health record software that is activated when a diagnosis of atrial fibrillation is entered, presents the risk of stroke according to the CHA_2DS_2-VASc score, and prompts ordering an anticoagulant. Secondly, potential barriers for CPR use related to individual clinicians and organisations need to be addressed (23, 372). For example, after seeking clinicians' input, Reilly et al. (7, 186) modified a CPR by Golman et al. (187) because many clinicians felt a new left bundle-branch block, which was not part of the CPR, was an important predictor and also were uncomfortable with not-admitting patients with certain clinical features such as pulmonary oedema to the coronary care unit. Lastly, there are many study design-related challenges in impact studies of CPRs. Although cluster randomised trial is the ideal study design (1, 8, 23, 33), it is often difficult to

undertake because it is cost-prohibitive, takes too long, or is unsuitable (303-305). Although there are acceptable alternative designs such as controlled-before and after study (1, 8, 23, 297), these designs are infrequently used as I found in this chapter. Uncontrolled before-after study, which was the most commonly used in impact studies, is difficult to interpret due to potential bias in results (8, 290, 297, 298, 312, 313).

Researchers, funding agencies, and journals should give priority to generating and accumulating high-quality evidence of generalizability and effectiveness for existing cardiovascular CPRs, especially for those that are widely used in clinical practice already, rather than developing and publishing more cardiovascular CPRs unnecessarily. Researchers should recognize that study designs suitable for impact assessment are different from appropriate study designs for derivation and external validation stages and therefore use appropriate study designs such as cluster randomised trial when conducting impact studies. When conducting and publishing impact studies, it is important for researchers to generate credible results and clear reports so that these can be used for clinical decision making or future research which might be promoted by considering relevant methodological standards and reporting guidelines. Journals and funders could play a crucial role in improving the value of CPR research by motivating researchers to shift their research conduct and publication.

Clinicians should beware that high-quality impact studies of cardiovascular CPRs are uncommon. When a study shows that CPR can benefit patients, reduce cost, or improve efficiency, clinicians should critically evaluate the evidence and understand the uncertainties in the results.

5.5.5. Conclusion

In this chapter, I showed that the vast majority of cardiovascular CPR impact studies suffered from flaws in design, methods, or reporting. This chapter's findings expand the understanding of waste from poor design, methods, and reporting into the impact assessment stage and highlight consistent and troublesome trend throughout CPR development. More impact studies are imminently needed because even for some of widely used CPRs, it is uncertain whether using them improves care. On the other hand, existing impact studies do not disclose unbiased information about the effectiveness of CPR nor can contribute to future research. As we expect all pharmacologic interventions to be developed and evaluated in a rigorous, multiphase process before they are used for patient care (374), the same standards should apply when developing and assessing any non-pharmacologic interventions before they are adopted in clinical practice (375), including cardiovascular CPRs.

Chapter 6. Familiarity and use of cardiovascular clinical prediction rules

A manuscript based on this chapter was published in the British Journal of General Practice Open:

> Ban J-W, Perera R, Stevens R. GPs' familiarity and use of cardiovascular clinical prediction rules: a UK survey study. British Journal of General Practice Open.

6.1. Abstract

Background: Clinical prediction rule (CPRs) can help general practitioners (GPs) address challenges in preventing and managing cardiovascular disease. However, most cardiovascular CPRs derived are not used in clinical practice and represent research waste. A study published in 2014 evaluated GPs' awareness and use of CPRs in the UK. However, many new CPRs have been published since then. Therefore, I assess which cardiovascular CPRs currently recognised and used by GPs and how GPs' familiarity and use of cardiovascular CPRs have changed in this chapter.

Methods: I conducted an internet-based survey of GPs in the UK. Using the same methods as the 2014 survey, I asked GPs how familiar they were with, how frequently they used, and why they used cardiovascular CPRs. I compared the results between the 2019 and 2014 surveys.

Results: Most of the 401 GPs were familiar with QRISK scores, ABCD scores, CHADS scores, HAS-BLED score, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism. The proportions of GPs using these CPRs were 96.3%, 65.1%, 97.3%, 93.0%, 92.5%, and 82.0%, respectively. Compared with the 2014 survey, GPs' use increased by 31.2% for QRISK scores, by 13.5% for ABCD scores, by 54.6% for CHADS scores, by 33.2% for Wells scores for deep vein thrombosis, and by 43.6% for Wells scores for pulmonary embolism; and decreased by 45.9% for JBS risk calculator, by 38.7% for Framingham risk scores, and by 8.7% for New Zealand tables. GPs most commonly used cardiovascular CPRs to guide therapy and referral.

Discussion: I found GPs' familiarity and use of cardiovascular CPRs changed substantially. Integrating CPRs into guidelines and practice software might increase familiarity and use.

6.2. Introduction

Morbidity and mortality from cardiovascular disease are some of the leading sources of the burden of disease in the UK (376) and pose many challenges to general practitioners (GPs). Cardiovascular clinical prediction rule (CPRs) are intended to help clinicians including GPs address these challenges in preventing and managing the cardiovascular disease by improving clinical outcomes, increasing efficiencies, and reducing costs (93, 377).

Although many cardiovascular CPRs exist (47, 53), only a few have been broadly used by GPs (3), which might be due to waste and inefficiency in derivation, external validation, and impact assessment stages of CPR development that I explored in previous chapters. In other words, most cardiovascular CPRs are never or seldom used in practice, and therefore are research waste and inefficiency that contribute little or nothing to clinical practice.

I searched Medline and EMBASE to find studies that assessed the familiarity or use of various cardiovascular CPRs using a modified version of the search strategy used in a systematic review by Damen et al. (53). For this chapter, I define familiarity with a prediction rule as a clinician's perceived and reported knowledge or awareness of the prediction rule rather than understanding in higher-level or proficiency in using the prediction rules. I provided the search strategy used in this chapter in Appendix 6.1. I found twenty-three studies that assessed the familiarity or use of various cardiovascular CPRs: twelve from European countries (3, 378-388), six from the USA (389-394), three from Australia and New Zealand (395-397), and two from other countries (398, 399). I summarised the characteristics of these studies that assessed the familiarity or use of various cardiovascular CPRs in Table 6.2.1.

Table 6.2.1 Studies that evaluated the familiarity or use of cardiovascular clinical prediction rule.

Study, year	Focus of evaluation	Methods	Clinical context	Geographic location
1. Pearson et al. (389), 1994	Use of Goldman rule for chest pain	Survey	Emergency physicians	USA
2. Thomsen et al. (378), 2001	Use of European risk chart	Postal survey	GPs	Denmark
3. Hobbs and Erhardt (379), 2002	Use of coronary CHD risk calculator	Semi-structured telephone interview	GPs and family physicians	France, Germany, Italy, Sweden, and UK
4. De Muylder et al. (380), 2003	Knowledge and use of global CVD risk tool	Postal survey	GPs	Belgium
5. Bonnevie et al. (381), 2005	Use and attitude regarding the Copenhagen Risk Score	Internet and postal survey	GPs	Denmark
6. Eaton et al. (390), 2006	Knowledge and use of Framingham ATP III CHD risk calculator	Postal survey	Family physicians	USA
7. Eichler et al. (382), 2007	Barriers to using CVD scores	Postal survey	GPs	Switzerland
8. Sposito et al. (399), 2009	Attitude and adherence to CPRs for primary prevention of CVD	Postal survey	Cardiologists, Endocrinologists, and GPs	Brazil, Chile, Greece, Netherlands, Portugal, Venezuela, and USA
9. Imms et al. (395), 2010	Knowledge and use of New Zealand CVD risk calculator	Postal survey	GPs	Australia
10. Müller-Riemenschneider et al. (383), 2010	Use and attitude regarding CPRs in six clinical areas	Postal survey and focus group	GPs	Germany
11. Dallongeville et al. (384) 2012	Awareness and use of CVD risk assessment tools	Internet and postal survey	Primary care physicians, Cardiologists, Internists, and Endocrinologists	Austria, Belgium, France, Germany, Greece, Norway, Russia, Spain, Sweden,

Study, year	Focus of evaluation	Methods	Clinical context	Geographic location
				Switzerland, Turkey, and UK
12. Shillinglaw et al. (391), 2012	Awareness and use of global CHD risk tool	Internet survey	Family physicians, general internists, and cardiologists	USA
13. Sarazin et al. (386), 2013	Use of seven diagnostic and prognostic CPRs	Internet survey and focus group	GPs	France
14. Elustondo et al. (385), 2013	Opinion regarding CVD risk table	Postal survey	Primary care physicians	Spain
15. Plüddemann et al. (3), 2014	Awareness and use of CPRs in six clinical areas	Internet survey	GPs	UK
16. Byrne et al. (387), 2015	Awareness and use of CVD risk tool	Postal survey	GPs	Ireland
17. Pokharel et al. (392), 2016	Knowledge of 10-year ASCVD risk estimator	Survey	Internists, Family physicians, cardiologists, and endocrinologists	USA
18. Kesieme et al. (398), 2016	Use of Wells score for DVT	Semi-structured questionnaire	Surgeons	Nigeria
19. Ferguson et al. (396), 2016	Use of CHA_2DS_2-VASc and HAS-BLED score	Survey	Nurses	Australia
20. Wu et al. (393), 2017	Use of Denver screening criteria	Survey	Radiologists	USA
21. Heidbuchel et al. (388), 2018	Knowledge and skills in using CHA_2DS_2-VASc and HAS-BLED score	Semi-structured interview and quantitative internet survey	Cardiologists, GPs, Family physicians, and Neurologists	France, Germany, Italy, Poland, Spain, and UK
22. Ferguson et al. (397), 2019	Use of CHA_2DS_2-VASc and HAS-BLED score	Internet survey	Nurses in Neuroscience, stroke, rehabilitation, and cardiology	Australia
23. Amroze et al. (394), 2019	Use of CHA_2DS_2-VASc and HAS-BLED, and ATRIA score	Internet survey	Cardiologists and primary care providers	USA

GP: general practitioner, CHD: coronary heart disease, CVD: cardiovascular disease, ATP III: The third adult treatment panel, CPR: clinical prediction rule, DVT: Deep vein thrombosis

These studies showed the familiarity and use of cardiovascular CPRs varied widely depending on the CPRs, clinical contexts, and geographic locations. For example, an internet-based survey conducted in six European countries showed that CHA_2DS_2-VASc and HAS-BLED scores were used by 94.0% and 70.0% of cardiologists but only by 58.0% and 42.0% of GPs (388). Another internet-based survey from the USA found that, although 100.0% of cardiologists and 84.4% of primary care physicians reported using the CHA_2DS_2-VASc score, only 52.0% of cardiologists and 17.8% of primary care physicians reported using the HAS-BLED score. Furthermore, nurses who participated in an Australian online survey indicated using CHA_2DS_2-VASc and HAS-BLED scores much less commonly, 5.5% and 5.4% of participants, respectively (397).

A study conducted in the UK by Plüddemann et al. (3) evaluated GPs' awareness and use of CPRs for multiple clinical areas. The study included a survey of CPRs for cardiovascular disease risk, stroke, and venous thromboembolism and found most GPs were familiar with and used Framingham risk scores, QRISK scores, and Wells rule for deep vein thrombosis (3). Cardiovascular CPRs assessed in their study were published between 1998 and 2008 because those CPRs were primarily identified from guidelines developed before 2011. In this chapter, I reassess the familiarity with and use of cardiovascular CPRs among GPs in the UK because many new

CPRs have been published since then and GP's familiarity and use may have changed.

6.2.1. Objectives

The objectives of this chapter are:

(4) to understand which cardiovascular CPRs are recognised and used by GPs in the UK by conducting a survey, and

(5) to assess changes in the GPs' familiarity and use of cardiovascular CPRs by comparing the results of the current survey with the results of the 2014 survey by Plüddemann et al. (3).

6.3. Methods

6.3.1. Survey method

I conducted an internet-based survey in 2019 to examine the GPs' familiarity with and use of cardiovascular CPRs using an online questionnaire and data collection methods equivalent to those used in the 2014 survey so that results from the 2019 survey could be compared with the results of 2014 survey. There were several advantages of using the internet instead of mail to conduct a survey of health professionals: it tends to be less expensive, less time-consuming, and more flexible in questionnaire design (400). Potential disadvantages of conducting the internet-based survey included a tendency to have a lower response rate compared with postal surveys as well as the risk of

collecting unrepresentative samples depending on the degree of access to and familiarity with technologies (e.g., internet, computer, and e-mail) (400, 401). Understanding these advantages and disadvantages, I chose to use the internet-based survey methods, especially through doctors.net.uk, which is a network of doctors in the UK with more than 236,000 members (402), so that parallel comparisons of the results between the current and 2014 surveys could be made.

The Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Oxford, the National Institute for Health Research (NIHR), provided the funding but had no role in designing, conducting, or publication of the study. The Medical Sciences Interdivisional Research Ethics Committee of the University of Oxford (MS IDREC) conducted an ethical review and approved the study (Reference: R63716/RE001).

6.3.2. Questionnaire development

I developed a questionnaire for the survey based on the questionnaire used in the 2014 survey of the GPs (3). The questionnaire consisted of four sections, as presented in Appendix 6.2. The first section had three questions for determining the eligibility of participants. The target participants were GPs who were practising medicine in the UK at the time of the survey. Hospital doctors, doctors in training, GP registrars, GPs practising outside of the UK, and retired GPs were excluded.

The second and third sections contained queries about the familiarity and use of cardiovascular CPRs: eleven CPRs for cardiovascular disease risk and eight CPRs for stroke and venous thromboembolism. For the familiarity of CPRs, participants were asked to indicate whether (a) the CPR was integrated in electronic health records (EHR), (b) they had heard of it, (c) they had never heard of it, or (d) they were not sure. For the use of CPRs, participants were asked to report whether they used the CPR (a) in most or all relevant cases, (b) occasionally, (c) rarely, or (d) never. Participants were also asked to indicate for which specific reasons they used cardiovascular CPRs.

The fourth section consisted of five questions about demographic information about participants. No personal data were collected. All questions and available answers were identical to those used in the 2014 survey, except for the additional option of “the CPR was integrated in electronic health record (EHR)” for the familiarity question in the 2019 survey. I chose to add this answer to distinguish GPs who were aware of the CPR because the “CPR was integrated in the electronic health record” from GPs who had just “heard of it.” I explored whether GPs frequently used cardiovascular CPRs integrated into their EHR software because a lack of integration into EHR has been identified as one of the key barriers for GPs’ to use CPRs in practice (403, 404).

In the questionnaire, I included cardiovascular CPRs that were likely to be recognised and used by GPs in the UK at the time of the survey. Cardiovascular CPRs surveyed in the 2014 and 2019 studies are compared in Table 6.3.1.

Table 6.3.1 Clinical prediction rule (CPRs) included in the 2014 and 2019 surveys.

CPRs in the 2014 survey	CPRs in the 2019 survey	Guidelines that recommended the CPR in 2019
Cardiovascular disease risk		
JBS2 risk charts or calculator (405)	JBS3 risk calculator (406)	Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3) (406)
QRISK or QRISK2 (20, 112)	QRISK2 or QRISK3 (20, 407)	NICE Clinical guideline CG181: Cardiovascular disease: risk assessment and reduction, including lipid modification (17)
ASSIGN score (21)	ASSIGN score (21)	SIGN 149: Risk estimation and the prevention of cardiovascular disease (18)
UKPDS risk engine (125)	UKPDS risk engine (125)	-
SCORE risk charts (129)	SCORE risk charts (129)	2016 European Guidelines on cardiovascular disease prevention in clinical practice (408)
PROCAM score (332)	PROCAM score (332)	-
Framingham risk scores (216)	Framingham risk scores (216)	-
-	ACC/AHA pooled cohort equation (22)	2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease (19)
New Zealand tables (409)	New Zealand tables (409)	-
-	New Zealand primary prevention equations (410)	New Zealand Ministry of Health, Cardiovascular Disease Risk Assessment, and Management for Primary Care 2018 (411)
-	WHO/ISH risk prediction charts (412)	World Health Organization, Prevention of cardiovascular disease (413)
Stroke and venous thromboembolism [a]		
ABCD or $ABCD^2$ (134, 244)	$ABCD^2$, $ABCD^3$, or $ABCD^3$-I score (244, 414)	NICE Clinical guideline CG68: Stroke and transient ischaemic attack in over 16s: diagnosis and initial management (415)
-		
California score (416)	California score (416)	-
CHADS or $CHADS_2$ (243)	$CHADS_2$ or CHA_2DS_2-VASc score (184, 243)	NICE Clinical guideline CG180: Atrial fibrillation: management (417), SIGN 129: Antithrombotics: indications and management (418), and 2016 ESC Guidelines for the management of atrial

CPRs in the 2014 survey	CPRs in the 2019 survey	Guidelines that recommended the CPR in 2019
		fibrillation developed in collaboration with EACTS (419)
-	HAS-BLED score (185)	NICE Clinical guideline CG180: Atrial fibrillation: management (417)
Wells score for deep vein thrombosis (245)	Wells scores for deep vein thrombosis (245)	NICE Clinical guideline CG144: Venous thromboembolic diseases: diagnosis, management and thrombophilia testing (420) and SIGN 122: Prevention and management of venous thromboembolism (421)
Wells score for pulmonary embolism (246)	Wells scores for pulmonary embolism (246)	NICE Clinical guideline CG144: Venous thromboembolic diseases: diagnosis, management and thrombophilia testing (420) and SIGN 122: Prevention and management of venous thromboembolism (421)
-	Geneva or revised Geneva score (108, 422)	2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism (423) and SIGN 122: Prevention and management of venous thromboembolism (421)
-	Pulmonary Embolism Severity Index (PESI) or simplified Pulmonary Embolism Severity Index (sPESI)	2014 ESC Guidelines on the diagnosis and management of acute pulmonary embolism (423)

[a] These CPRs were included under the 'General Medical' category in the 2014 survey. ACC/AHA = American College of Cardiology/Amercian Heart Association. CPR = clinical prediction rule. ESC = European Society of Cardiology. ISH = International Society of Hypertension. JBS = Joint British Societies. NICE = National Institute for Health and Care Excellence. SIGN = Scottish Intercollegiate Guidelines Network. WHO = World Health Organization.

Firstly, because the second aim of this chapter was to assess changes in the GPs' familiarity and use of cardiovascular CPRs by comparing the current survey results with the results of the 2014 survey (3), I included in the questionnaire all CPRs for cardiovascular disease risk, stroke, and venous

thromboembolism from the 2014 survey. When a CPR had been updated, I selected the updated versions.

Secondly, I added the American College of Cardiology/American Heart Association (ACC/AHA) pooled cohort equation (22) and New Zealand Primary Prevention Equations (410). In the 2014 survey, Plüddemann et al. (3) found the Framingham risk scores (216) and New Zealand tables (409) were better recognised and more frequently used by GPs than other CPRs derived in Europe, such as the SCORE risk chart (129) and PROCAM score (332). However, in 2019, major guidelines in the USA (19) and New Zealand (411) recommended the American College of Cardiology/American Heart Association (ACC/AHA) pooled cohort equation (22) and New Zealand Primary Prevention Equations (410) instead. Therefore, these CPRs were included in the questionnaire.

Thirdly, I included any additional cardiovascular CPRs recommended by the major UK (17, 18, 417, 418, 420, 421), European (408, 419, 423-425), and World Health Organization (413) guidelines. GPs were also asked to name any other cardiovascular CPRs or CPRs in other clinical areas that they were using.

6.3.3. Data collection

GPs were recruited from doctors.net.uk using identical recruitment and survey delivery methods to the ones used in the 2014 survey (3). Doctors.net.uk

promoted the study and recruited participants by advertising the study on the GP members' homepage of doctors.net.uk and sending e-mails to invite GP members who opted to receive information about research participation. These advertisement and invitation e-mails included an online link for the study information that outlined the study's aim, why they were invited, who the researchers were, how long it would take to complete, how data would be stored, and how to raise any concern. Materials used to recruit survey participants are provided in Appendix 6.3. Participants anonymously consented online and confirmed they were 18 years or older, read and understood the information page, and agreed to participate voluntarily by ticking a box before they could take part.

For participating, 1,000 Electronic Surfing Reward points were offered as an incentive, which is equivalent to £5.00. Each GP participant was required to enter a unique user ID when completing the questionnaire so that participants could take part in the survey no more than once. In addition, potential duplicate entries were prevented by blocking questionnaire submissions from the same IP address. Doctors.net.uk did not retain these IP addresses beyond the survey completion date and did not share them with any third party. The order of CPRs presented to each survey participant in the second and third section of the questionnaire was randomised. Participants could review and change their answers before submitting the questionnaire. To be comparable with the 2014 survey, I aimed to recruit a convenience sample of 401 GPs stratified by geographic regions, which was approximately 0.9% of GP members from each

region. Doctors.net.uk closed the survey in a region when the target sample size for the region was reached.

6.3.4. Analysis

The characteristics of participants were described using medians and interquartile ranges (IQRs) for continuous variables. For categorical variables, numbers and proportions were used to present the characteristics of participants. Rather than the number and proportion of GPs who are familiar with cardiovascular CPR, I reported the number and proportion of GPs who are unfamiliar with cardiovascular CPR for the following reasons. Firstly, the 2014 survey by Plüddemann et al. (3) reported the number and proportion of GPs who had “never heard of it” and were “not sure” for each CPR, instead of the number and proportion of GPs who had “heard of it.” Therefore, to be consistent with the 2014 survey, I also presented the number and proportion of GPs unfamiliar with cardiovascular CPR: those who had “never heard of it” and were “not sure.” Secondly, in the 2019 survey, GPs who were familiar with cardiovascular CPR had an option of “CPR was integrated in EHR” in addition to “heard of it.” Therefore, answers for familiarity were less comparable than answers for unfamiliarity between the 2014 and 2019 survey. I presented the number and proportion of GPs who used cardiovascular CPR and GPs’ reported reasons for using CPRs. Although our survey responses included self-administered health questionnaires, these are beyond the scope of this thesis and not reported here.

I used the two-sample test of proportions (426) to assess a null hypothesis that the proportion of GPs unfamiliar with a cardiovascular CPR is the same between the current and 2014 survey, and to evaluate a null hypothesis that the proportions of GPs using a cardiovascular CPR is the same between the 2019 and 2014 survey. Statistical significance tests were conducted for multiple outcomes within the 2019 and 2014 survey data: 9 comparisons for GPs' familiarity with cardiovascular CPRs and 9 comparisons for GPs' use of cardiovascular CPRs. Due to the increased risk of type I error from the multiple statistical significance tests, I examined the robustness of the results by adjusting *p* values using the Bonferroni method (427). I presented 95% confidence intervals (CIs) as well as corresponding 99.5% CIs obtained using the Bonferroni method. I chose to use the Bonferroni method to correct for multiple statistical significance tests despite the criticism that this method is overly conservative (428, 429) because it was simple to calculate and likely that any differences found in the familiarity with and use of cardiovascular CPRs were real. I used Stata (Release 14. College Station, TX: StataCorp LP) for all analyses.

6.4. Results

6.4.1. Survey participants

The survey commenced on 18 June 2019 and concluded on 1 July 2019 when the target sample size of 401 was reached. Characteristics of GPs who participated and their practices were similar between the 2019 and 2014 survey,

as presented in Table 6.4.1. Most GPs who participated were practising in either urban or suburban setting: 269 (67.1%) GPs in the 2014 survey and 266 (66.3%) GPs in the 2019 survey. Similarly, most GPs' practices were located in England: 323 (80.5%) in the 2014 survey and 335 (83.5%) in the 2019 survey. The numbers of GPs in other geographic regions were also similar between the 2014 and 2019 survey. Of 401 GPs who participated in the 2019 survey, 356 (88.8%) were not involved in research: 197 with a teaching role only (49.1%) and 159 (39.7%) with neither research nor teaching role. The academic role of GP was not included in the 2014 survey.

Table 6.4.1 Characteristics of general practitioners (GPs) and their practices.

Characteristic	2014 survey, n = 401	2019 survey, n = 401
Gender [a]		
Male	243 (60.6)	245 (61.1)
Female	158 (39.4)	153 (38.2)
Other	-	3 (0.7)
Median year qualified, IQR	1995, 1986-2000	1998, 1991-2004
Type of GP		
GP Partner or principal	267 (66.6)	222 (55.4)
Salaried GP	95 (23.7)	117 (29.2)
Locum GP	32 (8.0)	54 (13.5)
Sessional GP	4 (1.0)	8 (2.0)
Retainer GP	3 (0.7)	0 (0.0)
Academic role [b]		
Both research and teaching	-	38 (9.5)
Research only	-	7 (1.7)
Teaching only	-	197 (49.1)
Neither	-	159 (39.7)
Median practice size [c], IQR	6 GPs, 4-7 GPs	6 GPs, 4-8 GPs
Practice type		
Urban area	150 (37.4)	161 (40.1)
Suburban area	119 (29.7)	105 (26.2)
Semi-rural area	87 (21.7)	89 (22.2)
Rural area	45 (11.2)	46 (11.5)
Region [d]		
England	323 (80.5)	335 (83.5)

Characteristic	2014 survey, n = 401	2019 survey, n = 401
East of England	35 (8.7)	34 (8.5)
London	48 (12.0)	56 (14.0)
East Midlands	25 (6.2)	26 (6.5)
West Midlands	33 (8.2)	36 (9.0)
North East	17 (4.2)	16 (4.0)
Yorkshire & Humber	33 (8.2)	33 (8.2)
North West	43 (10.7)	43 (10.7)
South East	51 (12.7)	54 (13.5)
South West	38 (9.5)	37 (9.2)
Scotland	48 (12.0)	38 (9.5)
Wales	18 (4.5)	17 (4.2)
Northern Ireland	12 (3.0)	11 (2.7)

Values are numbers (%) unless specified otherwise.
[a] "Other" option was not available in the 2014 survey.
[b] Academic role of GP was not included in the 2014 survey.
[c] 16 of 401 GPs did not know practice size in the current survey.
[d] England was further stratified by the Strategic Health Authority regions of the National Health Service.

6.4.2. GPs' familiarity with cardiovascular CPRs

The proportions of GPs unfamiliar with cardiovascular CPRs are presented in Table 6.4.2. Compared with the 2014 survey, the proportion of GPs unfamiliar with CPR decreased for QRISK scores by 20.9% (95% CI, 16.8-25.1; 99.5% CI, 15.0-26.9), ABCD scores by 15.2% (95% CI, 8.7-21.8; 99.5% CI, 5.8-24.6), CHADS scores by 50.1% (95% CI, 45.1-55.2; 99.5% CI, 42.9-57.4), Wells scores for deep vein thrombosis by 30.4% (95% CI, 25.4-35.5; 99 .5% CI 23.2-37.7), and Wells scores for pulmonary embolism by 43.9% (95% CI, 38.0-49.7; 99.5% CI, 35.5-52.3). The proportion of GPs who were unfamiliar with California score decrease by 3.5% with *p*-value of 0.030 and Bonferroni adjusted *p*-value of 0.329 (95% CI, 0.3-6.6; 99.5% CI, -1.0-8.0).

Table 6.4.2 General practitioners (GPs) unfamiliar with clinical prediction rules (CPRs): GPs who had never heard of or who were not sure whether they heard of respective CPRs.

Clinical prediction rule	2014 survey, N = 401			2019 survey, N = 401			Change in proportion of GPs unfamiliar to CPR				
	"Never heard of"	"Not sure"	Total	"Never heard of"	"Not sure"	Total	(%)	[95% CI]	*p* value	[99.5% CI][a]	*p* value[a]
Cardiovascular disease											
JBS3 risk calculator[b]	19 (4.7)	8 (2.0)	27 (6.7)	122 (30.4)	28 (7.0)	150 (37.4)	(30.7)	[25.3, 36.0]	< 0.001	[23.0, 38.3]	< 0.011
QRISK2 or QRISK3[c]	64 (16.0)	24 (6.0)	88 (21.9)	4 (1.0)	0 (0.0)	4 (1.0)	(-20.9)	[-25.1, -16.8]	< 0.001	[-26.9, -15.0]	< 0.011
ASSIGN score	-	-	-	222 (55.4)	32 (8.0)	254 (63.3)	-	-	-	-	-
UKPDS risk engine	-	-	-	140 (34.9)	23 (5.7)	163 (40.6)	-	-	-	-	-
SCORE risk charts	265 (66.1)	67 (16.7)	332 (82.8)	273 (68.1)	37 (9.2)	310 (77.3)	(-5.5)	[-11.0, 0.0]	0.052	[-13.4, 2.4]	0.571
PROCAM score	328 (81.8)	50 (12.5)	378 (94.3)	337 (84.0)	36 (9.0)	373 (93.0)	(-1.2)	[-4.6, 2.1]	0.469	[-6.1, 3.6]	1.000
Framingham risk scores	4 (1.0)	2 (0.5)	6 (1.5)	18 (4.5)	4 (1.0)	22 (5.5)	(4.0)	[1.5, 6.5]	0.002	[0.4, 7.6]	0.023
ACC/AHA pooled cohort equation[d]	-	-	-	347 (86.5)	30 (7.5)	377 (94.0)		-	-	-	-
New Zealand tables	249 (62.1)	64 (16.0)	313 (78.1)	328 (81.8)	26 (6.5)	354 (88.3)	(10.2)	[5.1, 15.4]	< 0.001	[2.9, 17.6]	< 0.011
New Zealand primary prevention equations[d]	-	-	-	352 (87.8)	28 (7.0)	380 (94.8)	-	-	-	-	-
WHO/ISH risk prediction charts[d]	-	-	-	262 (65.3)	33 (8.2)	295 (73.6)	-	-	-	-	-
Stroke and venous thromboembolism											
$ABCD^2$, $ABCD^3$ or $ABCD^3$-I score[e]	128 (31.9)	45 (11.2)	173 (43.1)	98 (24.4)	14 (3.5)	112 (27.9)	(- 15.2)	[-21.8, -8.7]	< 0.001	[-24.6, -5.8]	< 0.011
California score	336 (83.8)	50 (12.5)	386 (96.3)	341 (85.0)	31 (7.7)	372 (92.8)	(-3.5)	[-6.6, -0.3]	0.030	[-8.0, 1.0]	0.329
$CHADS_2$ or CHA_2DS_2-VASc score[f]	160 (39.9)	48 (12.0)	208 (51.9)	5 (1.2)	2 (0.5)	7 (1.7)	(-50.1)	[-55.2, -45.1]	< 0.001	[-57.4, -42.9]	< 0.011
HAS-BLED score[d]	-	-	-	12 (3.0)	4 (1.0)	16 (4.0)	-	-	-	-	-
Wells scores for deep vein thrombosis	109 (27.2)	30 (7.5)	139 (34.7)	12 (3.0)	5 (1.2)	17 (4.2)	(-30.4)	[-35.5, -25.4]	< 0.001	[-37.7, -23.2]	< 0.011
Wells scores for pulmonary embolism	169 (42.1)	59 (14.7)	228 (56.9)	31 (7.7)	21 (5.2)	52 (13.0)	(-43.9)	[-49.7, -38.0]	< 0.001	[-52.3, -35.5]	< 0.011
Geneva or revised Geneva score[d]	-	-	-	315 (78.6)	32 (8.0)	347 (86.5)	-	-	-	-	-
PESI or simplified PESI[d]	-	-	-	318 (79.3)	30 (7.5)	348 (86.8)	-	-	-	-	-

[a] Adjusted for conducting 11 significance tests using the Bonferroni method.
[b] "JBS2 risk calculator or JBS2 risk charts" were included in the 2014 survey.
[c] "QRISK or QRISK2" were included in the 2014 survey.
[d] These CPRs were not included in the 2014 survey.
[e] "ABCD or $ABCD^2$" were included in the 2014 survey.
[f] "CHADS or $CHADS_2$" were included in the 2014 survey.

6.4.3. GPs' use of cardiovascular CPRs

The proportions of GPs using CPRs in the 2014 and 2019 survey are presented in Figure 6.4.1 and Figure 6.4.2. For cardiovascular disease risk CPRs, the proportion of GPs using QRISK scores increased by 31.2% (95% CI, 26.2-36.2; 99.5% CI, 24.0-38.4) whereas the proportion of GPs using CPR decreased for JBS risk calculator by 45.9% (95% CI, 39.9-51.8; 99.5% CI, 37.4-54.4), and Framingham risk scores by 38.7% (95% CI, 32.7-44.6; 99.5% CI, 30.1-47.2), and New Zealand tables by 8.7% (95% CI, 5.2-12.2; 99.5% CI, 3.7-13.7), as shown in Figure 6.4.1.

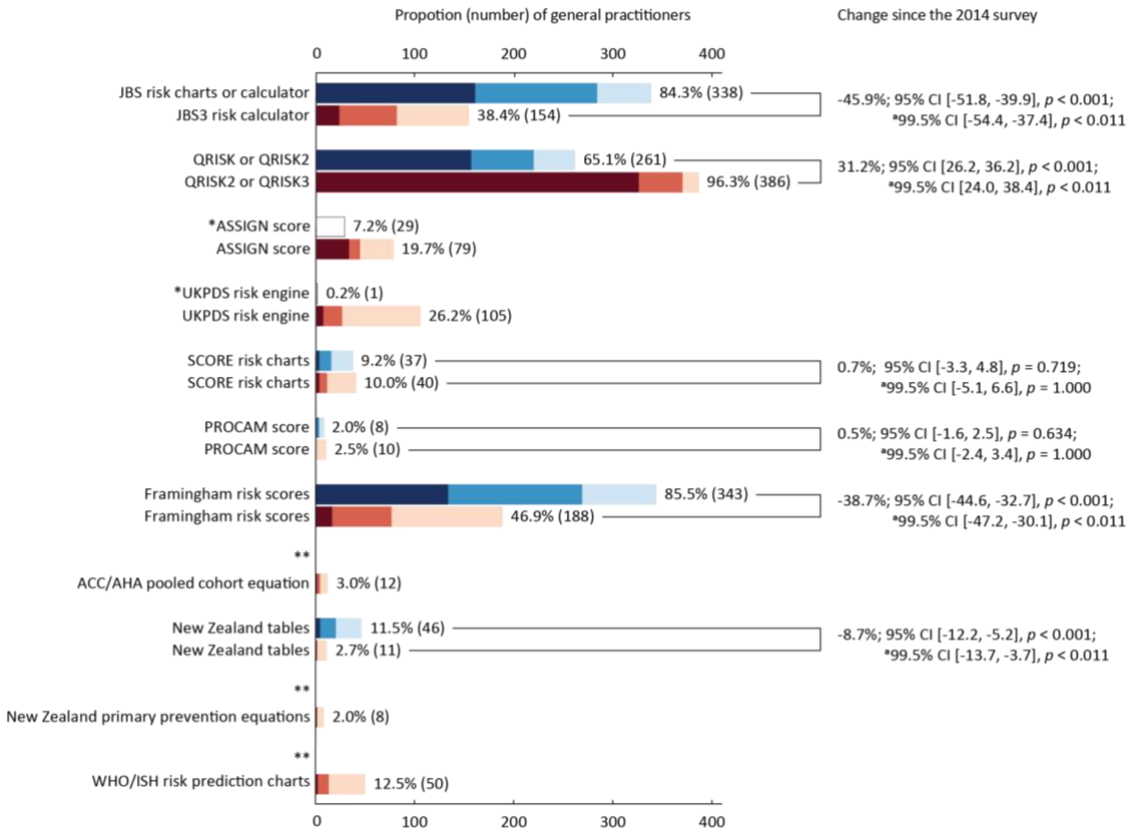


Figure 6.4.1 The proportion of general practitioners using clinical prediction rules for cardiovascular disease in the 2014 and 2019 survey.

[a]Confidence interval (CI) and p-value adjusted using the Bonferrroni method

Among CPRs for stroke and venous thromboembolism, the proportion of GPs using CPR increased for ABCD scores by 13.5% (95% CI, 6.7-20.2; 99.5% CI, 3.8-23.1), CHADS scores by 54.6% (95% CI, 49.5-59.7; 99.5% CI, 47.3-61.9), Wells scores for deep vein thrombosis by 33.2% (95% CI, 27.7-38.6; 99.5% CI, 25.4-41.0), and Well scores for pulmonary embolism by 43.6% (95% CI, 37.6-49.7; 99.5% CI, 35.0-52.3), as presented in Figure 6.4.2.

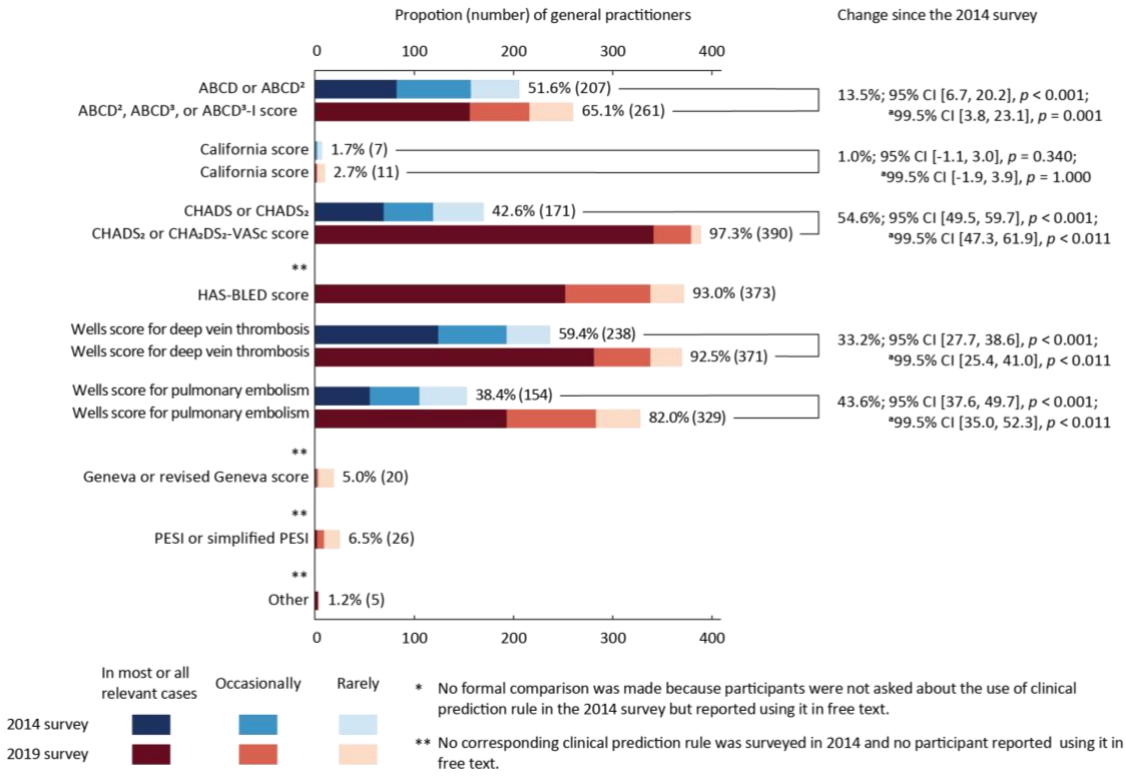


Figure 6.4.2 The proportion of general practitioners using clinical prediction rules for stroke and venous thromboembolism in the 2014 and 2019 survey.

[a]Confidence interval (CI) and p-value adjusted using the Bonferrroni method

The reported frequencies of using cardiovascular CPRs by GPs who indicated the CPR was integrated into their EHRs and all GPs who participated in the survey are compared in Table 6.4.3. Cardiovascular CPRs that GPs most frequently reported using “most or all relevant cases” were QRISK scores (325 GPs, 81.0%), ABCD score (156 GPs 38.9%), CHADS scores (341 GPs, 85.0%), HAS-BLED score (252 GPs, 62.8%), Wells scores for deep vein thrombosis (281 GPs, 70.1%), and Wells scores for pulmonary embolism (193 GPs, 48.1%). These cardiovascular CPRs were also the ones that GPs most often reported being integrated into their EHR software: 280 (69.8%) GPs for

QRISK scores, 107 (26.7%) GPs for ABCD score, 273 (68.1%) GPs for CHADS scores, 207 (51.6%) GPs for HAS-BLED score, 184 (45.9%) GPs for Wells scores for deep vein thrombosis, and 153 (38.2%) GPs for Wells scores for pulmonary embolism.

Table 6.4.3 Frequency of using cardiovascular clinical prediction rules (CPRs) by all general practitioners (GPs) who participated and GPs who reported that CPR was integrated in electronic health record, n (%).

Clinical prediction rule	GPs who reported that "CPR was integrated in electronic health record"					All GPs who participated in the survey				
	Number of GPs	Use CPR			Do not use CPR	Number of GPs	Use CPR			Do not use CPR
		Most or all cases	Occasionally	Rarely			Most or all cases	Occasionally	Rarely	
Cardiovascular disease										
JBS3 risk calculator	43	12 (27.9)	16 (37.2)	9 (20.9)	6 (14.0)	401	23 (5.7)	58 (14.5)	73 (18.2)	247 (61.6)
QRISK2 or QRISK3	280	257 (91.8)	21 (7.5)	1 (0.4)	1 (0.4)	401	325 (81.0)	44 (11.0)	17 (4.2)	15 (3.7)
ASSIGN score	24	21 (87.5)	2 (8.3)	1 (4.2)	0 (0.0)	401	33 (8.2)	11 (2.7)	35 (8.7)	322 (80.3)
UKPDS risk engine	14	3 (21.4)	3 (21.4)	4 (28.6)	4 (28.6)	401	7 (1.7)	19 (4.7)	79 (19.7)	296 (73.8)
SCORE risk charts	3	1 (33.3)	1 (33.3)	1 (33.3)	0 (0.0)	401	3 (0.7)	8 (2.0)	29 (7.2)	361 (90.0)
PROCAM score	2	0 (0.0)	0 (0.0)	0 (0.0)	2 (100.0)	401	0 (0.0)	0 (0.0)	10 (2.5)	391 (97.5)
Framingham risk scores	56	7 (12.5)	19 (33.9)	19 (33.9)	11 (19.6)	401	16 (4.0)	60 (15.0)	112 (27.9)	213 (53.1)
ACC/AHA pooled cohort equation	1	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	401	1 (0.2)	3 (0.7)	8 (2.0)	389 (97.0)
New Zealand tables	1	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	401	1 (0.2)	0 (0.0)	10 (2.5)	390 (97.3)
New Zealand primary prevention equation	1	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	401	1 (0.2)	0 (0.0)	7 (1.7)	393 (98.0)
WHO/ISH risk prediction charts	2	0 (0.0)	1 (50.0)	1 (50.0)	0 (0.0)	401	2 (0.5)	11 (2.7)	37 (9.2)	351 (87.5)
Stroke and venous thromboembolism										
$ABCD^2$, $ABCD^3$ or $ABCD^3$-I score	107	81 (75.7)	21 (19.6)	3 (2.8)	2 (1.9)	401	156 (38.9)	61 (15.2)	44 (11.0)	140 (34.9)
California score	1	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	401	1 (0.2)	2 (0.5)	8 (2.0)	390 (97.3)
$CHADS_2$ or CHA_2DS_2-VASc score	273	255 (93.4)	13 (4.8)	3 (1.1)	2 (0.7)	401	341 (85.0)	39 (9.7)	10 (2.5)	11 (2.7)
HAS-BLED score	207	168 (81.2)	29 (14.0)	10 (4.8)	0 (0.0)	401	252 (62.8)	87 (21.7)	34 (8.5)	28 (7.0)
Wells scores for deep vein thrombosis	184	166 (90.2)	14 (7.6)	1 (0.5)	3 (1.6)	401	281 (70.1)	58 (14.5)	32 (8.0)	30 (7.5)
Wells scores for pulmonary embolism	153	112 (73.2)	29 (19.0)	9 (5.9)	3 (2.0)	401	193 (48.1)	91 (22.7)	45 (11.2)	72 (18.0)
Geneva or revised Geneva score	1	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	401	1 (0.2)	3 (0.7)	16 (4.0)	381 (95.0)
PESI or simplified PESI	3	1 (33.3)	1 (33.3)	1 (33.3)	0 (0.0)	401	2 (0.5)	8 (2.0)	16 (4.0)	375 (93.5)

The reported reasons for using CPRs are presented in Table 6.4.4. The most common reasons for using CPRs for cardiovascular disease risk were to guide therapy (91.0% in 2019 survey, 84.0% in 2014 survey), comply with clinical guidelines (63.4% in 2019 survey, 67.0% in 2014 survey), and inform or educate patients (55.1% in 2019 survey, 55.0% in 2014 survey). The most common reasons for using CPRs for stroke and venous thromboembolism were to guide referral (71.4% in 2019 survey, 54.0% in 2014 survey), guide therapy (69.8% in 2019 survey, 47.0% in 2014 survey), and aid diagnosis (68.6% in 2019 survey, 55.0% in 2014 survey).

Table 6.4.4 General practitioners (GPs) reporting respective reasons for using clinical prediction rules.

	2014 survey, N = 401		2019 survey, N = 401	
	Cardiovascular disease risk	Stroke and venous thromboembolism	Cardiovascular disease risk	Stroke and venous thromboembolism
Use clinical prediction rule	395 (98.5)	328 (81.8)	399 (99.5)	398 (99.3)
• To aid diagnosis	143 (35.7)	219 (54.6)	145 (36.2)	273 (68.1)
• To assess disease severity	191 (47.6)	142 (35.4)	163 (40.6)	141 (35.2)
• To guide therapy	336 (83.8)	188 (46.9)	363 (90.5)	278 (69.3)
• To guide referral	89 (22.2)	218 (54.4)	95 (23.7)	284 (70.8)
• To comply with clinical guidelines or Qualities and Outcomes Framework (QOF)	267 (66.6)	106 (26.4)	253 (63.1)	220 (54.9)
• To inform or educate patients	220 (54.9)	75 (18.7)	220 (54.9)	137 (34.2)
• Automatically generated by practice software	106 (26.4)	19 (4.7)	67 (16.7)	52 (13.0)
• Other	1 (0.2)	3 (0.7)	2 (0.5)	3 (0.7)
Do not use clinical prediction rule	6 (1.5)	73 (18.2)	2 (0.5)	3 (0.7)

Values are numbers (%) of GPs.

Five GPs named four other CPRs that they were using for stroke and venous thromboembolism but were not asked about in the survey. They were the Cincinnati Prehospital Stroke Scale (CPSS, also known as FAST) (430) (n = 2), NIH stroke scale (431) (n = 1), PERC rule (141) (n = 1), and Q stroke score (432) (n = 1).

CPRs in other clinical areas GPs most commonly reported using were the Fracture Risk Assessment Tool (FRAX) (433) (n = 24), Centor score (434) (n = 16), FeverPain score (12) (n = 16), CRB-65 or CURB-65 (13) (n = 7), and QCANCER Risk Assessment Tools (435) (n = 6).

6.5. Discussion

In this chapter, I evaluated which cardiovascular CPRs were recognised and used by GPs in the UK, and assessed how GPs' familiarity with and use of cardiovascular CPRs have evolved by comparing the results of the 2019 survey with the results of the 2014 survey.

6.5.1. Summary of findings

I adjusted all *p*-values and confidence intervals using the Bonferroni methods to account for the increased risk for incorrectly rejecting a true null hypothesis owing to the multiple statistical significance tests. Although these adjustments can increase the risk of incorrectly accepting a false null hypothesis, I found that almost all the comparisons remained statistically significant after adjusting for

multiple testing. Only an apparent increase in GPs' familiarity with the California score in the 2019 survey did not remain statistically significant. However, this inconsistency did not affect my overall conclusions.

I found that the vast majority of GPs who participated in the survey were familiar with the QRISK score and used it to assess cardiovascular disease risk. The CPRs for stroke and venous thromboembolism that most GPs recognised and used were ABCD scores, CHADS scores, HAS-BLED score, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism. These cardiovascular CPRs were also the CPRs recommended by national guidelines in the UK and most frequently integrated into GPs' EHR software. QRISK scores have become dominant CPRs for cardiovascular disease risk assessment in the UK while the popularity of Framingham risk scores waned compared with the 2014 survey. For stroke and venous thromboembolism, substantially more GPs are now using ABCD scores, CHADS scores, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism. It may be hypothesised that integrating CPRs into national guidelines and GPs' EHR software increases the familiarity with and use of CPRs. GPs used CPRs for cardiovascular disease risk most commonly to guide therapy, comply with clinical guidelines, and inform or educate patients; and CPRs for stroke and venous thromboembolism mainly to guide referral, guide therapy, and aid diagnosis.

6.5.2. Comparison with existing literature

Older studies published before 2010 often concluded that CPRs for cardiovascular disease risk were infrequently used in primary care (378-382, 389, 390, 399). For example, a US national survey from 2006 found only 17% of family physicians used the coronary heart disease risk calculator recommended by national guidelines (390). Similarly, a postal survey from Switzerland showed that just 46.8% of GPs either frequently or occasionally used a CPR to assess global cardiovascular disease risk (380).

On the other hand, more recently published studies tend to show higher rates of awareness and use of CPRs for cardiovascular disease risk in primary care (3, 383, 385, 387, 395). For example, a study published in 2015 found 92.5% of Irish GPs were aware of cardiovascular disease risk assessment tools, and 72.8% of them used them, most commonly the SCORE risk calculator (387). Their results match the results of this chapter that nearly all GPs who took part in the 2019 survey were familiar with and used CPRs for cardiovascular disease risk, especially QRISK scores. This chapter's findings confirm the trends observed in existing evidence that primary care physicians have become more familiar with CPRs for cardiovascular disease risk and are more frequently using them in practice. One exception to this trend found in evidence review was a study from the USA found that only 61.2% of physicians were aware of the internet version of the ACC/AHA pooled cohort equation, and 56.5% of physicians felt comfortable using it (392). Although it might be postulated that

the US physicians are less familiar with cardiovascular CPRs recommended by national guidelines, the fact that 78.2% of participants were physicians in training (392) could potentially explain participants' relative unfamiliarity to the ACC/AHA pooled cohort equation.

The vast majority of GPs in the 2019 survey indicated they were using CPRs for stroke and venous thromboembolism such as $CHADS_2$ or CHA_2DS_2-VASc score, HAS-BLED score, and Wells score for deep vein thrombosis. The proportions of GPs using these CPRs were generally higher than the ones reported in other recently published studies (388, 394, 396-398). For example, only 58% of GP participants in an internet survey conducted in six European countries (388) and 84.0% of primary care physicians in an internet survey from the USA (394) reported using CHA_2DS_2-VASc score, which is lower than the proportion (97.3%) of GPs that indicated using CHA_2DS_2-VASc score in this chapter. The reported use of the HAS-BLED score was just 42.0% in the European survey (388) and 18% in the US survey (394), which was also far less than the GPs' reported use of HAS-BLED score of 93.0% in this chapter. These contrasts suggest the use of CHA_2DS_2-VASc and HAS-BLED score in primary care might differ between countries. Furthermore, the GPs' use of Wells score for deep vein thrombosis in the current survey was much higher than the reported use by surgeons in Nigeria (398), suggesting there might be differences in the degree of cardiovascular CPR use between specialities.

Common reasons for using cardiovascular CPRs in the existing literature were to educate patients (385, 387, 395), motivate lifestyle changes (384, 387, 395), guide drug therapy (384, 387, 391, 395), and establish treatment goal (387, 395). In addition to these, GPs in our study frequently used cardiovascular CPRs to comply with guidelines or Qualities and Outcomes Framework (QOF), which might be unique for GPs in the UK. The NICE guideline published in 2008 advised using the Framingham risk score to assess cardiovascular risk (436). After external validation studies in the UK showed QRISK scores consistently performed better than the Framingham risk score (219, 220, 437), QRISK2 was recommended by the 2014 NICE guideline (17). This shift in guideline recommendation might explain notable changes in the use of CPRs for assessing the risk of cardiovascular disease. Similarly, most GPs in our survey used CHADS scores and HAS-BLED score that was first recommended by the 2012 SIGN (418) and 2014 NICE guideline (417).

The authors of the 2014 survey (3) noted that GPs were relatively unfamiliar with Centor score (434) despite having strong evidence of generalizability in primary care, whereas many GPs used NICE traffic light system (438) that lacked evidence to support the use of it. These findings are consistent with the results of chapters 5 and 6 that some cardiovascular CPRs commonly recognised and used by GPs such as QRISK scores and ACBD scores lack high-quality evidence of impact. In other words, factors other than research evidence might have an important influence on the use of cardiovascular CPRs in primary care practice.

6.5.3. Strengths and limitations

According to the systematic literature search results presented in the introduction section, this is the first study that evaluated changes in the familiarity with and use of CPRs over time in a country. I was able to assess temporal trends in the familiarity with and use of CPRs among GPs in the UK by applying recruitment strategy, sampling mechanism, questionnaire design, and analysis methods equivalent to the ones used in the 2014 survey.

I recruited GPs from doctors.net.uk for compatibility with the 2014 survey. This strategy allowed me to recruit a geographically representative sample of GPs from all UK regions efficiently and gave direct comparability to the 2014 survey. However, there were several weaknesses in using doctors.net.uk. Firstly, it was impossible to calculate the response rate because doctors.net.uk partly relied on the advertisement on the GP members' homepage to recruit participants. Secondly, I could not investigate whether GPs who participated differed from those who did not participate. Therefore, I cannot exclude the possibility that the survey sample over-represented GPs who were interested in cardiovascular CPRs and those who frequently used computers and the internet, especially the doctors.net.uk homepage. Lastly, it is unclear whether the survey findings are applicable to all GPs in the UK because the participants were sampled from doctors.net.uk, which has a large number of GP members but is not necessarily representative of all GPs. Furthermore, the survey findings may have limited generalisability outside of the UK.

The vast majority of GPs who participated in the survey indicated that they used CPRs for cardiovascular disease risk to guide therapy which seems rationale considering all CPRs for cardiovascular disease risk included in the survey were prognostic rules. However, some GPs indicated that they used CPRs for cardiovascular disease risk to aid diagnosis. This shows one of the weaknesses of survey research; “data are useful only when survey instruments can prompt accurate and consistent information from participants” (439). Although inconsistent responses could have been reduced by asking GPs why they were using each CPR for cardiovascular disease risk in retrospect, this approach would have made the survey substantially longer and mode difficult to conduct due to the prohibitive costs.

6.5.4. Implications for research and practice

The study by Plüddemann et al. (3) concluded that GPs’ lack of familiarity with CPRs was one of the main reasons that GPs were not using CPRs in practice. The findings of this chapter suggest that integrating CPRs into national guidelines and EHR software might be important factors for increasing GPs’ familiarity and use of cardiovascular CPRs. Conducting an international survey in countries where guidelines recommend different CPRs and where EHR software has a varying degree of CPR integrations might help assess these associations. It is possible that the associations between integrating CPRs into national guidelines or EHR and GP’s familiarity with and use of cardiovascular CPR, are not causal. Ultimately, these hypotheses can be tested by conducting

interrupted time series studies and comparing the familiarity with and use of cardiovascular CPRs in multiple time points before and after publishing a new guideline that recommends a cardiovascular CPR, or before and after integrating a cardiovascular CPR into EHR software. Interrupted time series design has been used to assess the impact of health policy (440) or guideline (441) on clinical practice.

Similar to the 2014 survey, the GPs' use of cardiovascular CPRs is focused on a few selected CPRs such as QRISK scores. In other words, even though CPRs are intended to be used to help decisions of clinicians and patients, most existing cardiovascular CPRs are never or rarely used. These cardiovascular CPRs are research waste and inefficiency that contribute little or nothing to clinical practice. Therefore, I believe the focus of cardiovascular CPR research should be shifted from deriving more CPRs that are unlikely to contribute to clinical practice to externally validating, and assessing the impact of existing CPRs to prevent further accumulation of research waste and inefficiency.

6.5.5. Conclusion

In this chapter, I showed that there had been notable changes in cardiovascular CPRs that GPs are familiar with and use since 2014. The most widely recognised and used cardiovascular CPRs by GPs were QRISK scores, ABCD scores, CHADS scores, HAS-BLED score, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism. Recommending cardiovascular CPRs in national guidelines and integrating them into GPs' EHR

software seem to promote familiarity with and use of cardiovascular CPRs. In the next chapter, I will explore how primary care providers' experiences with, and perceptions of cardiovascular CPR research influence their use of cardiovascular CPRs in practice.

Chapter 7. Influence of research evidence on the uptake of cardiovascular CPRs

7.1. Abstract

Background: Even for high-quality research evidence about cardiovascular CPRs to be valuable, the evidence should contribute to improving the healthcare decisions of patients, clinicians, or policymakers. However, existing knowledge about the influence of research evidence on the use of CPRs is limited. Therefore, I evaluate in this chapter how perceptions of and experiences with cardiovascular CPR research influence primary care providers' (PCPs) decisions about using cardiovascular CPRs.

Methods: I conducted an exploratory qualitative interview study with thematic analysis. I recruited PCPs from the WWAMI (Washington, Wyoming, Alaska, Montana and Idaho) region Practice and Research Network (WPRN). The WPRN sent e-mails to WPRN Survey Research Panel members and invited them to express the interest in participating by completing an online questionnaire. From PCPs who completed the online questionnaire, I purposefully selected PCPs for interview using a sampling frame to maximise diversities. I conducted semi-structured online interviews with PCPs using a pilot-tested topic guide. Interview recordings were transcribed verbatim. I analysed data using an inductive thematic analysis approach and generated

semantic themes. Each PCP who completed the interview received a $50.00 Amazon eGift Card for their time.

Results: Fifteen PCPs were interviewed. I identified two main themes relating to the influence of PCPs' perceptions of and experiences with cardiovascular CPR research on their decisions about using cardiovascular CPRs: "Seek and Judge" and "Be acquainted and Assume". When PCPs are familiar with, trust, and feel confident in using research evidence, they might actively "Seek and Judge" the evidence which then could influence their decisions about using cardiovascular CPRs. On the other hand, when PCPs are unfamiliar with, distrust, or find it challenging to use research evidence, they might be passively acquainted with evidence and assume the evidence is adequate. Therefore, these PCPs might not rely on research evidence when making decisions about using cardiovascular CPRs.

Conclusions: PCPs' perceptions of and experiences could influence how they use research evidence in decisions about using cardiovascular CPRs. An intervention targeted at unfamiliarity, distrust, and challenges regarding the use of research evidence might promote the use of research evidence in their decisions. Further, because PCPs often rely on factors unrelated to evidence, guideline developers and policymakers should recommend cardiovascular CPRs supported by robust evidence.

7.2. Introduction

Health research is valuable when it creates evidence that improves the healthcare decisions of patients, clinicians, or policymakers; or when other researchers can use the evidence to synthesise such knowledge (34). High-quality evidence about a clearly justified CPR can be accumulated through well designed and conducted derivation, external validation, and impact studies by demonstrating the CPR's ability to accurately predict an outcome, verifying satisfactory performance in different populations and settings, and confirming a potential to improve patient care. But for evidence accumulated through these stages of CPR development to be valuable, the evidence in its original form or knowledge synthesised using the evidence should contribute to improving the healthcare decisions of patients, clinicians, or policymakers. In other words, even high-quality research evidence about cardiovascular CPRs is still research waste if it is not used by patients, clinicians, or policymakers to help their decisions.

To understand existing knowledge about this topic, I initially searched electronic databases for studies that evaluated the influence of research evidence on the uptake of CPRs. Because no relevant study was identified, I sought studies that examined factors influencing the use of CPRs by searching Medline (Ovid) and EMBASE (Ovid) using the strategy provided in Appendix 7.1. I also reviewed the electronic database search results for studies that assessed the familiarity or use of cardiovascular CPRs in Chapter 6 and my personal collection of CPR literature to find relevant studies. I found 33 studies that evaluated barriers or

facilitators for using CPRs in various clinical domains: 17 for cardiovascular (97, 378-382, 384, 385, 387-391, 394, 399, 403, 442), six for musculoskeletal (443-448), four for CPR in general (49, 383, 386, 404), three for neurological (289, 449, 450), one for cancer (451), one for digestive (452), and one for respiratory (453). Only eight of these studies reported some results about the influence of evidence on the uptake of CPRs. The findings from these studies are summarised in Table 7.2.1.

Table 7.2.1 Studies that reported the influence of research evidence on using clinical prediction rule (CPRs).

Author, year	Methods and context	CPR	Findings
Graham et al. (443), 1998	Postal survey of Emergency Physicians	Ottawa ankle rule, Ottawa knee rule, Canadian CT (computerized tomography) head rule, and Canadian cervical spine rule	Physicians were willing to use well-validated and highly sensitive CPRs.
Eichler et al. (382), 2007	Postal survey of general practitioners (GPs)	Cardiovascular CPRs	Some GPs distrusted the CPRs' validity because their accuracy was insufficient.
Müller-Riemenschneide et al. (383), 2010	Postal survey and focus group of GPs	CPRs in six clinical domains	GPs "challenged the evidence base of CPRs for their particular patient population" (383).
Dowling and Wishart (446), 2011	Postal survey of Paediatric Emergency Physicians	Ottawa ankle rule	10.5% did not use Ottawa ankle rule because of insufficient evidence to support the use in children.
Ballard et al. (450), 2013	Internet survey of Emergency Physicians	PECARN Paediatric head injury rule	Most physicians were more likely to use a CPR if it has a high sensitivity or specificity.

Author, year	Methods and context	CPR	Findings
Brown et al. (404), 2016	Internet survey and focus group of Family Physicians	CPRs in general	“Most GPs were unaware of the evidence behind CPRs and felt addressing this would promote their use” (404). GPs discussed evidence about limitation of CPRs with patients.
Kennedy and Gallego (49), 2018	Systematic review of studies collecting healthcare providers’ opinion about using CPRs	CPRs in general	Healthcare providers were willing to implement CPRs when “confident in the evidence supporting the tool“ such as validation study (49).
Muthee et al. (97), 2020	Systematic review of studies evaluating factors for the uptake of CPRs in primary care	Cardiovascular CPRs	Evidence of accuracy and effectiveness facilitated the use of cardiovascular CPRs. Uncertainty in the evidence of impact also influenced the implementation.

These eight studies reported that research evidence about CPRs (382, 404), in particular for accuracy (97, 382, 443, 450), generalizability (49, 383, 443, 446), and effectiveness (97), influenced the use of CPRs. However, findings from these studies primarily related to whether having or not-having sufficient evidence affected the use of CPRs. Therefore, more substantive knowledge about how research evidence impacts cardiovascular CPRs' uptake is still absent. In this chapter, I explore how research evidence influences the use of cardiovascular CPRs because understanding the role of evidence in this process might help develop strategies to promote the use of evidence in decisions for using CPRs in practice.

7.2.1. Objectives

The objectives of this chapter are:

(1) To explore primary care providers' (PCPs') perceptions of and experiences with cardiovascular CPR research, and

(2) To understand how PCPs' perceptions and experiences influence their perceived uptake of cardiovascular CPR in practice.

7.3. Methods

In this chapter, I chose to use a qualitative research approach. Although quantitative research methodologies were used mostly in the previous chapters, I determined that a qualitative research approach, specifically an exploratory qualitative interview study with thematic analysis, was more appropriate to address this chapter's research question and objectives. Furthermore, I decided to use an inductive thematic analysis because there was little existing knowledge about how PCPs' perceptions of and experiences with cardiovascular CPR research influence the uptake of the CPRs.

7.3.1. Participant selection

Initially, I planned to recruit general practitioners (GPs) in the UK with support from the Clinical Research Network Thames Valley and South Midlands (CRN TVSM) and Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC). However, I could not secure the needed funding,

which was estimated to be over £39,000.00 for recruiting 20 GPs. Therefore, I modified my protocol to work with the WWAMI region Practice and Research Network (WPRN) and recruit primary care providers (PCPs) in the USA instead. The WRPN is a network of diverse primary care practices located in urban, suburban, and rural areas across the five-state WWAMI (Washington, Wyoming, Alaska, Montana and Idaho) region in academic, private, community, and public settings (454). One of the functions of the WPRN is to link researchers with primary care practices for research participation and provide support in conducting primary care research (454).

Based on the numbers of participants included in existing exploratory qualitative interview studies conducted in primary care (455-457), I aimed to recruit 15 PCPs initially with a plan to recruit more PCPs as needed. I recruited PCPs who were (1) 18 years or older, (2) willing and able to give informed consent, and (3) currently practising office-based primary care medicine such as Family Practice, Internal Medicine, Geriatrics, or General Practice in the USA. The WPRN sent e-mails with recruitment materials to the PCP members of the WPRN Survey Research Panel inviting them to take part in the study. The PCP members were asked to complete a brief online questionnaire to express the interest in participating in the interview. Also, they were asked to provide demographic, practice, and contact information. PCPs were required to consent for taking part in the online questionnaire by ticking a box before they could commence. A $5.00 Amazon eGift Card was provided to PCPs who completed the online questionnaire. The recruitment letter and online questionnaire are provided in Appendix 7.2 and Appendix 7.3.

From 29 PCP members who completed the online questionnaire and expressed the interest in taking part, I selected 15 PCPs for interview using a sampling frame aimed at maximizing diversities in their use of cardiovascular CPRs, research experience, and consideration of research evidence in decisions about using cardiovascular CPRs. I determined that thematic saturation was achieved after analysing data from 15 PCP interviews and did not recruit additional PCPs for interview.

7.3.2. Data collection

I conducted qualitative semi-structured interviews with PCPs. The Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) of the University of Oxford conducted an ethical review and approved the study (Ethics Approval Reference: R69695/RE001). The initial protocol included conducting interviews in-person or online according to participating PCPs' preference. However, due to Coronavirus disease (Covid-19) pandemic, all interviews were conducted online using Microsoft Teams. Interviews took place between 3 August 2020 and 10 September 2020. All participants consented to participate in the interview: by sending a completed and signed written consent form via post or e-mail in advance or giving verbal consent immediately before starting the interview.

I used a topic guide to ask PCPs about their perceptions of and experiences with cardiovascular CPR research and about the influence of these perceptions

and experiences on their use of cardiovascular CPRs. I updated the initially developed topic guide after conducting three pilot interviews with primary care physician colleagues. The topic guide was revised again after interviewing five PCP participants to include questions about PCPs' familiarity with cardiovascular CPR research because of substantial heterogeneity in participating PCPs' familiarity with cardiovascular CPR research. The final version of the topic guide is provided in Appendix 7.4.

After a brief introduction, I explained the purpose, scope, and process of the interview to the participant and confirmed they understood how to raise concerns. By loosely following the topic guide's structure but maintaining the natural flow of conversation, I asked PCPs about their use of cardiovascular CPRs, reasons for choosing the CPRs, views on research evidence, experiences with research evidence, and research evidence's role in their decisions about using cardiovascular CPRs. I audio-recorded the interviews using QuickTime Player on a password secured desktop computer with no automatic cloud back-up. Each PCP who completed the interview received a $50.00 Amazon eGift Card.

7.3.3. Data management and analysis

I transcribed interview recordings verbatim using Otter (Otter.ai), an online software that generates transcriptions from live or recorded voices. In all data, except for the consent forms and interview recordings, I removed any information that could directly discern an individual participant and identified

each participant only by an alphanumeric code assigned to. I stored all data, including interview recordings, interview transcripts, consent forms, answers from the screening questionnaire, and participant identifiers, in a dedicated folder within the University of Oxford's Nexus 365 OneDrive for Business.

I analysed data using inductive thematic analysis approach described by Braun and Clarke (105) and generated semantic themes. After importing each interview transcript to NVivo 12, a qualitative data analysis software, I corrected errors in interview transcripts made by Otter (Otter.ai) as accurately as possible while repeatedly listening to the corresponding interview recordings. Then, I generated initial codes by finding small data segments that might be meaningful, thinking about and focusing on, but not necessarily restricted to, my specific research questions. After coding all interview transcripts, I categorised codes into themes while reviewing data within each theme and revising codes as needed. At the same time, I assessed the relationship between the themes and reorganized them using an evolving concept map as a guide. I continued this process until I was no longer making significant revisions to the themes.

7.4. Results

7.4.1. Participants

Of 29 PCPs who completed the online questionnaire, 15 PCPs were interviewed. The characteristics of these 29 PCPs are presented in Table 7.4.1. The median year when 15 interviewed PCPs started practising as a PCP was

2007 (IQR, 1991-2017). They were in various career stages (between second-year resident in training and semiretired), and practised in Washington or Idaho.

Table 7.4.1 Characteristics of primary care providers (PCPs) who completed the online questionnaire.

	Interviewed, n = 15	Not interviewed, n = 14	All, n = 29
The year started as a PCP, median (IQR)	2007 (1991-2017)	2004 (1995-2019)	2007 (1995-2018)
Practice size, median (IQR)	30 (7-48)	27.5 (10-40)	30 (10-40)
Speciality			
Family medicine	10 (66.7)	10 (71.4)	20 (69.0)
Internal medicine	5 (33.3)	4 (28.6)	9 (31.0)
PCP type			
Physician	12 (80.0)	12 (85.7)	24 (82.8)
Resident	2 (13.3)	2 (14.3)	4 (13.8)
Physician assistant	1 (6.7)	0 (0.0)	1 (3.4)
Teaching or research experience			
Both research and teaching	5 (33.3)	8 (57.1)	13 (44.8)
Research only	0 (0.0)	1 (7.1)	1 (3.4)
Teaching only	8 (53.3)	5 (35.7)	13 (44.8)
Neither	2 (13.3)	0 (0.0)	2 (6.9)
Practice setting			
Urban	6 (40.0)	10 (71.4)	16 (55.2)
Suburban	6 (40.0)	2 (14.3)	8 (27.6)
Semi-rural	2 (13.3)	2 (14.3)	4 (13.8)
Rural	1 (6.7)	0 (0.0)	1 (3.4)
Practice location			
Western Washington	8 (55.3)	8 (57.1)	16 (55.2)
Eastern Washington	4 (26.7)	3 (21.4)	7 (24.1)
Idaho	3 (20.0)	1 (7.1)	4 (13.8)
Montana	0 (0.0)	1 (7.1)	1 (3.4)
Wyoming	0 (0.0)	1 (7.1)	1 (3.4)
Integration in practice software			
At least one clinical prediction rule (CPR)	9 (60.0)	5 (35.7)	14 (48.3)
No CPR	6 (40.0)	9 (64.3)	15 (51.7)
Use of CPR in practice			
In most or all cases	6 (40.0)	5 (35.7)	11 (37.9)
Occasionally	6 (40.0)	7 (50.0)	13 (44.8)

	Interviewed, n = 15	Not interviewed, n = 14	All, n = 29
Rarely	3 (20.0)	1 (7.1)	4 (13.8)
Never	0 (0.0)	1 (7.1)	1 (3.4)
When choosing CPR			
Considered evidence	14 (93.3)	11 (78.6)	25 (86.2)
Did not consider evidence	1 (6.7)	3 (21.4)	4 (13.8)

Values are numbers (%) of PCPs unless indicated otherwise.

The PCPs were purposefully selected for interview according to the sampling frame as presented in Table 7.4.2. Demographic data of these PCPs who were interviewed are provided in Appendix 7.5.

Table 7.4.2 Selection of primary care providers (PCPs) for interview according to the sampling frame.

	Has research experience		Has no research experience	
	Considered evidence	Did not consider evidence	Considered evidence	Did not consider evidence
Use CPR in most or all relevant cases	2 (4)	0 (1)	3 (4)	1 (2)
Use CPR occasionally	3 (7)	0 (1)	3 (5)	0 (0)
Use CPR rarely or never	0 (1)	0 (0)	3 (4)	0 (0)

Values are numbers of PCPs selected for interviews (numbers of PCPs who completed the online questionnaire).

7.4.2. Main themes

From data collected from interviews, I identified two main themes regarding the influence of PCPs' perceptions of and experiences with cardiovascular CPR

research on their decisions about using cardiovascular CPRs: “Seek and Judge” and “Be acquainted and Assume.” I presented concept maps for these themes in Figure 7.4.1 and 7.4.2 illustrating how PCPs’ perceptions and experiences in relation to cardiovascular CPR research might influence how PCPs interact with research evidence and make decisions about using cardiovascular CPRs.

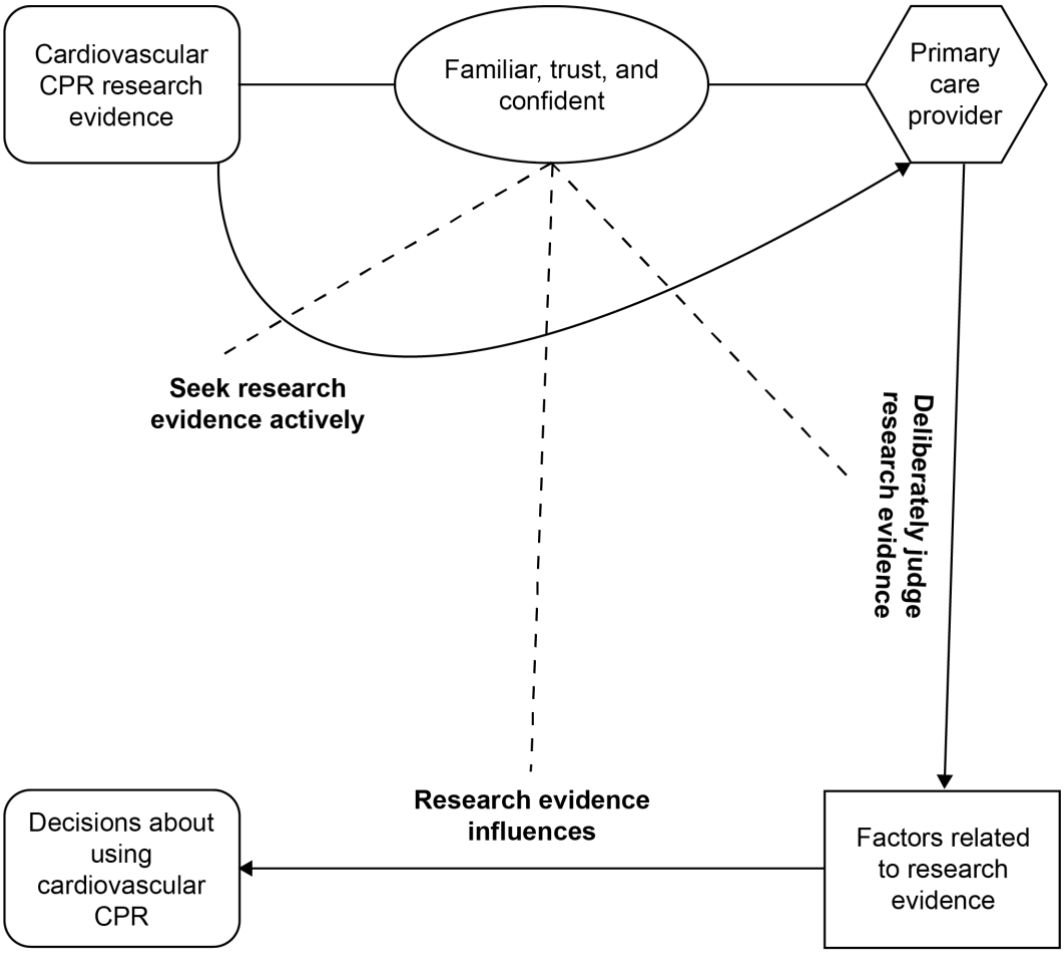


Figure 7.4.1 A concept map of the theme "Seek and Judge."

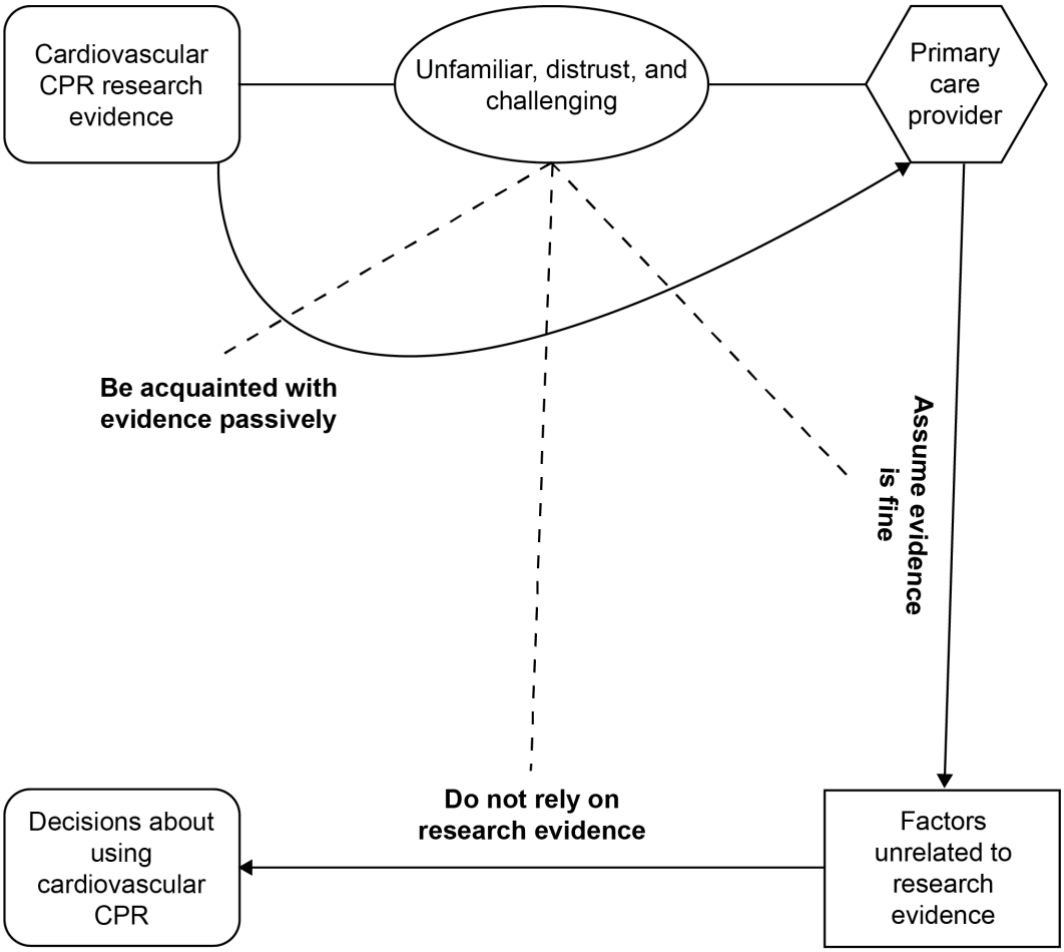


Figure 7.4.2 A concept map of the theme "Be acquainted and Assume."

7.4.3. “Seek and Judge”

This main theme includes four subthemes: “familiar, trust, and confident”, “seek and judge”, “influence of evidence”, and “evidence related factors”. The first subtheme, “familiar, trust, and confident” represents PCPs’ perceptions of and experiences with cardiovascular CPR research. The second and third subthemes, “seek and judge” and “influence of evidence” explain how PCPs’ perceptions and experiences might affect the use of research evidence. The

last subtheme, “evidence related factors” describes some of the aspects in research evidence that might be important to PCPs’ decisions about using cardiovascular CPRs.

A. Familiar, trust, and confident

Some of the PCPs’ perceptions of and experiences with cardiovascular CPR research might help PCPs include research evidence in their decisions about using cardiovascular CPRs. When a PCP is familiar with cardiovascular CPR research in general or is aware of research evidence about a cardiovascular CPR, the PCP’s decision about using the cardiovascular CPR might involve the research evidence. A few interviewed PCPs were familiar with cardiovascular CPR research to a certain extent and knew some research evidence behind cardiovascular CPRs they were using. A General Internist who learned about various cardiovascular CPRs such as the ACC/AHA pooled cohort equation, PERC rule, CHA_2DS_2-VASc score, and Wells scores during medical school and residency training had some knowledge about what cardiovascular CPR research was:

> “Cardiovascular clinical prediction rule research. … It could mean kind of like what you're doing, seeing how much people understand about these targets, you know, what they mean, and how you use them. So I think it's a very broad category, but it kind of encompasses this whole how do you validate these things? How do you use them are people using them accurately?” (IS78)

Also, he recalled the fact that some of the cardiovascular CPRs had cost-effectiveness analyses:

> "Well, I think some of these studies have cost-benefit analyses. Some of these predictors do have cost-benefit analyses that show, you know, that there are cost benefits to using these clinical predictors in the right setting." (IS78)

Similarly, research evidence might influence a PCP's decision on whether or how the PCP uses a cardiovascular CPR when the PCP trusts and values the research evidence about the cardiovascular CPR. Some PCPs conveyed that they trusted, had faith in, and had a favourable opinion on cardiovascular CPR research and evidence the research generated. A Family Physician who used the ACC/AHA pooled cohort equation in most or all relevant cases trusted evidence around the ACC/AHA pooled cohort equation because she did not feel there was any potential conflict of interest:

> "I'm pretty sceptical pharma's involvement in things in general. … So, I think, like for example, it ASCVD the like, moderate versus high-intensity statin type of thing, it seems like, just instinctively like, there isn't like a ton of possible, um, corruption with like, that versus other things like asthma guidelines where the inhalers are super expensive and the, you know, some much newer and more expensive and not generic than others. But with like, ASCVD it doesn't seem to play as much, which is reassuring." (BG35)

Further, when a PCP feels confident in using research evidence, the PCP's decision on whether or how the PCP uses a cardiovascular CPR might be influenced by the research evidence behind the cardiovascular CPR. Some PCPs, especially younger and newly trained physicians interviewed, expressed their confidence in searching for, accessing, reading, and appraising primary research evidence about cardiovascular CPRs. A third-year Internal Medicine resident who rarely used the ACC/AHA pooled cohort equation felt confident

about searching for research evidence using PubMed and indicated PubMed filters were helpful:

> "When I'm trying to broaden my own knowledge, I tend to use PubMed or part of the University of Washington libraries search, search tools. … Uh, yeah, I think PubMed is pretty easy. There's also a section of PubMed that only looks at clinical studies as well rather than more primary bench top type research. So that tends to be helpful." (FA84)

Similarly, a second-year Internal Medicine resident who used various guideline-recommended cardiovascular CPRs felt comfortable reading and interpreting primary studies because he had a fair amount of training in biostatistics and critical appraisal:

> "My comfort level with reading primary literature is fairly good. I feel like I had pretty decent training in like Biostatistics and ability to to read primary literature and interpret what they mean when they use sensitivity, specificity, specificity, number needed to treat things like that. … I think that our biostatistics particularly at med school, I felt like we did a strong amount of that." (SY54)

These PCPs who expressed confidence in using research evidence about cardiovascular CPRs learned the knowledge and skills to search for, read, and appraise research evidence during medical school and residency training through a formal curriculum on evidence-based medicine (EBM) including journal club, or informal teachings such as discussions with mentors and colleagues.

B. Seek and judge

When a PCP has favourable perceptions of or experiences with cardiovascular CPR research, the PCP might actively seek and judge research evidence about cardiovascular CPRs. Some interviewed PCPs, who were familiar with, trusted, and felt confident in using research evidence about cardiovascular CPRs, actively sought and searched for primary research evidence about cardiovascular CPRs using electronic databases, libraries, and online resources. A General Internist who was experienced with and felt confident about searching research evidence looked up a study in PubMed when a question came up during rounding about using the PERC rule in patients with a high risk of pulmonary embolism:

> "I'm a big fan of MeSH headings in PubMed. That's how I find most things if there's a particular question that we're asking. … When we were on rounds one day at the hospital, and the residents were sort of using the PE rule criteria in a very high-risk patient. … And we so the question was, you know, 'In high-risk patients, does the PE rule criteria apply?' And so I usually use a MeSH heading on PubMed to just look for whatever the title is, and then for diagnosis, and and that was one of the one that had just come out in the last four to five years. And so we pulled it in that was the that was the one we picked." (GO38)

Some PCPs who were familiar with, trusted, and felt confident in using research evidence about cardiovascular CPRs, questioned others' assessments of evidence and tried to judge cardiovascular CPR studies' trustworthiness independently or with other PCPs at journal club. Some found evaluating research evidence about cardiovascular CPRs easy and valuable, but others felt lost and disappointed when they did not understand information. A General Internist who felt confident about his searching and appraising skills that he picked up during his residency training, shared how he assessed cardiovascular CPR studies:

> "Oh, mostly, I kind of go through their methodology and see what what that's like. You know, I look at the kind of generalizability to my patient population as well. Um, I mean, I basically do a literature appraisal to see if it's appropriate to the clinical question I have in terms of validation for these particular risks, risk stratification tool. … Some papers are real easier to read and understand than others." (IS78)

But, when judging the trustworthiness of evidence about cardiovascular CPRs, PCPs relied on certain study characteristics and results such as sample size, population, setting, outcome, funding source, or *p*-value, rather than study design and internal validity. A newly trained Family Physician who had limited time reading research articles focused on the sample size, type of population, and study results to determine the paper's trustworthiness.

> "Oh, I look at… the number of participants, I look at… what kind of a study it was. And what the patient population was made up of, if it's similar to my patient population … and then what the, I guess, the confidence interval is." (PR88)

Therefore, although some PCPs seemed to be actively seeking and judging the research evidence about cardiovascular CPRs, their searches and appraisals might not necessarily have sufficient thoroughness and rigour according to the classic definition of EBM (458).

C. Influence of evidence

When a PCP has positive perceptions of and experiences with cardiovascular CPR research, the PCP's decision about using cardiovascular CPRs might be influenced by research evidence. Some PCPs, who were familiar with, trusted, and felt confident in using research evidence, argued that PCPs should use

research evidence when making decisions about using cardiovascular CPRs and indicated that reviewing research evidence influenced whether or how they used cardiovascular CPRs. A General Internist who was experienced in reading cardiovascular CPR studies believed that PCPs should review research evidence about cardiovascular CPRs before using them because some were created in different settings and might not have been validated in primary care:

> "Because a lot of them are … designed in an emergency setting, … I do think sometimes they get misused in an outpatient setting. I wish that we had more that were very clearly outpatient validated. … It is easy to just plug the numbers into a rule, um, and sort of make snap judgments that don't always consider all the research that went into building that rule." (GO38)

Further, she indicated that reading articles about the TIMI and Heart scores led her to change the way she used these cardiovascular CPRs in her evaluation and management of patients with chest pain:

> "I think every study I read is useful in some way. … I don't think I've ever read a study that told that made me abandon a decision rule. Um, but it really just helps me hone, 'when do I use it?' and on 'who do I use it?' And so I think that's what I've seen evolve in my practice is it used to be anytime, you know, anytime someone had chest pain, I would use a TIMI, and that was like the thing that we did every chest pain, you know, that I saw. And, so now I really tried to refine and say, 'is this the right person?' or 'is this someone where I think the HEART score makes more sense' in this patient, you know, based on some of those initial studies." (GO38)

Similarly, after reviewing an article about a new cardiovascular CPR, a Family Physician who was familiar with cardiovascular CPR research felt the evidence about the new CPR was "compelling" and recommended residents to use the CPR:

> "I found it compelling. Yes, I found it that it will be worth to change it … to the new numbers that they proposed compared to the older one. Yes, I did. ... I think they push us to use. I think it, knowing my bias, knowing my credibility of the people who are writing things that I listening to the research. … I have banked on those articles, decided to use those. I guess in this case, and I tried to teach it to the residents also. Like two or three years ago, when I was training the first years, I asked all of them to download the ACVD app to their phone. And then EPIC added it to our modules probably a year ago. So now we are using it in the computer." (FX91)

Therefore, having familiarity, trust, and confidence regarding the use of research evidence might promote the involvement of research evidence in PCPs' decisions about using cardiovascular CPRs.

D. Evidence related factors

When a PCP's decision about using a cardiovascular CPR involves research evidence, certain evidence related factors might play an important role in the PCP's decision. PCPs who took part in this study considered three evidence related factors in their decisions about using cardiovascular CPRs: having evidence behind, evidence of applicability, and features of evidence.

Some PCPs thought it was important for the cardiovascular CPRs they used had robust and valid evidence of accuracy, generalizability, and effectiveness. For examples, an experienced faculty Family Physician at a small residency programme taught his residents to use cardiovascular CPRs that had been evaluated in an external validation study:

> "Certainly it's really nice to see a validation study in a different setting, before rolling out anywhere else. … Yeah, for them, it's almost branding.

> That's what's most famous? What's everybody using? What's the community standard? And that's, a little bit they're relying on us, the faculty to be picking the ones that are best to use. … I think I've taught most of them that most prediction rules have never been validated. So you better be really sure you're using one that's at least validated. So baby steps." (GR91)

However, the most important factor related to research evidence that influenced the PCPs' decisions on using cardiovascular CPRs, was the CPRs' evidence of applicability to PCPs' own patients. These PCPs evaluated populations and settings included in CPR studies and assessed whether they were sufficiently similar to the PCPs' patients and settings. The PCPs felt confident when applying cardiovascular CPRs evaluated in populations and environments similar to the PCPs' but were reluctant to use cardiovascular CPRs studied only in different populations or settings. A third-year Internal Medicine resident who believed PCPs should consider research evidence when choosing cardiovascular CPRs explained how he decided to use the ACA/AHA pooled cohort equation:

> "For the ASCVD risk score, it seemed the most broad to me. I know there are different rule sets and whether you're looking at females or you're looking at different ethnic groups. Because I mostly deal with populations that seemed to fall most in line with what the ASCVD population was. … It doesn't mean that I don't have black patients or female patients. … I tried to look at my patient population on my outpatient panel. And I tried to select one that was, knowing that if I was going to just select one, which one was going to envelop my patients most appropriately. So, looking at it from a population standpoint and then making sure that there is research to support whichever tool I was using." (FA84)

Similarly, a Family Physician argued that, "as a good doctor", it was important to know that the research evidence applied to his patient populations fairly early in his decision about choosing a cardiovascular CPR:

> “I think, a good doctor, you know what I mean, would know that, right? Or, if something didn't apply to your population, then the results aren't as you can't extrapolate them as nicely as you could to a different population. So, we definitely talk to, talk to our residents about those kinds of things, when you're looking at studies and that kind of thing, you know, ‘does it apply to your patient population?’ … Yeah, I think that would be important. … That's probably a fairly early decision in using that tool. ‘Does, do the my patients kind of match the patients they used to come up with this? or not?’ Yeah, I think that's important.” (IN63)

For some PCPs interviewed, certain features about research evidence such as having a large sample size, higher power, extended follow-up, multiple locations, and statistically significant results were important when choosing to use cardiovascular CPR. A General Internist, who modified how he applied the CPRs in his practice after critically appraising cardiovascular CPR studies, discussed some of the features of research evidence about cardiovascular CPRs that was important to him:

> “I would say, you know, like, for me, the biggest thing would be the sample size. If, like in some of these research tools of validating some of these predictors, if the sample size is super small, and you know, it's just done in a very specialized group of people, probably wouldn't do it, use those tools. You know, studies are underpowered, I probably wouldn't use those tools as well. The *p*-value was not, you know, appropriate, I probably wouldn't use those tools as well.” (IS78)

Therefore, even when PCPs believed their decisions about using cardiovascular CPRs were influenced by research evidence, they were usually guided by the aspect of research evidence that might not be the most crucial in ensuring the evidence was high-quality (6, 199, 459).

7.4.4. “Be acquainted and Assume”

This main theme also consists of four subthemes: “unfamiliar, distrust, and challenging”, “be acquainted and assume”, “do not rely on evidence”, and “evidence unrated factors”. The first subtheme, “unfamiliar, distrust, and challenging” characterises PCPs’ perceptions of and experiences with cardiovascular CPR research. The second and third subthemes, “be acquainted and assume” and “do not rely on evidence” describe how PCPs’ perceptions and experiences might influence the use of research evidence. The last subtheme, “factors unrelated to evidence” explains how PCPs make decisions about using cardiovascular CPRs when they do not rely on research evidence.

A. Unfamiliar, distrust, and challenging

Some of the PCPs’ perceptions of or experiences with cardiovascular CPR research might impede directly involving research evidence in the PCPs’ decisions about using cardiovascular CPR. For example, when a PCP is unfamiliar with cardiovascular CPR research or unaware of research evidence about a cardiovascular CPR, it would be implausible for research evidence to directly influence whether or how the PCP use the cardiovascular CPR. In fact, many interviewed PCPs did not know any research evidence about cardiovascular CPRs that they were using or could not recall any specifics about the research evidence they might have read. A second-year Internal Medicine resident who frequently used various guideline-recommended cardiovascular CPRs including the ACC/AHA pooled cohort equation, PERC

rule, CHA_2DS_2-VASc score, and HAS-BLED score admitted that he had limited knowledge about the evidence behind these CPRs:

> "Yeah, so this is where I may have knowledge about ... the scores, but I have limited knowledge about necessarily the research that goes into them because I can't say that I've done like a lot of digging behind them. I somewhat … have trust in my training, in my training that I'm learning these scores in some part due to my attendings who have helped me. And I think that they have looked more into the research." (SY54)

Similarly, when a PCP distrusts research evidence about a cardiovascular CPR despite being aware of the evidence, the research evidence could have limited influence on whether or how the PCP uses the cardiovascular CPR. Some interviewed PCPs had doubts and scepticism for research evidence in general and shared the same views on cardiovascular CPR research. For example, a Family Physician who had postgraduate degrees (MPH and PhD) in Public Health indicated she read a paper about CHA_2DS_2-VASc score multiple times because she had seen widely accepted evidence refuted over time and, therefore, sometimes "did not believe it":

> "I think I had to read like three times. … like 35, 25 years ago when I was in medical, in my graduate school, they were all talking about hormone replacement therapy. And then I had a teacher who said, 'This is not going to be true, this is not going to be true.' And everybody was saying, 'Oh, this is the panacea, everybody should be taking HRTs.' So, I have these in my mind that I don't believe everything people say all the time. I have seen so many things changing on everything in medicine that I have a bias. So, I usually read things often and read again and read again, because sometimes I don't believe it." (FX91)

In addition, when a PCP had a difficult experience with cardiovascular CPR research or views using research evidence about a cardiovascular CPR as challenging, the research evidence might have limited influence on the PCP's

decision about using the cardiovascular CPR. Some of the common challenges in using research evidence about cardiovascular CPRs among interviewed PCPs were lack of time, ability, and resource. PCPs interviewed often had difficulties using research evidence because they had a busy clinical schedule and little time to read. A Family Physician who was a professor at a University and had more time than typical PCPs practising in the community discussed his challenge in using research evidence due to lack of time:

> “Yeah, I just think people are busy. You know, even for me as an academic, I don't have the time. I haven't made the time to go back and look at the original data and research under them. So I think that's really hard for practising physicians to do that.” (KJ34)

A recently-graduated Family Physician who was focused on building her patient panel at a suburban primary care practice explained why it was challenging for her look up research evidence when she chose to use the ACC/AHA pooled cohort equation:

> “So, but … why wouldn’t, why don’t I do just that? Having a busy practice. … So, I tried to read the AFP journal articles. But that’s about all I have time for. So if they're not there, or they're not, you know, I’m not having any sort of benchmark and notifications about whether I'm meeting certain goals, then, then I don’t know if I’m, if I’m even aware of certain tools that are available.” (PR88)

Further, some PCPs pointed out challenges related to lack of resources. For example, an Internist who worked at a University-affiliated residency programme speculated that PCPs in the community might have difficulty using evidence about cardiovascular CPRs because they lacked resources to access research evidence:

> "… with my affiliation with like, UW (University of Washington) and WSU (Washington State University) Elson Floyd College of Medicine, I get free access to most all, you know, these things like UpToDate, DynaMed, paper access. But I suspect not everybody else does. So, I think that might be a challenge sometimes." (IS78)

Lastly, several PCPs, especially those older PCPs with more clinical, research, and teaching experiences, felt it was challenging for them to use research evidence about cardiovascular CPRs because they had limited ability, training, knowledge, or skills. These PCPs indicated that they were unfamiliar with methods, found research evidence complex, and were unconfident in assessing research evidence independently. A Family Physician who trained about 250 Family Medicine residents over 30 years throughout his career doubted Family Medicine residents' ability to critically appraise cardiovascular CPR studies:

> "By the end of their residency, I might have a few Family Medicine residents who I think are able to judge the quality of a prediction rule. … I don't think most of them leave, the certainly 90% of my residents, don't leave their training with those skills. They probably could sweat their way through it, but it is not going to be pleasant and the quality of their outcomes." (GR91)

An experienced Family Physician who completed residency and fellowship training in Sports Medicine 10 years ago also did not feel that he had skills or knowledge to evaluate research evidence about cardiovascular CPRs independently:

> "I can't say I spent a lot of time picking apart, how these tools got created and what data it came from. I don't know that I have the mental capacity… or the knowledge to trust what I think more than what other people think that are more savvy with that kind of information." (IN63)

Even a Family Physician who was also a very experienced researcher felt it was difficult for her to read and critically appraise cardiovascular CPR studies:

> "Oh, I mean, we all were, quote, unquote, 'taught'. But, um actually even as a researcher and someone who has been trained quite extensively in research methods, I think it's extremely difficult for even someone like myself to critically read an article well and interpret the results." (PX33)

Therefore, she argued it was "unrealistic" to think that typical PCPs could critically appraise research evidence about cardiovascular CPRs:

> "I think that… it's unrealistic to think that students, residents, practising providers are actually going to be able to critically appraise the literature and come to their own conclusions… about very much." (PX33)

Although it was counterintuitive, some of these PCPs who were in a position to educate students, residents, other PCPs tended to be sceptical about their ability to search, appraise, and use evidence about cardiovascular CPRs.

B. Be acquainted and assume

When a PCP has unfavourable perceptions of or experiences with cardiovascular CPR research, the PCP may not actively seek and evaluate research evidence about cardiovascular CPRs. Even so, the PCP could be passively acquainted with research evidence about cardiovascular CPRs. In this case, the PCP might not make a deliberate judgment about whether the PCPs can trust the research evidence. For example, some PCPs who were unfamiliar with, distrusted, or found it challenging to use research evidence, came in contact with research evidence about cardiovascular CPRs through didactic

sessions, rounding, or journal club during medical school or residency training. Other PCPs were introduced to research evidence about cardiovascular CPRs via their clinical colleagues, administrative role, and online resources. For examples, a newly graduated Family Physician who found it challenging to search for and read cardiovascular CPR studies got to learn research evidence behind some cardiovascular CPRs from medical school, didactic session, journal club discussion, and clinical pharmacist's talk:

> "In med school, there was a big emphasis on the JNC eight. … And then, two years ago … when things changed, and it became like putting less people on Statins, we had a big didactic session about it and a journal club. And, I remember us having a discussion about the research behind it. … I think what I learned about Framingham, that's one of the things that I learned in med school, was that because it was like, I think it was in like Massachusetts or something. So, it was like a lot of old white men in like the 80s. That kind of stuff. … Yeah, we have a clinical pharmacist at our clinic who's here two days a week. And he also teaches the residents. And so, twice a month, he does a little lunchtime 30-minute talk where he goes through primary research, a couple of new primary research papers on medication-related things." (BG35)

A Physician Assistant who did not have time to look up research evidence about cardiovascular CPRs because of busy work and family life came to know about cardiovascular CPR studies by reading e-mails that she received from online resources she subscribed to:

> "I get like, these e-mails from like Healio is one of them. And … it's just like a research study program that you have subscribed to, and I do some continuing medical education on. … And then, sometimes just reading the article, it was having a study on the presentation, I mentioned on. … Like, if I log on them, yeah, Healio Minute is one. I get tons of Healio Minute e-mails, and then I read all those because they're just a variety of medical stuff. … And then, I also get e-mails from PA news, Doximity. I do a lot of Doximity, that's right. … I do ECHO with ECHO Idaho … And then, they'll present, you know, research and those sometimes. … Presentations, yeah, some of them are like webinars." (TF63)

Some PCPs who were unfamiliar with, distrusted, or found it challenging to use research evidence, did not make an explicit judgement about the validity of research evidence. For example, a second-year resident in General Internal Medicine residency programme “skimmed through abstract”, rather than poring over the full article and judging the credibility, because he found it difficult to understand the details:

> “Of course, there it's, it's not so hard to get into the details that maybe go, not over my head. But like, do I really care about all of these little details? Yeah. So, I do more abstracts skim them than like full like methodology and dive into the articles. … So that that is something that, you know when judging credibility, I could use personal improvement on on when I look through articles to make sure that the methodology seems sound.” (SY54)

Instead of critically evaluating research evidence behind cardiovascular CPRs they were using, these PCPs assumed others who had more time, resources, and skills had already assessed the evidence. A Family Physician who did not feel confident in evaluating research evidence assumed someone smarter than him had already evaluated the evidence behind cardiovascular CPRs he used:

> “I assumed… that there are people out there smarter than me that have thought through this. And if it's a generally accepted thing that, you know, the standard of care is blah, blah, blah, then I should probably be on board with that. … But the short answer is I assumed that somebody smarter than me already thought through this. And so I'm gonna go with that. … I don't feel confident in assessing some of these things, myself, and I assume that there's smarter people out there that have figured this out, and I don't have to go down this path myself. … But there are people that are fantastic at this stuff and they are and they're really good at it kind of dumbing it down for people like me.” (IN63)

Some PCPs who were unable to critically appraise cardiovascular CPR studies felt they could not help but trust, have faith in, and go with others' judgment on the evidence. A family physician who was unfamiliar with methods used in cardiovascular CPR studies and had difficulty interpreting the results felt he was at the mercy of authors, journal editors, professional societies, and guideline developers and could not help but trust their interpretations:

> "It's difficult to figure out how to interpret those. … And so I think largely we're, we are left to trust that the authors and editors of those. And then also relying on our specialty societies and other guideline writing organizations to help us interpret that information. … I specifically think about my colleagues who are seeing patients all day every day that we're largely at the mercy of the medical journal editors … or what we are then taught by other societies, or … guideline writing organizations. But yeah, I think like … for me to dive into that and assess all the methods that they utilized, and really pretend that I could make a better conclusion than the authors or editors of the paper would probably be a little bit presumptuous." (ZY58)

Some PCPs who did not critically evaluate the evidence behind information assumed that information was credible if it came from certain sources. A busy senior Family Physician who felt he was not experienced with critically appraising cardiovascular CPR studies judged the credibility of articles depending on which journal they were published:

> "As a primary care provider, you're busy. And, you don't have a lot of background in determining the quality of the paper. I think it's helpful to go to a journal … like New England Journal or Annals of Internal Medicine. … It's not 100%, but that's pretty helpful. And then … if people are asking me … 'hey … go to Cochrane or go to the US Preventive Services Task Force.' I think those those papers are pretty tried and true." (JA36)

Similarly, a family physician and researcher who did not critically appraise research evidence determined the credibility of information according to who the authors were.

> “I don't. I mean, technically, I could sort of look at who are these authors are they, of the UpToDate, you know. Are they people who I think are trustworthy? Um, you know, I think as a practitioner, I think that there are a lot of leaps of faith. … I depend on particular authors.” (PX33)

Therefore, when PCPs did not make a deliberate judgement about research evidence, they assumed that someone else had evaluated the evidence, they could go with others’ interpretation of the evidence, or they could trust the evidence if it came from certain sources.

C. Do not rely on evidence

When a PCP has negative perceptions of or experiences with cardiovascular CPR research, the PCP may not rely on research evidence when making decisions about using cardiovascular CPRs. PCPs who were unfamiliar with, distrusted, or found it challenging to use research evidence about cardiovascular CPRs, did not seek, did not rely on, or were not influenced by research evidence when deciding whether or how they used cardiovascular CPRs. For example, a second-year General Internal Medicine resident who frequently used the ACC/AHA pooled cohort equation had not looked up research evidence behind this CPR because he found it challenging to identify important studies when searching for research evidence:

> "But, honestly, I haven’t… done a whole lot of my own digging there. … Something like ASCVD risk, I … probably haven’t looked up the the

> primary research behind it. … Challenges with searching is maybe trying to know what the, trying to get a good … large scale study, or maybe the most recent study. Sometimes, it is challenging in knowing that this is the most up to date or the best study that's out there because … when you get the search, you just get what pops up best with your keywords that matches and not necessarily like the signature study … or the classic study or that's your thing. That can be challenging, I think." (SY54)

Similarly, a Family Physician who found it difficult to read and interpret CPR studies did not rely on evidence of accuracy when he chose to use the CHA_2DS_2-VASc score and HAS-BLED score:

> "Whoo. … I know, I've seen some papers in the last few years. … It's difficult to figure out how to interpret those. The, you know, the, the statistical analysis involved is not something that probably most primary care clinicians are going to be immediately familiar with. … I have chosen tools to use often based on guidelines, or sort of what's often typical standard practice. So, I wouldn't say that the… research suggesting that the tools are one tool may or may not be as accurate has affected my choice of the tool because I don't feel that there's necessarily a superior alternative." (ZY58)

Even when some interviewed PCPs were aware of research evidence showing limitations of the cardiovascular CPRs, their use of the cardiovascular CPRs were not influenced by the evidence. For example, a young Family Physician who learned about the ACC/AHA pooled cohort equation during her residency training which she completed a few months ago continued to use the CPR despite knowing its inaccuracy:

> "And … the ASCVD, I still have learned a lot about … problems with it, like with under predicting or over predicting for non-white people and stuff like that. … I don't necessarily like just follow the guidelines. But I still use it. … But … I've learned more about how they're particularly inaccurate, maybe for certain populations. And that's been something that I haven't really known how to incorporate that into my practice, but something to just be aware of, I guess?" (BG35)

One of the reasons that she “had not known how to incorporate” the evidence of inaccuracy might be from challenges related to the complexity of evidence about cardiovascular CPRs, as she explained:

> “So … the papers about this are actually pretty complex because there's so many different aspects to it. But yeah, I think like a lot of sort of like internal validity stuff and then also, like external, we're talking about before about, like, does it apply to the population that you're actually treating.” (BG35)

Therefore, even when PCPs were aware of some research evidence, the research evidence had a limited role in PCPs’ decisions about using cardiovascular CPRs.

D. Evidence-unrelated factors

When a PCP does not rely on research evidence to make a decision about using a cardiovascular CPR, the PCP might depend on factors unrelated to research evidence to make the decision. PCPs interviewed decided whether or how they used cardiovascular CPRs based on six factors unrelated to research evidence: guideline recommendations, going along with others, feeling comfortable, own intuitive judgement, assuming infallibility, and features of cardiovascular CPRs. Many PCPs trusted guideline recommendations and relied on them to select cardiovascular CPRs that they used. A third-year resident in an Internal Medicine residency programme explained reasons for using the CHA_2DS_2-VASc score although he was unsure whether using it would improve care:

> “Um, really, it seems to be the one that is used by most of the guidelines. … I mean, it's part of the ACP guidelines. … In terms of actually knowing for certain, I mean, I suppose I don't. … Um, sometimes I just rely on what the guidelines say. Um, otherwise I don't know if I have quite, quite a good linear way of determining whether it's a useful tool or not.” (FA84)

Similarly, a Family Physician who did not look up or read any studies about the cardiovascular CPR he was using, assumed that he had to trust guideline developing organizations and “be on board with” their recommendations:

> “I bet I can't say because I read the research. … I think I assumed … if it's endorsed by, you know, the groups that have a stake in the matter, like the American Heart Association, or the American College of Cardiology, or the American Academy of Family Physicians. But I also understand that some sometimes there are political pressures with some of these organizations as well. So, so for that reason, you know, we tend to go with unbiased groups like the US Preventive Services Task Force recommendations, that kind of thing. … Yeah, I guess … there are people out there, they get paid to figure this stuff out. And, and I kind of have to trust them.” (IN63)

Some PCPs went along with other people’s decisions and chose cardiovascular CPRs that were a part of standard practice in their community and were used or suggested by their colleagues, consultants, pharmacists, and supervisors. An experienced Family Physician explained that he used cardiovascular CPRs typically expected to be used among medical staff members of his organisation, although research evidence supporting them might be outdated:

> “These are the ones that are community standard here. So, that's, and they may be a bit old, their research may have moved on from the ones we're using. But they're the ones that the medical staff here is most are most comfortable with most familiar with. … Yeah, if you want to talk to your consultant, those are the words, the magic words, they'll understand what you're trying to say.” (GR91)

Other PCPs used cardiovascular CPRs that they felt comfortable because they learned the CPRs in medical school or residency training programme, were aware of and familiar with the CPRs, and were already used to the CPRs. An Internist who completed his residency training eight years ago continued to use a cardiovascular CPR that he learned during his medical school and residency training and felt comfortable with, rather than evaluating the advantages and disadvantages of other available CPRs:

> "Probably just because… I've been using it for a long period of time. … I suspect it's just because it's become habit, that it's the one that I've chosen and decided to use on a regular basis. … I was introduced to it. … Yeah, probably did happen during training and throughout residency, medical school clinic clinical rotations. … I've come across a few others. And I suspect I just don't use them because I'm familiar with the ones that I have used." (IS78)

Some PCPs relied on their own intuitive judgement about the accuracy and effectiveness of cardiovascular CPRs rather than research evidence. A Family Physician who also had an editorial role for a medical journal, intuitively sensed that he improved care by using Wells scores when communicating with consultants or evaluating patients:

> "I generally haven't gone out and and researched the question on my own. … Um, ... improve my care. In some cases, it improves communication with others because I'm getting a lot of consultation for these problems. So yes, it improves care because I can communicate with my my consultants. The other places where it improves care is where I didn't use them, and things didn't go well. That's an, that's not research based, that's an intuitive thing. But having missed a PE now and again, when I forgot to use my … When I forgotten to use my Wells, … I've missed PE, and I get put back on the right track." (GR91)

Rather than making a judgement about the evidence behind cardiovascular CPRs, some PCPs assumed the cardiovascular CPRs they used were infallible, accurate, and effective. A teaching Family Physician who did not verify evidence behind the cardiovascular CPRs because he did not have the ability or time just assumed that the cardiovascular CPRs he used were as “perfect as it can be”:

> “Well, … I think this is, it's common in medicine, and one of the things we talk with our residents about is to assume that these things are infallible. You know, the lab tests are perfect, imaging studies are perfect, and prediction stuff is perfect as it can be. So, I probably just kind of assume that it's as good as we got. And it's probably pretty good if the patient fits into the demographic. … I guess we're kind of hypocritical a little bit in that. And we teach residents, you know, you should trust this information, but you should always verify, you know, trust but verify. But then we don't always do that. And why not? Well, maybe I don't know how, maybe I don't have the time. Maybe, I, yeah, it's probably both those things. I don't feel confident in assessing some of these things, myself.” (IN63)

Some interviewed PCPs chose to use cardiovascular CPRs that had sensible, available, and comprehensive predictors, assessed clinically relevant outcomes, were easy to use, and were useful when managing PCPs’ patients. A Family Physician who worked at a University clinic explained she used the ACC/AHA pooled cohort equation because it was convenient to use in point of care and included more comprehensive predictors:

> “Our EPIC has a cardiovascular disease risk calculator that's embedded in a dot phrase. … I can put in ‘.ascvd’ when I'm in a patient's chart, and… it will use the data that are in the chart to pull up their risk based on the most recent American College of Cardiology risk calculator. It used to use Framingham, but now it uses the ACC/AHA calculator.” (PX33)

These features of cardiovascular CPRs seemed to play an important role in PCPs' decisions about using cardiovascular CPRs when their decisions were not based on research evidence.

7.5. Discussion

7.5.1. Summary of findings

I identified two main themes relating to the influence of PCPs' perceptions of and experiences with cardiovascular CPR research on their decisions about using cardiovascular CPRs: "Seek and Judge" and "Be acquainted and Assume". Some PCPs were familiar with, trusted, or felt confident in using research evidence about cardiovascular CPRs. These favourable perceptions of and experiences with cardiovascular CPR research, which were often noted in more junior PCPs who recently completed their residency training, led some PCPs to actively "Seek and Judge" the evidence about cardiovascular CPRs. These PCPs appeared to involve research evidence about cardiovascular CPRs when deciding whether or how to use the CPRs. However, rather than depending on study design or internal validity, PCPs often relied on certain features of cardiovascular CPRs studies when making judgements about the evidence. Whether the evidence was applicable to their own patients seemed to be a key feature of research evidence for PCPs' judgements about the evidence.

On the other hand, some PCPs were unfamiliar with, distrusted, or found it challenging to use research evidence about cardiovascular CPRs. Surprisingly, these negative perceptions of and experiences with cardiovascular CPR research were often seen in older PCPs who had more clinical, teaching, or research experiences. Instead of actively “Seeking and Judging”, these PCPs tended to be passively acquainted with research evidence about cardiovascular CPRs through their training, colleagues, and online resources. These PCPs seemed to assume that someone else had evaluated the evidence, they could go with others’ interpretation of the evidence, or they could trust the evidence if it came from certain sources. Therefore, these PCPs did not rely on research evidence when deciding about using cardiovascular CPRs but looked at guideline recommendations, their intuition, or features of cardiovascular CPRs.

7.5.2. Comparison with existing literature

In an ethnographic study of two general practices by Gabbay and May (460), authors found that general practitioners relied on “Mindline” for their clinical decisions but never went through the linear steps of EBM during two years of intermittent observations. Contrary to this, one of the main themes, “Seek and Judge” suggests that PCPs might actively search for and evaluate research evidence about cardiovascular CPRs and use the information when making decisions about the CPRs. Although the theme, “Seek and Judge”, does not necessarily imply that PCPs are going through the linear processes of EBM, it is possible that use of research evidence in primary care has progressed since Gabby and May published their paper in 2004 (460).

Further, based on the findings that younger, more recently trained PCPs tended to express confidence in using research evidence than older, more experienced PCPs, it can be hypothesised that advances in the EBM education have affected PCPs perceptions about using research evidence about cardiovascular CPRs. Existing evidence seems to support this hypothesis. A 1998 national survey of Internal Medicine residency programmes in the USA showed that only 36.8% (99 of 269) offered a dedicated EBM curriculum (461). In a 2000 national survey of Emergency Medicine residency programmes in the USA, although 81.5% (53 of 65) of residency programmes included some degree of EBM teaching, 66.2% (43 of 65) spent less than five hours annually on EBM education (462). On the other hand, 84.6% (11 of 13) of Canadian Emergency Medicine residency programmes participated in a 2010 national survey offered formal EBM curricula, 61.5% (8 of 11) spent at least ten or more sessions annually, and 53.8% (7 of 13) had online EBM resources for education (463).

I also found that more senior PCPs with more clinical, research, or teaching experiences who might be in a position to teach EBM tended to be less confident about their ability to search for, appraise, and use research evidence about cardiovascular CPRs. Existing literature also consistently shows that a lack of appropriately trained enthusiastic faculties is one of the biggest barriers to EBM education (462-465). For example, in a 1999 national survey of Internal Medicine clerkship directors in the USA, inadequately trained faculties (82.6%,

90 of 109) and a lack of faculty enthusiasm (43.1%, 47 of 109) were identified as some of the barriers for EBM education (464).

One of the most important evidence related factors for PCPs' decisions about using cardiovascular CPRs in this chapter was their understandings from research evidence that cardiovascular CPRs were applicable to PCPs' own patients. Similarly, a survey and focus group study that evaluated German GPs' use of, attitude towards, and barriers of using CPRs also found that GPs disputed the applicability of evidence about cardiovascular CPRs for their specific patient populations (383).

Some interviewed PCPs' decision about using cardiovascular CPRs were influenced by factors unrelated to research evidence. One of these evidence-unrelated factors that influenced the PCPs' decisions about using cardiovascular CPRs was the feature of cardiovascular CPRs. Many existing studies also identified CPR related features such as whether CPRs included adequate predictors, evaluated relevant outcomes, were easy to use, or integrated into electronic health records (EHRs), as the key factors that influenced the use of CPRs (49, 97, 289, 381, 382, 385, 387, 404, 443, 447, 448, 450, 452, 453). The other evidence unrelated factor that influenced the PCPs decisions about using cardiovascular CPRs in this chapter was the PCPs' own intuitive judgement that using CPRs improved care, rather than based on evidence. Several existing studies also showed that the use of CPRs is promoted when users perceive CPRs are clinically helpful (289, 381, 384, 399,

404, 445, 450). Previous studies found that clinicians' knowledge of (382, 403, 448) and familiarity with (289, 384, 399) CPRs also influenced the use of CPRs which is consistent with this study's finding that PCPs used cardiovascular CPRs they felt comfortable with because they were aware, familiar, and using already. In an internet survey of Emergency Physicians, 78.0% of participants indicated that they were "more likely to use a CPR if it was used regularly by their colleagues" (450) which is also in line with PCPs in this chapter that used cardiovascular CPRs that were a part of standard practice in their community and were used or suggested by their colleagues, consultants, pharmacists, and supervisors.

7.5.3. Strengths and limitations

I purposively selected a maximum variation sample of 15 PCPs regarding factors that might determine perceptions and experiences with cardiovascular CPR research. Hence, I was able to identify two common themes among PCPs with a wide range of use of cardiovascular CPRs, research experience, and consideration of research evidence in decisions about using cardiovascular CPRs. However, because the participants were recruited from the WWAMI region Practice and Research Network (WPRN), PCPs with interests in research could have been overrepresented.

I only relied on PCP participants' statements to understand their perceptions, experiences, use of evidence, and decision-making process. Therefore, I do not know nor can verify whether their accounts were accurate. For example, I found

that some PCPs, especially those who were younger, recently trained, and less experienced, had confidence in searching for, reading, and appraising, and a tendency for using research evidence in their decision about using cardiovascular CPRs. It is possible that some of these PCPs felt they should give “correct” answers to questions, and the answers were different from what they indeed perceived, experienced, or did.

The findings from this chapter might have limited transferability to UK primary care because there are some unique differences in how cardiovascular CPRs are used between the USA and the UK. Firstly, although some cardiovascular CPRs, such as CHA2DS2-VASc score (184), are recommended by national guidelines from both countries (417, 418, 466), there are differences in CPRs recommended by the US and UK guidelines. For example, in the UK, QRISK2 (20) is included in the NICE guideline (17), whereas the US guideline (19) recommends the ACC/AHA pooled cohort equation (22). Therefore, clinicians' perceptions of and experiences with research about these CPRs might differ between the USA and the UK. Secondly, there might be differences in how cardiovascular CPRs are integrated into electronic health record (EHR) software. In the UK, certain cardiovascular CPRs are often integrated into GP’s EHR software which seems to promote the use of CPRs in practice (467). For example, a case study showed that QRISK2 is implemented in EHR software systems used by greater than 90% of general practices in the UK (468). On the other hand, less than half of primary care providers who completed the online questionnaire in this chapter indicated that at least one cardiovascular CPR was available in their EHR software. Also, the USA's EHR software systems often

tend to implement cardiovascular CPRs for venous thromboembolism or anticoagulation management, rather than CPRs for cardiovascular disease risk, and present these CPRs passively using non-interruptive methods (469). Therefore, clinicians' perceptions and uses of cardiovascular CPRs might be influenced by these differences.

Because there was little prior knowledge about this topic, I used an inductive approach to generate hypotheses about how PCP's perceptions of and experiences with cardiovascular CPR research influence their decisions about using cardiovascular CPRs. However, I did not test the credibility of these hypotheses. Although there is no reason to suspect that this chapter's findings might not apply to CPRs in other clinical domains, the findings' transferability should be confirmed in a new study.

7.5.4. Research and clinical implications

A systematic review showed that teaching EBM to trainees improved only knowledge but did not change attitude, skills, or behaviour (470). One of the reasons that the improvement is limited to knowledge might be because the focus of educational interventions to promote EBM and curricula for EBM in residency programmes is mainly on knowledge and skills of EBM steps (461, 471). Further, existing evidence about the effectiveness of educational interventions to promote EBM primarily evaluated the impact on knowledge and skills but had been less interested in assessing the effects on attitude, confidence, and behaviour (472, 473). Contrary to these practices in current

efforts to promote EBM, I found that PCPs' familiarity, trust, and confidence might play an important role in their encounter, evaluation, and use of research evidence about cardiovascular CPRs. Therefore, researchers designing interventions to promote the utilization of high-quality evidence in decisions about using cardiovascular CPRs should consider targeting the interventions to improve clinicians' familiarity, trust, and confidence regarding using research evidence.

All stakeholders involved in generating research evidence about cardiovascular CPRs, including researchers, journal editors, and funders, should understand unfamiliarity, distrust, and challenges regarding the use of research evidence might hinder clinicians form using research evidence in clinical decision making. These stakeholders of cardiovascular research should ensure that the evidence generated is high-quality and that it is easy for clinicians to access, read and understand, and use in their clinical decisions.

Those PCPs who are already familiar with, trust, and confident in using research evidence about cardiovascular CPRs might be able to maintain, promote, and reinforce their familiarity, trust, and confidence to use research evidence by continuing to seek, judge, and use research evidence in their decisions. However, understanding that many PCPs might depend on various factors unrelated to research evidence such as guideline recommendations and EHR integrations for choosing cardiovascular CPRs, guideline developers and policymakers should systematically evaluate existing evidence behind

cardiovascular CPRs and carefully base their recommendations about cardiovascular CPRs on the best available evidence.

7.5.5. Conclusion

CPRs are designed to be used by clinicians and patients to make evidence-informed healthcare decisions. Evidence from cardiovascular CPR research cannot be valuable unless it is utilised by clinicians, patients, or policymakers to support the use of cardiovascular CPRs. In this chapter, I found that PCPs' perceptions of and experiences with cardiovascular CPR research might influence the use of research evidence in this process. Based on this finding, future efforts to promote the use of research evidence in decisions about cardiovascular CPRs should consider targeting and evaluating clinicians' familiarity, trust, and confidence, in addition to knowledge and skills to use evidence. At the same time, I believe it is important to recognise clinicians' unfamiliarity, distrust, and challenges regarding the use of evidence might not be easily resolved. Because these clinicians often rely on factors other than research evidence such as guideline recommendations or integrations of CPRs into EHRs, stakeholders, especially guideline developers and policymakers, should ensure that their recommendations are based on the knowledge from systematically reviewing existing evidence of accuracy, generalizability, and impact.

Chapter 8. Discussion and conclusion

In this thesis, I evaluated research waste and inefficiency throughout all stages of cardiovascular CPR development and explored their causes and implications.

8.1. Summary of findings

In chapter 2, I found that authors of many derivation studies do not cite any existing CPR when justifying the need for a new cardiovascular CPR. The thematic content analysis showed justifications for deriving a new cardiovascular CPR are concerned with existing CPRs' derivation, construct, performance, transferability, and evidence, some of which could have been addressed by externally validating or updating the existing CPRs. I also found that, in the survey of the same derivation study authors, the vast majority of authors believe it is important to cite existing CPR when justifying a new cardiovascular CPR, which contradicts findings from the review of derivation studies that showed the majority of them do not cite any existing CPR. This might be one of the reasons researchers keep deriving many new cardiovascular CPRs despite similar CPRs for the same problem already exist.

I showed in chapter 3 that the probability for a CPR for cardiovascular disease risk to have an independent external validation study is low even many years after its derivation. This significantly limits the value of studies deriving new

CPRs, because CPRs cannot make an evidence-based contribution to clinical practice without strong evidence of generalizability. I also found that by providing the results of internal validation and sufficient information for risk calculation, authors of new cardiovascular risk CPRs might improve the chance for their CPRs to be externally validated and progress through CPR development stages towards a potential integration into evidence-based practice. Cardiovascular risk CPRs derived from studies that are conducted in the USA, use a larger sample size, and are published in a journal with a higher impact factor may also have a better chance of having an independent external validation study.

In chapter 4, focussing on Framingham Wilson CHD risk rule, I found that most external validation studies fail to generate meaningful evidence about Framingham Wilson CHD risk rule's generalizability to their full potential because authors of external validation studies do not update the CPR when it performs poorly, or external validation studies do not contribute to understanding the overall performance of Framingham Wilson CHD risk rule. Further, those external validation studies of Framingham Wilson CHD risk rule that do not update the CPR when it performs poorly are often used to justify deriving many new, potentially redundant, cardiovascular CPRs.

I systematically summarised the design, methods, and reporting of impact studies of cardiovascular CPRs in chapter 5. I found that many impact studies of cardiovascular CPRs are conducted using inappropriate study designs such

as uncontrolled before-after study. Even when impact studies are conducted using appropriate study designs, most have a substantial risk of bias in results or provide insufficient details in their reporting. Therefore, the vast majority of impact studies are research waste or inefficiency that are unusable in clinical decisions or future research. I also found that the quality of methods and reporting does not differ between impact studies of cardiovascular CPRs and studies assessing other types of nonpharmacologic interventions. Hence, poor methods and reporting seem to be a common feature of studies evaluating the effects of non-pharmacologic, complex interventions and not necessarily specific to impact studies of cardiovascular CPRs.

In Chapter 6, I surveyed GPs to evaluate which cardiovascular CPRs are recognised and used in UK primary care. I found that the vast majority of GPs are familiar with the QRISK score and use it to assess cardiovascular disease risk. For stroke and venous thromboembolism, most GPs recognise and use ABCD scores, CHADS scores, HAS-BLED score, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism. These cardiovascular CPRs are also the CPRs recommended by national guidelines in the UK and most frequently integrated into GPs' EHR software. Compared with the results of the 2014 study by Plüddemann et al. (3), I found that QRISK scores have become dominant CPRs for cardiovascular disease risk assessment in the UK while Framingham risk scores' popularity has waned. For stroke and venous thromboembolism, substantially more GPs use ABCD scores, CHADS scores, Wells scores for deep vein thrombosis, and Wells scores for pulmonary embolism, than GPs in the 2014 study (3).

In chapter 7, by analysing data from interviews with primary care providers (PCPs) in the USA, I identified two main themes relating to the influence of PCPs' perceptions of and experiences with cardiovascular CPR research on their decisions about using cardiovascular CPRs: "Seek and Judge" and "Be acquainted and Assume". When PCPs are familiar with, trust, or feel confident in using research evidence about cardiovascular CPRs, they tend to actively "Seek and Judge" the evidence about cardiovascular CPRs and involve research evidence in their decisions about whether or how they use the CPRs. On the other hand, when PCPs are unfamiliar with, distrust, or find it challenging to use research evidence about cardiovascular CPRs, they are passively acquainted with research evidence about cardiovascular CPRs and assume that someone else has evaluated the evidence, they can go with others' interpretation of the evidence, or they can trust the evidence if it is from certain sources. Instead of relying on research evidence when deciding about using cardiovascular CPRs, these PCPs tend to be influenced by guideline recommendations, their intuition, or features of cardiovascular CPRs.

In summary, I found that researchers might keep deriving new cardiovascular CPRs when similar CPRs exist because they do not examine existing evidence when creating a new cardiovascular CPR. Most of these cardiovascular CPRs lack evidence of generalizability because they do not get externally validated by independent researchers timely. Even when external validation studies are conducted, as shown in the case of Framingham Wilson CHD risk rule, most of

them fail to generate the evidence of generalizability to their full potential. Further, I found that the vast majority of impact studies are unusable because of flaws in study design, methods or reporting. GPs' use of cardiovascular CPRs in the UK has evolved and is focused on a few CPRs, such as the QRISK scores. PCPs' decisions about using cardiovascular CPRs seem to be influenced by their familiarity, trust, and confidence in using research evidence as well as factors unrelated to evidence such as guideline recommendations.

8.2. Strengths and limitations

8.2.1. Strengths

As part of my research, I approached the data using varied methods that best suited to answer research questions raised in each chapter which helped me gain a deeper understanding of the results obtained. The strengths of each research project have been discussed in the discussion section of chapters 2,3,4,5,6, and 7. Therefore, I briefly summarise some of the key strengths here.

In chapter 2, I conducted a mixed-methods study to explore why researchers might keep deriving new cardiovascular CPRs despite similar CPRs for the same problem already existed. Using mixed-methods, I was able to further augment the findings from the review of derivation studies by thematic content analysis and author survey.

In chapter 3, instead of merely assessing the proportion of the cardiovascular CPRs with an independent external validation study, I conducted survival analysis because it was important to evaluate how promptly a cardiovascular CPR was externally validated by independent researchers after its derivation. As a result, I was able to show that cardiovascular CPR's probability of having a timely independent external validation is low.

I primarily used forward citation search of derivation studies in chapters 3, 4, and 5 to identify external validation and impact studies of cardiovascular CPRs, which was supplemented by electronic database searches when necessary. Using this strategy, I collected larger numbers of external validation and impact studies than those included in previously existing systematic reviews (53, 236) which used exhaustive electric database searches. This shows the efficiency and robustness of the forward citation search used in three chapters of this thesis as the method to identify external validation and impact studies when conducting systematic reviews of CPR studies.

I tried to make the survey methods used in chapter 6 as equivalent as possible to the ones used in the 2014 study by Plüddemann et al. (3) so that it was possible to compare the results between the 2014 and 2019 studies directly. As a result, I was able to evaluate temporal trends in the familiarity with and use of cardiovascular CPRs among GPs in the UK. I adjusted all *p*-values and CIs using the Bonferroni methods to account for the multiple statistical significance tests. Even after these adjustments, all comparisons but one remained

statistically significant, indicating these observed trends in the familiarity and use of cardiovascular CPRs are likely real.

In chapter 7, I purposively selected a maximum variation sample of PCPs regarding factors that might determine their perceptions of and experiences with cardiovascular CPR research. This sampling strategy enabled me to identify two common themes using rich data generated from interviews with PCPs who had diverse uses of cardiovascular CPRs, experiences with research, and considerations of research evidence in decisions about using cardiovascular CPRs.

Lastly, I investigated research waste and inefficiency by following the entire stages of CPR development from derivation to external validation to impact assessment to uptake in practice. Therefore, through this thesis's findings, I was able to broaden the understandings about how research waste and inefficiency are generated throughout the “evidence pipeline” (474) of cardiovascular CPR development.

8.2.2. Limitations

Each research project's limitations have been discussed in the discussion section of chapters 2,3,4,5,6, and 7. Here, I briefly summarise some of the key limitations.

I used the International Register of Clinical Prediction Rules for Primary Care (47) as the information source for cardiovascular CPRs and their derivation studies in chapters 2 and 5. Although the International Register was the most comprehensive source of CPRs relevant to primary care at the time, it is possible that some cardiovascular CPRs were missed. I evaluated derivation studies of cardiovascular CPRs from the original version of the International Register in chapter 2, which only included cardiovascular CPRs published before 2010 (47). The updated version of the International Register was used in chapter 5, which had cardiovascular CPRs published up to 2015. Therefore, the findings from these chapters may not apply to cardiovascular CPRs published more recently.

Another limitation in chapter 2 was the small number of authors surveyed which might have made it difficult to detect the associations between citation of existing CPR and the survey answers. In retrospect, this could have been prevented by estimating the necessary sample size for planned statistical analyses and reflecting that into the survey design at the protocol stage.

I could not evaluate predictor variables in a multivariable analysis in chapter 3 because there were a limited number of cardiovascular risk CPRs and cardiovascular CPRs with an independent external validation study. Therefore, findings from univariable analyses should be interpreted as hypothesis-generating.

In chapter 4, I only assessed research inefficiency from external validation studies that do not update Framingham Wilson CHD risk rule when it performs poorly and from external validation studies that do not contribute to understanding the overall performance of Framingham Wilson CHD risk rule. I did not consider other sources of waste and inefficiency, such as design, methods, reporting, or publication. Therefore, despite finding the majority of external validation studies in chapter 4 to be examples of research inefficiency, it is very likely that I underestimated the amount of research waste and inefficiency among external validation studies of Framingham Wilson CHD risk rule.

In chapter 5, I used tools for assessing the risk of bias and recommendations from reporting guidelines to evaluate the quality of methods and reporting among impact studies of cardiovascular CPRs. Assessing the risk of bias and reporting quality using these tools was challenging because it often required making the best judgements based on the limited information provided in impact studies. Although my interpretations and judgements were adjudicated by the reviewers who independently evaluated impact studies, it is unclear whether findings can be easily replicated by other researchers (475).

I recruited GPs from doctors.net.uk for compatibility with the 2014 study in chapter 6. However, there were several weaknesses in this approach. Although doctors.net.uk had a large number of GP members across all regions of the UK, GP participants recruited from doctors.net.uk might not have been truly

representative. Hence, the findings may not be generalisable to all GPs in the UK. Also, the recruitment mechanism did not allow me to calculate the response rate nor investigate whether GPs who participated differed from those who did not.

It would have been ideal for me to interview GPs in the UK in chapter 7 so that findings from chapters 6 and 7 are consistently applicable to primary care in the UK. However, constraints in research funding prohibited me from recruiting GPs in the UK. Also, this was an exploratory qualitative interview study with thematic analysis to generate hypotheses about how PCPs' perceptions of and experiences with cardiovascular CPR research influence their decisions about using cardiovascular CPRs. Therefore, the hypotheses generated in this chapter have not been further investigated.

Lastly, the scope of this thesis is limited to cardiovascular CPR relevant to primary care. Therefore, knowledge about research waste and inefficiency in cardiovascular CPRs gained from this thesis may not necessarily be generalisable to CPRs in other clinical areas.

8.3. Implications of this thesis

8.3.1. A framework for CPR research

In chapter 1, I described research inefficiency from disparities in the evidence accumulation for the different stages of cardiovascular CPR development. Here, I propose the following framework in Figure 8.3.1 to reduce research waste and inefficiency from disparities in evidence accumulation, increase the value of CPR research, and promote the implementation of CPRs supported by evidence. I constructed this framework by building on the framework to reduce the development of redundant CPRs presented in chapter 2, Figure 2.5.1, and updating it with findings from succeeding chapters.

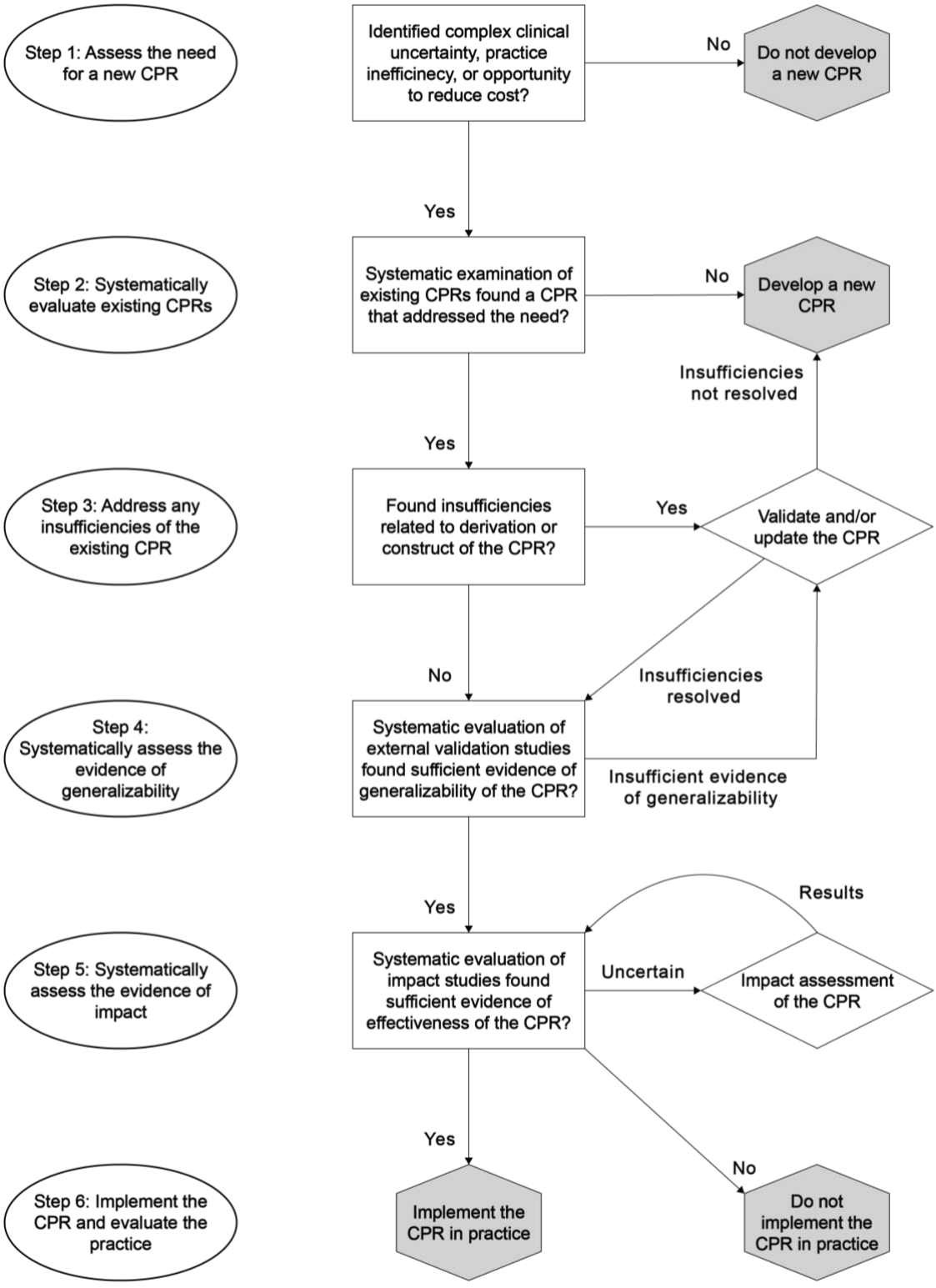


Figure 8.3.1 A framework to reduce research waste and inefficiency from disparities in evidence accumulation, increase the value of clinical prediction rule (CPR) research, and promote the implementation of CPRs supported by evidence.

CPR research has been disproportionately focused on deriving a new CPR compared with externally validating and assessing the impact of existing CPRs (7, 47). I recommend researchers who want to derive a new CPR consider whether the CPR is necessary by going through the following steps, starting at step 1.

Step 1: Researchers should establish how a CPR could be useful for the clinical context. A CPR might be helpful for complicated clinical situations with uncertainties, practice inefficiencies, or cost-saving opportunities. Researchers should not derive a new CPR without understanding how the CPR might contribute to practice.

Step 2: Once how a CPR could help practice is recognised, researchers should systematically search for existing CPRs relevant to the clinical context which can be done by finding an up-to-date systematic review, updating an existing systematic review, or, when needed, conducting a new systematic review of CPRs for the clinical context. Researchers should derive a new CPR when no relevant CPR is found.

Step 3: When a CPR is found in the systematic search, researchers should evaluate the CPR and determine whether it was created using valid methods and is sufficient for given clinical context regarding predictors, outcomes, presentations, and usability. Researchers should attempt to resolve any

insufficiencies by externally validating and updating the CPR. Deriving a new CPR is justified when the insufficiencies cannot be resolved.

Step 4: When no insufficiency of the CPR's derivation or construct was identified, researchers should systematically evaluate whether there is sufficient evidence of generalizability for the CPR. Researchers should then consider externally validating the CPR and updating it as needed to address any insufficiencies of the CPR's generalizability.

Step 5: For CPRs with sufficient evidence of generalizability, researchers can systematically evaluate the evidence of effectiveness. When a CPR with sufficient evidence of effectiveness is found, the CPR can be implemented in practice. However, when the evidence is uncertain about whether implementing a CPR can improve practice, researchers should consider evaluating the CPR's effectiveness in an impact analysis. When there is sufficient evidence that using a CPR is ineffective, the CPR should not be implemented in practice.

Researchers interested in externally validating a CPR can also use this framework to evaluate the external validation study's need by starting at step 4. Similarly, researchers who want to conduct an impact analysis of a CPR can start at step 5 of this framework to evaluate the impact study's need.

Further, funders and journals can also use this framework to determine the priorities of studies evaluating CPRs, and promote and disseminate necessary independent external validation studies, impact studies, and systematic reviews of CPR studies. This framework might also help guideline developers, policymakers, and clinicians identify gaps in evidence of a CPR they are interested in adopting.

8.3.2. Reducing waste and inefficiency within each stage of CPR development

Many systematic reviews have shown that the evidence-base supporting the use of CPRs at derivation (4, 50, 57-67, 71, 73, 284) and external validation (58, 60, 65, 66, 79, 233) stages is weak because of flaws in design, methods, and reporting. Although evaluating the quality of design, methods, or reporting of studies deriving or externally validating cardiovascular CPRs was not the primary focus of this thesis, these trends were visible in findings of chapters 3 and 4. For example, in chapter 3, most derivation studies of cardiovascular risk CPRs did not clearly describe participants and outcomes, and, in chapter 4, complete case analysis was used to handle missing data in the majority of external validation studies of Framingham Wilson CHR risk rule. Promoting the adoption of Prediction model Risk Of Bias ASsessment Tool (PROBAST) and the TRIPOD statement by researchers might reduce the flaws in design, methods, and reporting among derivation and external validation studies, and strengthen evidence-base to support the use of cardiovascular CPRs at derivation and external validation stages.

Impact studies of cardiovascular CPRs are rarely conducted (47, 91), and as found in chapter 5, even those impact studies are of poor quality due to inappropriate study design, suboptimal methods, and unclear reporting. Researchers need to recognize that study designs appropriate for impact assessment of a CPR are different from proper study designs for derivation and external validation studies. Researchers also should consider using methodological standards and reporting guidelines relevant to impact studies so that the evidence they generate is credible, understandable, and usable in clinical decision making or future research. Journals and funders could play a crucial role in reducing waste and inefficiency in cardiovascular CPR research by motivating researchers to shift their research conduct and publication.

8.3.3. Increasing the uptake of CPRs supported by evidence

As we expect all pharmacologic interventions to be developed and evaluated in a rigorous, multiphase process before they are used in practice (374), the same standards should apply when developing and assessing any non-pharmacologic interventions (375), including cardiovascular CPRs. Unfortunately, clinicians interested in using cardiovascular CPRs should be aware that, although there might be many redundant CPRs for the same purposes, cardiovascular CPRs externally validated by independent researchers and assessed in high-quality impact studies are rare. Further, findings from chapters 3 and 5 showed that even some of the widely used cardiovascular CPRs lack sufficient evidence of generalizability or effectiveness to support their use. For example, although

nearly all GPs who took part in the survey in chapter 6 indicated they used QRISK scores, I did not find any impact study that evaluated the QRISK scores using appropriate study design in chapter 5. Therefore, cardiovascular CPRs should be implemented only after accumulating the sufficient evidence of accuracy, generalizability, and effectiveness through evaluations with the same level of rigour expected in the assessment of pharmacologic interventions.

The findings of chapter 7 suggest that the clinicians' uptake of CPRs informed by research evidence might be promoted by improving their familiarity, trust, and confidence regarding the use of CPR research evidence. This hypothesis could be tested by developing an educational intervention and evaluating it in an experiment. Further, findings from chapters 6 and 7 indicate that factors other than research evidence such as guideline recommendations and integrations of CPR into EHR might influence the use of CPRs. Therefore, guideline developers and policymakers should systematically evaluate the evidence behind CPRs and recommend or implement CPRs supported by evidence of accuracy, generalizability, and effectiveness. This hypothesis should also be tested in an experiment using interrupted time series study or other appropriate designs that allow adequate evaluation.

8.4. Conclusion

Research waste and inefficiency can arise from disparities in the accumulation of evidence throughout the stages of cardiovascular CPR development and

from poor design, methods, and reporting of cardiovascular CPR research in each stage of cardiovascular CPR development.

By systematically evaluating existing evidence and rigorously assessing the need for any new cardiovascular CPR research, researchers might be able to reduce research waste and inefficiency from disparities in evidence accumulation, increase the value of CPR research, and promote the implementation of CPRs supported by evidence. Journals and funders could play a crucial role in shifting the paradigm of cardiovascular CPR research and improving research conduct and publication. At the same time, endeavours to promote the uptake of cardiovascular CPRs supported by research evidence should further probe the role of improving clinicians' familiarity, trust, and confidence regarding the use of CPR research evidence as well as providing directions to clinicians via guideline recommendations or EHR integrations.

References

1. 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.
2. 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.
3. Plüddemann 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. Br J Gen Pract. 2014;64(621):e233-42.
4. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. JAMA : the journal of the American Medical Association. 1997;277(6):488-94.
5. Adams ST, Leveson SH. Clinical prediction rules. BMJ. 2012;344:d8312.
6. 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.
7. 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.
8. 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.
9. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res. 2018;2:11.
10. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagn Progn Res. 2019;3:16.
11. Little P, Hobbs FD, Moore M, Mant D, Williamson I, McNulty C, et al. PRImary care Streptococcal Management (PRISM) study: in vitro study, diagnostic cohorts and a pragmatic adaptive randomised controlled trial with nested qualitative study and cost-effectiveness study. Health Technol Assess. 2014;18(6):vii-xxv, 1-101.
12. Little P, Hobbs FD, Moore M, Mant D, Williamson I, McNulty C, et al. Clinical score and rapid antigen detection test to guide antibiotic use for sore throats: randomised controlled trial of PRISM (primary care streptococcal management). BMJ. 2013;347:f5806.
13. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58(5):377-82.
14. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. The New England journal of medicine. 1997;336(4):243-50.

15. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Respir Crit Care Med. 2019;200(7):e45-e67.
16. Lim WS, Baudouin SV, George RC, Hill AT, Jamieson C, Le Jeune I, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax. 2009;64 Suppl 3:iii1-55.
17. National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modifification 2014 [updated September 2016; cited 2019 OCtober 17]. Available from: https://www.nice.org.uk/guidance/cg181.
18. Scottish Intercollegiate Guidelines Network. Risk estimation and the prevention of cardiovascular disease : a national clinical guidleine 2017 [cited 2019 October 18]. Available from: https://www.sign.ac.uk/sign-149-risk-estimation-and-the-prevention-of-cardiovascular-disease.html.
19. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019;140(11):e596-e646.
20. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ. 2008;336(7659):1475.
21. Woodward M, Brindle P, Tunstall-Pedoe H, estimation Sgor. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart. 2007;93(2):172-6.
22. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/ AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/ American Heart Association Task Force on Practice Guidelines. Journal of the American College of Cardiology. 2014;63(25 Pt B):2935.
23. Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, et al. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). BMC medical informatics and decision making. 2011;11:62.
24. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98(9):683-90.
25. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-38.
26. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Annals of internal medicine. 1999;130(6):515-24.
27. 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.

28. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. Journal of clinical epidemiology. 2001;54(8):774-81.
29. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013;13:33.
30. Altman DG, Royston P. What do we mean by validating a prognostic model? Statistics in medicine. 2000;19(4):453-73.
31. 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.
32. 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.
33. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ. 2009;338:b606.
34. Ioannidis JP. Why Most Clinical Research Is Not Useful. PLoS medicine. 2016;13(6):e1002049.
35. Glasziou P, Chalmers I. The BMJ Opinion [Internet]: BMJ Publishing Group. 2016. [cited 2019].
36. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet. 2009;374(9683):86-9.
37. Altman DG. The scandal of poor medical research. BMJ. 1994;308(6924):283-4.
38. 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.
39. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014;383(9912):166-75.
40. Al-Shahi Salman R, Beller E, Kagan J, Hemminki E, Phillips RS, Savulescu J, et al. Increasing value and reducing waste in biomedical research regulation and management. Lancet. 2014;383(9912):176-85.
41. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet. 2014;383(9913):257-66.
42. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. Lancet. 2014;383(9913):267-76.
43. James Lind Alliance. About Priority Setting Partnerships [cited 2020 6 October]. Available from: https://www.jla.nihr.ac.uk/about-the-james-lind-alliance/about-psps.htm.
44. Lund H, Brunnhuber K, Juhl C, Robinson K, Leenaars M, Dorch BF, et al. Towards evidence based research. BMJ. 2016;355:i5440.
45. Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. BMC Med. 2010;8:24.

46. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: increasing value, reducing waste. Lancet. 2014;383(9912):101-4.
47. 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.
48. Muthee TB, Kimathi D, Richards G, Roberts N, Williams V, Nunan D, et al. Factors affecting the implementation of cardiovascular risk scoring in primary care; a mixed-method systematic review. 2019.
49. Kennedy G, Gallego B. Clinical prediction rules: A systematic review of healthcare provider opinions and preferences. Int J Med Inform. 2019;123:1-10.
50. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. The New England journal of medicine. 1985;313(13):793-9.
51. 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.
52. Wessler BS, Paulus J, Lundquist CM, Ajlan M, Natto Z, Janes WA, et al. Tufts PACE Clinical Predictive Model Registry: update 1990 through 2015. Diagn Progn Res. 2017;1:20.
53. 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.
54. The Predictive Analytics and Comparative Effectiveness (PACE) Center. Tufts PACE Clinical Predictive Model Registry [cited 2020 25 October]. Available from: https://www.nice.org.uk/guidance/cg181.
55. Alvarado A. A practical score for the early diagnosis of acute appendicitis. Ann Emerg Med. 1986;15(5):557-64.
56. 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.
57. Mushkudiani NA, Hukkelhoven CW, Hernandez AV, Murray GD, Choi SC, Maas AI, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. Journal of clinical epidemiology. 2008;61(4):331-43.
58. Altman DG. Prognostic models: a methodological framework and review of models for breast cancer. Cancer Invest. 2009;27(3):235-43.
59. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. BMC Med. 2011;9:103.
60. 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.
61. Hussain A, Choukairi F, Dunn K. Predicting survival in thermal injury: a systematic review of methodology of composite prediction models. Burns. 2013;39(5):835-50.
62. 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.
63. Wen Z, Guo Y, Xu B, Xiao K, Peng T, Peng M. Developing Risk Prediction Models for Postoperative Pancreatic Fistula: a Systematic Review of Methodology and Reporting Quality. Indian J Surg. 2016;78(2):136-43.
64. Sahle BW, Owen AJ, Chin KL, Reid CM. Risk Prediction Models for Incident Heart Failure: A Systematic Review of Methodology and Model Performance. J Card Fail. 2017;23(9):680-7.
65. Heus P, Damen J, Pajouheshnia R, Scholten R, Reitsma JB, Collins GS, et al. Poor reporting of multivariable prediction model studies: towards a targeted implementation strategy of the TRIPOD statement. BMC Med. 2018;16(1):120.
66. Tan J, Qi Y, Liu C, Xiong Y, He Q, Zhang G, et al. The use of rigorous methods was strongly warranted among prognostic prediction models for obstetric care. Journal of clinical epidemiology. 2019;115:98-105.
67. Bradley A, Van Der Meer R, McKay CJ. A systematic review of methodological quality of model development studies predicting prognostic outcome for resectable pancreatic cancer. BMJ Open. 2019;9(8):e027192.
68. van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. Journal of clinical epidemiology. 2006;59(10):1102-9.
69. Janssen KJ, Donders AR, Harrell FE, Jr., Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. Journal of clinical epidemiology. 2010;63(7):721-7.
70. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. Journal of clinical epidemiology. 2010;63(2):205-14.
71. Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. BMC Med. 2010;8:20.
72. Steyerberg EW. Clinical prediction models : a practical approach to development, validation, and updating. 2nd ed. Cham, Switzerland: Springer; 2019.
73. Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. BMC Med. 2010;8:21.
74. 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.
75. Medlock S, Ravelli AC, Tamminga P, Mol BW, Abu-Hanna A. Prediction of mortality in very premature infants: a systematic review of prediction models. PLoS One. 2011;6(9):e23441.
76. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. Breast Cancer Res Treat. 2012;132(2):365-77.
77. Jaja BN, Cusimano MD, Etminan N, Hanggi D, Hasan D, Ilodigwe D, et al. Clinical prediction models for aneurysmal subarachnoid hemorrhage: a systematic review. Neurocrit Care. 2013;18(1):143-53.

78. Vuong K, McGeechan K, Armstrong BK, Cust AE. Risk prediction models for incident primary cutaneous melanoma: a systematic review. JAMA Dermatol. 2014;150(4):434-44.
79. 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 Med Res Methodol. 2014;14:40.
80. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Statistics in medicine. 2016;35(2):214-26.
81. 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.
82. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. Journal of clinical epidemiology. 2014.
83. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. Journal of clinical epidemiology. 2016;74:167-76.
84. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic t, et al. Calibration: the Achilles heel of predictive analytics. BMC Med. 2019;17(1):230.
85. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med. 2007;35(9):2052-6.
86. May S, Hosmer DW. A cautionary note on the use of the Gronnesby and Borgan goodness-of-fit test for the Cox proportional hazards model. Lifetime Data Anal. 2004;10(3):283-91.
87. Hosmer DW, Hjort NL. Goodness-of-fit processes for logistic regression: simulation results. Statistics in medicine. 2002;21(18):2723-38.
88. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36.
89. Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. Biom J. 2008;50(4):457-79.
90. Maguire JL, Kulik DM, Laupacis A, Kuppermann N, Uleryk EM, Parkin PC. Clinical prediction rules for children: a systematic review. Pediatrics. 2011;128(3):e666-77.
91. Wallace E, Uijen MJ, Clyne B, Zarabzadeh A, Keogh C, Galvin R, et al. Impact analysis studies of clinical prediction rules relevant to primary care: a systematic review. BMJ Open. 2016;6(3):e009957.
92. Sanders SL, Rathbone J, Bell KJL, Glasziou PP, Doust JA. Systematic review of the effects of care provided with and without diagnostic clinical prediction rules. Diagn Progn Res. 2017;1:13.
93. Collins DR, Tompson AC, Onakpoya IJ, Roberts N, Ward AM, Heneghan CJ. Global cardiovascular risk assessment in the primary prevention of cardiovascular disease in adults: systematic review of systematic reviews. BMJ Open. 2017;7(3):e013650.
94. 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.

95. Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk prediction in patients with heart failure: a systematic review and analysis. JACC Heart Fail. 2014;2(5):440-6.
96. Ban J-W, Stevens R, Perera R. Predictors for independent external validation of cardiovascular risk clinical prediction rules: Cox proportional hazards regression analyses. Diagnostic and Prognostic Research. 2018;2(1):3.
97. 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.
98. Clarke M. Doing new research? Don't forget the old. PLoS medicine. 2004;1(2):e35.
99. World Organization of National Colleges Academies and Academic Associations of General Practitioners/Family Physicians. Classification Committee. ICPC-2 : international classification of primary care. 2nd ed. Oxford: Oxford University Press; 1998. x, 190 p. p.
100. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009;338:b604.
101. Steyerberg EW. Clinical prediction models : a practical approach to development, validation, and updating. New York: Springer; 2009. xxviii, 497 p. p.
102. Green J, Thorogood N. Qualitative Methods for Health Research. London: SAGE Publications Ltd,; 2018.
103. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-88.
104. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-12.
105. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.
106. Norris RM, Brandt PW, Caughey DE, Lee AJ, Scott PJ. A new coronary prognostic index. Lancet. 1969;1(7589):274-8.
107. Kernan WN, Horwitz RI, Brass LM, Viscoli CM, Taylor KJ. A prognostic system for transient ischemia or minor stroke. Annals of internal medicine. 1991;114(7):552-7.
108. Wicki J, Perneger TV, Junod AF, Bounameaux H, Perrier A. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. Arch Intern Med. 2001;161(1):92-7.
109. Wang TJ, Massaro JM, Levy D, Vasan RS, Wolf PA, D'Agostino RB, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. JAMA : the journal of the American Medical Association. 2003;290(8):1049-56.
110. Subramaniam RM, Chou T, Heath R, Allen R. Importance of pretest probability score and D-dimer assay before sonography for lower limb deep venous thrombosis. AJR Am J Roentgenol. 2006;186(1):206-12.
111. Hoellerich VL, Wigton RS. Diagnosing pulmonary embolism using clinical findings. Arch Intern Med. 1986;146(9):1699-704.
112. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score

for the United Kingdom: prospective open cohort study. BMJ. 2007;335(7611):136.
113. Pryor DB, Harrell FE, Jr., Lee KL, Califf RM, Rosati RA. Estimating the likelihood of significant coronary artery disease. Am J Med. 1983;75(5):771-80.
114. Tierney WM, Roth BJ, Psaty B, McHenry R, Fitzgerald J, Stump DL, et al. Predictors of myocardial infarction in emergency room patients. Crit Care Med. 1985;13(7):526-31.
115. Mark DB, Hlatky MA, Harrell FE, Jr., Lee KL, Califf RM, Pryor DB. Exercise treadmill score for predicting prognosis in coronary artery disease. Annals of internal medicine. 1987;106(6):793-800.
116. McNamara RF, Carleen E, Moss AJ. Estimating left ventricular ejection fraction after myocardial infarction by various clinical parameters. Am J Cardiol. 1988;62(4):192-6.
117. Selker HP, Griffith JL, D'Agostino RB. A time-insensitive predictive instrument for acute myocardial infarction mortality: a multicenter study. Med Care. 1991;29(12):1196-211.
118. Hubbard BL, Gibbons RJ, Lapeyre AC, 3rd, Zinsmeister AR, Clements IP. Identification of severe coronary artery disease using simple clinical parameters. Arch Intern Med. 1992;152(2):309-12.
119. Silver MT, Rose GA, Paul SD, O'Donnell CJ, O'Gara PT, Eagle KA. A clinical rule to predict preserved left ventricular ejection fraction in patients after myocardial infarction. Annals of internal medicine. 1994;121(10):750-6.
120. Morise AP, Haddad WJ, Beckner D. Development and validation of a clinical score to estimate the probability of coronary artery disease in men and women presenting with suspected coronary disease. Am J Med. 1997;102(4):350-6.
121. Selker HP, Griffith JL, Beshansky JR, Schmid CH, Califf RM, D'Agostino RB, et al. Patient-specific predictions of outcomes in myocardial infarction for real-time emergency use: a thrombolytic predictive instrument. Annals of internal medicine. 1997;127(7):538-56.
122. Kidwell CS, Saver JL, Schubert GB, Eckstein M, Starkman S. Design and retrospective analysis of the Los Angeles Prehospital Stroke Screen (LAPSS). Prehosp Emerg Care. 1998;2(4):267-73.
123. Wells PS, Ginsberg JS, Anderson DR, Kearon C, Gent M, Turpie AG, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. Annals of internal medicine. 1998;129(12):997-1005.
124. Pocock SJ, McCormack V, Gueyffier F, Boutitie F, Fagard RH, Boissel JP. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. BMJ. 2001;323(7304):75-81.
125. Stevens RJ, Kothari V, Adler AI, Stratton IM, United Kingdom Prospective Diabetes Study G. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci (Lond). 2001;101(6):671-9.
126. Wang SJ, Ohno-Machado L, Fraser HS, Kennedy RL. Using patient-reportable clinical history factors to predict myocardial infarction. Comput Biol Med. 2001;31(1):1-13.
127. Farkouh M, Oddone E, Simel D. Improving the clinical examination for a low ankle-brachial index. International Journal of Angiology. 2002;11(1):41-5.

128. Moons KG, Bots ML, Salonen JT, Elwood PC, Freire de Concalves A, Nikitin Y, et al. Prediction of stroke in the general population in Europe (EUROSTROKE): Is there a role for fibrinogen and electrocardiography? J Epidemiol Community Health. 2002;56 Suppl 1:i30-6.
129. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J. 2003;24(11):987-1003.
130. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model. JAMA : the journal of the American Medical Association. 2003;290(19):2581-7.
131. van Walraven C, Hart RG, Wells GA, Petersen P, Koudstaal PJ, Gullov AL, et al. A clinical prediction rule to identify patients with atrial fibrillation and a low risk for stroke while taking aspirin. Arch Intern Med. 2003;163(8):936-43.
132. Auble TE, Hsieh M, Gardner W, Cooper GF, Stone RA, McCausland JB, et al. A prediction rule to identify low-risk patients with heart failure. Acad Emerg Med. 2005;12(6):514-21.
133. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. Derivation and validation of a prognostic model for pulmonary embolism. Am J Respir Crit Care Med. 2005;172(8):1041-6.
134. Rothwell PM, Giles MF, Flossmann E, Lovelock CE, Redgrave JN, Warlow CP, et al. A simple score (ABCD) to identify individuals at high early risk of stroke after transient ischaemic attack. Lancet. 2005;366(9479):29-36.
135. Christenson J, Innes G, McKnight D, Thompson CR, Wong H, Yu E, et al. A clinical prediction rule for early discharge of patients with chest pain. Ann Emerg Med. 2006;47(1):1-10.
136. Lauer MS, Pothier CE, Magid DJ, Smith SS, Kattan MW. An externally validated model for predicting long-term survival after exercise treadmill testing in patients with suspected coronary artery disease and a normal electrocardiogram. Annals of internal medicine. 2007;147(12):821-8.
137. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA : the journal of the American Medical Association. 2007;297(6):611-9.
138. Uresandi F, Otero R, Cayuela A, Cabezudo MA, Jimenez D, Laserna E, et al. [A clinical prediction rule for identifying short-term risk of adverse events in patients with pulmonary thromboembolism]. Arch Bronconeumol. 2007;43(11):617-22.
139. Gagne P, Simon L, Le Pape F, Bressollette L, Mottier D, Le Gal G, et al. [Clinical prediction rule for diagnosing deep vein thrombosis in primary care]. Presse Med. 2009;38(4):525-33.
140. Clayton TC, Lubsen J, Pocock SJ, Voko Z, Kirwan BA, Fox KA, et al. Risk score for predicting death, myocardial infarction, and stroke in patients with stable angina, based on a large randomised trial cohort of patients. BMJ. 2005;331(7521):869.
141. Kline JA, Mitchell AM, Kabrhel C, Richman PB, Courtney DM. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. Journal of thrombosis and haemostasis : JTH. 2004;2(8):1247-55.

142. Younge BR, Cook BE, Jr., Bartley GB, Hodge DO, Hunder GG. Initiation of glucocorticoid therapy: before or after temporal artery biopsy? Mayo Clin Proc. 2004;79(4):483-91.
143. Daly CA, De Stavola B, Sendon JL, Tavazzi L, Boersma E, Clemens F, et al. Predicting prognosis in stable angina--results from the Euro heart survey of stable angina: prospective observational study. BMJ. 2006;332(7536):262-7.
144. Chan WS, Lee A, Spencer FA, Crowther M, Rodger M, Ramsay T, et al. Predicting deep venous thrombosis in pregnancy: out in "LEFt" field? Annals of internal medicine. 2009;151(2):85-92.
145. Allen CM. Predicting the outcome of acute stroke: a prognostic score. J Neurol Neurosurg Psychiatry. 1984;47(5):475-80.
146. Walker NJ, Sites FD, Shofer FS, Hollander JE. Characteristics and outcomes of young adults who present to the emergency department with chest pain. Acad Emerg Med. 2001;8(7):703-8.
147. Oudega R, Moons KG, Hoes AW. Ruling out deep venous thrombosis in primary care. A simple diagnostic algorithm including D-dimer testing. Thromb Haemost. 2005;94(1):200-5.
148. Zhang XF, Attia J, D'Este C, Yu XH, Wu XG. A risk score predicted coronary heart disease and stroke in a Chinese cohort. Journal of clinical epidemiology. 2005;58(9):951-8.
149. Coppola WG, Whincup PH, Papacosta O, Walker M, Ebrahim S. Scoring system to identify men at high risk of stroke: a strategy for general practice. Br J Gen Pract. 1995;45(393):185-9.
150. Jain AP, Gupta OP, Jajoo UN, Kumar A, Trivedi SK. Evaluation of clinical variables as predicators of left ventricular function in acute myocardial infarction. J Assoc Physicians India. 1993;41(1):17-9.
151. Stoffers HE, Kester AD, Kaiser V, Rinkens PE, Knottnerus JA. Diagnostic value of signs and symptoms associated with peripheral arterial occlusive disease seen in general practice: a multivariable approach. Medical decision making : an international journal of the Society for Medical Decision Making. 1997;17(1):61-70.
152. Lumley T, Kronmal RA, Cushman M, Manolio TA, Goldstein S. A stroke prediction score in the elderly: validation and Web-based application. Journal of clinical epidemiology. 2002;55(2):129-36.
153. Zamanipoor Najafabadi AH, Ramspek CL, Dekker FW, Heus P, Hooft L, Moons KGM, et al. TRIPOD statement: a preliminary pre-post analysis of reporting and methods of prediction models. BMJ Open. 2020;10(9):e041537.
154. Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. Journal of the Royal Society of Medicine. 2007;100(4):187-90.
155. Pandis N, Fleming PS, Koletsi D, Hopewell S. The citation of relevant systematic reviews and randomised trials in published reports of trial protocols. Trials. 2016;17(1):581.
156. Johnson AL, Walters C, Gray H, Torgerson T, Checketts JX, Boose M, et al. The use of systematic reviews to justify orthopaedic trauma randomized controlled trials: A cross-sectional analysis. Injury. 2020;51(2):212-7.
157. Torgerson T, Evans S, Johnson BS, Vassar M. The use of systematic reviews to justify phase III ophthalmology trials: an analysis. Eye (Lond). 2020.

158. Clarke M, Hopewell S. Many reports of randomised trials still don't begin or end with a systematic review of the relevant evidence. J Bahrain Med Soc. 2013;24(3):145-8.
159. Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, et al. A prediction rule to identify low-risk patients with pulmonary embolism. Arch Intern Med. 2006;166(2):169-75.
160. Panzer RJ, Feibel JH, Barker WH, Griner PF. Predicting the likelihood of hemorrhage in patients with stroke. Arch Intern Med. 1985;145(10):1800-3.
161. Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RB, Sr., et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. Lancet. 2009;373(9665):739-45.
162. Feringa HH, Bax JJ, Hoeks S, van Waning VH, Elhendy A, Karagiannis S, et al. A prognostic risk index for long-term mortality in patients with peripheral arterial disease. Arch Intern Med. 2007;167(22):2482-9.
163. Squizzato A, Donadini MP, Galli L, Dentali F, Aujesky D, Ageno W. Prognostic clinical prediction rules to identify a low-risk pulmonary embolism: a systematic review and meta-analysis. Journal of thrombosis and haemostasis : JTH. 2012;10(7):1276-90.
164. Elias A, Mallett S, Daoud-Elias M, Poggi JN, Clarke M. Prognostic models in acute pulmonary embolism: a systematic review and meta-analysis. BMJ Open. 2016;6(4):e010324.
165. Janjua M, Badshah A, Matta F, Danescu LG, Yaekoub AY, Stein PD. Treatment of acute pulmonary embolism as outpatients or following early discharge. A systematic review. Thromb Haemost. 2008;100(5):756-61.
166. Runchey S, McGee S. Does this patient have a hemorrhagic stroke?: clinical findings distinguishing hemorrhagic stroke from ischemic stroke. JAMA : the journal of the American Medical Association. 2010;303(22):2280-6.
167. Goldstein LB, Simel DL. Is this patient having a stroke? JAMA : the journal of the American Medical Association. 2005;293(19):2391-402.
168. Allen CM. Clinical diagnosis of the acute stroke syndrome. Q J Med. 1983;52(208):515-23.
169. Poorthuis MHF, Jones NR, Sherliker P, Clack R, de Borst GJ, Clarke R, et al. Utility of risk prediction models to detect atrial fibrillation in screened participants. European journal of preventive cardiology. 2021;28(6):586-95.
170. Himmelreich JCL, Veelers L, Lucassen WAM, Schnabel RB, Rienstra M, van Weert H, et al. Prediction models for atrial fibrillation applicable in the community: a systematic review and meta-analysis. Europace. 2020;22(5):684-94.
171. Sharma V, Ali I, van der Veer S, Martin G, Ainsworth J, Augustine T. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. BMJ Health Care Inform. 2021;28(1).
172. Hutson M. Artificial intelligence faces reproducibility crisis. Science. 2018;359(6377):725-6.
173. McDermott MBA, Wang S, Marinsek N, Ranganath R, Foschini L, Ghassemi M. Reproducibility in machine learning for health research: Still a ways to go. Sci Transl Med. 2021;13(586).
174. Haibe-Kains B, Adam GA, Hosny A, Khodakarami F, Massive Analysis Quality Control Society Board of D, Waldron L, et al. Transparency and reproducibility in artificial intelligence. Nature. 2020;586(7829):E14-E6.

175. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care. Journal of the American Medical Informatics Association : JAMIA. 2020;27(12):2011-5.
176. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, Spirit AI, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. BMJ. 2020;370:m3164.
177. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med. 2020;26(9):1320-4.
178. Sengupta PP, Shrestha S, Berthon B, Messas E, Donal E, Tison GH, et al. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. JACC Cardiovasc Imaging. 2020;13(9):2017-35.
179. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for Reporting Machine Learning Analyses in Clinical Research. Circ Cardiovasc Qual Outcomes. 2020;13(10):e006556.
180. Schwendicke F, Singh T, Lee JH, Gaudin R, Chaurasia A, Wiegand T, et al. Artificial intelligence in dental research: Checklist for authors, reviewers, readers. J Dent. 2021;107:103610.
181. Cook JA, Collins GS. The rise of big clinical databases. Br J Surg. 2015;102(2):e93-e101.
182. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. Ann Emerg Med. 1999;33(4):437-47.
183. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? Aust J Physiother. 2006;52(3):157-63.
184. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010;137(2):263-72.
185. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ, Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. Chest. 2010;138(5):1093-100.
186. Reilly BM, Evans AT, Schaider JJ, Das K, Calvin JE, Moran LA, et al. Impact of a clinical decision rule on hospital triage of patients with suspected acute cardiac ischemia in the emergency department. JAMA : the journal of the American Medical Association. 2002;288(3):342-50.
187. Goldman L, Cook EF, Johnson PA, Brand DA, Rouan GW, Lee TH. Prediction of the need for intensive care in patients who come to emergency departments with acute chest pain. The New England journal of medicine. 1996;334(23):1498-504.
188. Lund H, Juhl CB, Norgaard B, Draborg E, Henriksen M, Andreasen J, et al. Evidence-Based Research Series-Paper 2 : Using an Evidence-Based Research approach before a new study is conducted to ensure value. Journal of clinical epidemiology. 2021;129:158-66.
189. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. Statistics in medicine. 2000;19(24):3401-15.

190. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. Journal of clinical epidemiology. 2010;63(8):e1-37.
191. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605.
192. Riley RD, Ensor J, Snell KI, Debray TP, Altman DG, Moons KG, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. Bmj. 2016;353:i3140.
193. Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ. 2012;344:e3318.
194. Ioannidis JP. Scientific inbreeding and same-team replication: type D personality as an example. Journal of psychosomatic research. 2012;73(6):408-10.
195. 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.
196. 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.
197. Pencina MJ, D'Agostino RB, Sr., Larson MG, Massaro JM, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the framingham heart study. Circulation. 2009;119(24):3078-84.
198. van Kempen BJ, Ferket BS, Kavousi M, Leening MJ, Steyerberg EW, Ikram MA, et al. Performance of Framingham cardiovascular disease (CVD) predictions in the Rotterdam Study taking into account competing risks and disentangling CVD into coronary heart disease (CHD) and stroke. Int J Cardiol. 2014;171(3):413-8.
199. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. Annals of internal medicine. 2019;170(1):W1-W33.
200. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine. 2013;10(2):e1001381.
201. Cartus AR, Bodnar LM, Naimi AI. The Impact of Undersampling on the Predictive Performance of Logistic Regression and Machine Learning Algorithms: A Simulation Study. Epidemiology. 2020;31(5):e42-e4.
202. United Nations Statistics Division. Standard Country or Area Codes for Statistical Use (M49). Series M: Miscellaneous Statistical Papers, No 491998.
203. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. Journal of clinical epidemiology. 1995;48(12):1503-10.
204. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology. 1996;49(12):1373-9.

205. van Smeden M, de Groot JA, Moons KG, Collins GS, Altman DG, Eijkemans MJ, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol. 2016;16(1):163.
206. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE, Jr., Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. Statistics in medicine. 2019;38(7):1262-75.
207. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr., Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Statistics in medicine. 2019;38(7):1276-96.
208. van Smeden M, Moons KG, de Groot JA, Collins GS, Altman DG, Eijkemans MJ, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Stat Methods Med Res. 2019;28(8):2455-74.
209. Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. Br J Cancer. 1995;72(2):511-8.
210. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials. 1996;17(4):343-6.
211. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81(3):515-26.
212. National Cholesterol Education Program Expert Panel on Detection E, Treatment of High Blood Cholesterol in A. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. 2002;106(25):3143-421.
213. McNeil JJ, Peeters A, Liew D, Lim S, Vos T. A model for predicting the future incidence of coronary heart disease within percentiles of coronary heart disease risk. J Cardiovasc Risk. 2001;8(1):31-7.
214. Keys A, Aravanis C, Blackburn H, Van Buchem FS, Buzina R, Djordjevic BS, et al. Probability of middle-aged men developing coronary heart disease in five years. Circulation. 1972;45(4):815-28.
215. Polonsky TS, McClelland RL, Jorgensen NW, Bild DE, Burke GL, Guerci AD, et al. Coronary artery calcium score and risk classification for coronary heart disease prediction. JAMA : the journal of the American Medical Association. 2010;303(16):1610-6.
216. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743-53.
217. Expert Panel on Detection EToHBCi, Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA : the journal of the American Medical Association. 2001;285(19):2486-97.
218. Wilson PW, Castelli WP, Kannel WB. Coronary risk prediction in adults (the Framingham Heart Study). Am J Cardiol. 1987;59(14):91G-4G.
219. Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. BMJ. 2009;339:b2584.
220. Collins GS, Altman DG. An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. BMJ. 2010;340:c2442.

221. Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. BMJ. 2011;342:d3651.
222. Lee A, Mu JL, Joynt GM, Chiu CH, Lai VKW, Gin T, et al. Risk prediction models for delirium in the intensive care unit after cardiac surgery: a systematic review and independent external validation. Br J Anaesth. 2017;118(3):391-9.
223. Ramspek CL, Voskamp PW, van Ittersum FJ, Krediet RT, Dekker FW, van Diepen M. Prediction models for the mortality risk in chronic dialysis patients: a systematic review and independent external validation study. Clin Epidemiol. 2017;9:451-64.
224. Hendriksen JM, Geersing GJ, Lucassen WA, Erkens PM, Stoffers HE, van Weert HC, et al. Diagnostic prediction models for suspected pulmonary embolism: systematic review and independent external validation in primary care. BMJ. 2015;351:h4438.
225. Meertens LJE, van Montfort P, Scheepers HCJ, van Kuijk SMJ, Aardenburg R, Langenveld J, et al. Prediction models for the risk of spontaneous preterm birth based on maternal characteristics: a systematic review and independent external validation. Acta Obstet Gynecol Scand. 2018.
226. He Y, Ong Y, Li X, Din FV, Brown E, Timofeeva M, et al. Performance of prediction models on survival outcomes of colorectal cancer with surgical resection: A systematic review and meta-analysis. Surg Oncol. 2019;29:196-202.
227. 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.
228. Firth D. Bias Reduction of Maximum Likelihood Estimates. Biometrika. 1993;80(1):27-38.
229. Lin IF, Chang WP, Liao Y-N. Shrinkage methods enhanced the accuracy of parameter estimation using Cox models with small number of events. Journal of clinical epidemiology. 2013;66(7):743-51.
230. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Annals of internal medicine. 2019;170(1):51-8.
231. Wolff R, Collins GS, Kleijnen J, Mallett S, Reitsma JB, Riley R, et al. PROBAST: a risk of bias tool for prediction modelling studies. Challenges to evidence-based health care and Cochrane, The 24th Cochrane Colloquium; 2016 23-27 Oc; Seoul, Korea: John Wiley & Sons; 2016.
232. Wolff R, Collins G, Kleijnen J, Mallett S, Reitsma J, Riley R, et al. PROBAST: a risk of bias tool for prediction modelling studies. In: Challenges to evidence-based health care and Cochrane. The 24th Cochrane Colloquium; Seoul, Korea: John Wiley & Sons; 2016.
233. Ban J-W, Ignacio Emparanza J, Urreta I, Burls A. Design Characteristics Influence Performance of Clinical Prediction Rules in Validation: A Meta-Epidemiological Study. Plos One. 2016;11(1).
234. Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. Journal of clinical epidemiology. 2016;69:245-7.
235. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol

to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014;129(25 Suppl 2):S1-45.
236. Damen JA, Pajouheshnia R, Heus P, Moons KGM, Reitsma JB, Scholten RJPM, 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 Medicine. 2019;17(1).
237. 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.
238. Barzi F, Patel A, Gu D, Sritara P, Lam TH, Rodgers A, et al. Cardiovascular risk prediction tools for populations in Asia. J Epidemiol Community Health. 2007;61(2):115-21.
239. Grundy SM. National Cholesterol Education Program. Second Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). Circulation. 1994;89(3):1333.
240. Roccella EJ. The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNC V). Archives of internal medicine. 1993;153(2):154.
241. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J. 1991;121(1 Pt 2):293-8.
242. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. Circulation. 1991;83(1):356-62.
243. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA : the journal of the American Medical Association. 2001;285(22):2864-70.
244. Johnston SC, Rothwell PM, Nguyen-Huynh MN, Giles MF, Elkins JS, Bernstein AL, et al. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. Lancet. 2007;369(9558):283-92.
245. Wells PS, Hirsh J, Anderson DR, Lensing AW, Foster G, Kearon C, et al. A simple clinical model for the diagnosis of deep-vein thrombosis combined with impedance plethysmography: potential for an improvement in the diagnostic process. J Intern Med. 1998;243(1):15-23.
246. Wells PS, Anderson DR, Rodger M, Ginsberg JS, Kearon C, Gent M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. Thromb Haemost. 2000;83(3):416-20.
247. Kannel W, Wolf P, Garrison R. Some risk factors related to the annual incidence of cardiovascular disease and death in pooled repeated biennial measurements: Framingham Heart Study, 30 year follow-up. The Framingham Study: An Epidemiological Investigation of Cardiovascular Disease. 1987.
248. 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.
249. Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, 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. Eur Heart J. 2013;34(38):2949-3003.
250. 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.
251. 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.
252. 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.
253. 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.
254. 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.
255. 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.
256. 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.
257. 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.
258. 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.
259. 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.
260. 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? Arch Intern Med. 2008;168(11):1209-16.
261. 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.
262. 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.
263. van der Heijden AA, Ortegon 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.
264. 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.
265. Empana JP, Tafflet M, Escolano S, Vergnaux 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.
266. Ducloux D, Kazory A, Chalopin JM. Predicting coronary heart disease in renal transplant recipients: a prospective study. Kidney Int. 2004;66(1):441-7.
267. 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.
268. 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.
269. 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.
270. 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.
271. 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.
272. Protopsaltis 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.
273. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions. Second ed. Hoboken, NJ: Wiley-Blackwell; 2020.
274. 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.
275. 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.
276. 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.
277. 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.
278. 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.
279. Buitrago F, Calvo-Hueros JI, 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.
280. 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.
281. 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.
282. 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.
283. Ban J-W, 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. Journal of clinical epidemiology. 2021.
284. Whittle R, Peat G, Belcher J, Collins GS, Riley RD. Measurement error and timing of predictor values for multivariable risk prediction models are poorly reported. Journal of clinical epidemiology. 2018;102:38-49.
285. 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.
286. Scopus: Elsevier; [Web site]. Available from: https://www.scopus.com.
287. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012;345:e5661.
288. Cochrane Effective Practice and Organisation of Care (EPOC). What outcomes should be reported in EPOC reviews? EPOC Resources for review authors 2017 [Available from: http://epoc.cochrane.org/resources/epoc-resources-review-authors.
289. Brehaut JC, Stiell IG, Graham ID. Will a new clinical decision rule be widely used? The case of the Canadian C-spine rule. Acad Emerg Med. 2006;13(4):413-20.

290. Hendriksen JM, Geersing GJ, Moons KG, de Groot JA. Diagnostic and prognostic prediction models. Journal of thrombosis and haemostasis : JTH. 2013;11 Suppl 1:129-41.
291. Zomer E, Osborn D, Nazareth I, Blackburn R, Burton A, Hardoon S, et al. Effectiveness and cost-effectiveness of a cardiovascular risk prediction algorithm for people with severe mental illness (PRIMROSE). BMJ Open. 2017;7(9):e018181.
292. 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. Journal of clinical epidemiology. 2021;133:111-20.
293. Hemming K, Ryan R, Gill P, Westerby P, Jolly K, Marshall T. Targeted case finding in the prevention of cardiovascular disease: a stepped wedge cluster randomised controlled trial. Br J Gen Pract. 2016;66(651):e758-67.
294. Cooper A, O'Flynn N, Guideline Development G. Risk assessment and lipid modification for primary and secondary prevention of cardiovascular disease: summary of NICE guidance. BMJ. 2008;336(7655):1246-8.
295. Sheibani R, Sheibani M, Heidari-Bakavoli A, Abu-Hanna A, Eslami S. The Effect of a Clinical Decision Support System on Improving Adherence to Guideline in the Treatment of Atrial Fibrillation: An Interrupted Time Series Study. J Med Syst. 2017;42(2):26.
296. Bessen T, Clark R, Shakib S, Hughes G. A multifaceted strategy for implementation of the Ottawa ankle rules in two emergency departments. BMJ. 2009;339:b3056.
297. Cochrane Effective Practice and Organisation of Care (EPOC). What study designs should be included in an EPOC review? EPOC Resources for review authors 2017 [Available from: http://epoc.cochrane.org/epoc-specific-resources-review-authors.
298. Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. Fam Pract. 2000;17 Suppl 1:S11-6.
299. Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. BMJ. 2015;350:h2750.
300. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet. 2002;359(9300):57-61.
301. Fayers PM, Jordhoy MS, Kaasa S. Cluster-randomized trials. Palliat Med. 2002;16(1):69-70.
302. Zhan Z, van den Heuvel ER, Doornbos PM, Burger H, Verberne CJ, Wiggers T, et al. Strengths and weaknesses of a stepped wedge cluster randomized design: its application in a colorectal cancer follow-up study. Journal of clinical epidemiology. 2014;67(4):454-61.
303. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med. 2007;33(2):155-61.
304. Zabor EC, Kaizer AM, Hobbs BP. Randomized Controlled Trials. Chest. 2020;158(1S):S79-S87.
305. Pocock SJ. Clinical trials : a practical approach. Chichester, England: John Wiley & Sons; 1983.

306. Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet. 2002;359(9301):145-9.
307. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002;359(9302):248-52.
308. Anaby D, Lal S, Huszczynski J, Maich J, Rogers J, Law M. Interrupted time series design: a useful approach for studying interventions targeting participation. Phys Occup Ther Pediatr. 2014;34(4):457-70.
309. Sullivan LM. Repeated measures. Circulation. 2008;117(9):1238-43.
310. Kraska M. Repeated Measures Design. In: Salkind NJ, editor. Encyclopedia of Research Design. Thousand Oaks: SAGE Publications, Inc.; 2010. p. 1243-7.
311. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27):iii-x, 1-173.
312. Sedgwick P. Before and after study designs. BMJ. 2014;349:g5074.
313. Ho AMH, Phelan R, Mizubuti GB, Murdoch JAC, Wickett S, Ho AK, et al. Bias in Before-After Studies: Narrative Overview for Anesthesiologists. Anesth Analg. 2018;126(5):1755-62.
314. Goodacre S. Uncontrolled before-after studies: discouraged by Cochrane and the EMJ. Emerg Med J. 2015;32(7):507-8.
315. Lipsey MW, Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. Am Psychol. 1993;48(12):1181-209.
316. Wilson DB, Lipsey MW. The role of method in treatment effectiveness research: evidence from meta-analysis. Psychological methods. 2001;6(4):413.
317. Shikata S, Nakayama T, Noguchi Y, Taji Y, Yamagishi H. Comparison of effects in randomized controlled trials with observational studies in digestive surgery. Ann Surg. 2006;244(5):668-76.
318. Higgins JPT, Savović J, Page MJ, Sterne JAC. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). 2018. In: Cochrane Methods [Internet]. Cochrane Database of Systematic Reviews. Available from: https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/current-version-of-rob-2.
319. Eldridge S, Campbell M, Campbell M, Dahota A, Giraudeau B, Higgins J, et al. Revised Cochrane risk of bias tool for randomized trials (RoB 2.0): Additional considerations for cluster-randomized trials. 2016. In: Cochrane Methods [Internet]. Cochrane Database of Systematic Reviews; [10 (Suppl 1)]. Available from: https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/archive-rob-2-0-cluster-randomized-trials-2016.
320. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
321. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
322. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
323. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, Group C. Extending the CONSORT statement to randomized trials of nonpharmacologic

treatment: explanation and elaboration. Annals of internal medicine. 2008;148(4):295-309.
324. Des Jarlais DC, Lyles C, Crepaz N, Group T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004;94(3):361-6.
325. Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. Lancet. 2010;376(9734):20-1.
326. Hopewell S, Hirst A, Collins GS, Mallett S, Yu LM, Altman DG. Reporting of participant flow diagrams in published reports of randomized trials. Trials. 2011;12:253.
327. Egger M, Juni P, Bartlett C, Group C. Value of flow diagrams in reports of randomized controlled trials. JAMA : the journal of the American Medical Association. 2001;285(15):1996-9.
328. Holman B, Elliott KC. The promise and perils of industry-funded science. Philosophy Compass. 2018;13(11):n/a-n/a.
329. Chan AW, Pello A, Kitchen J, Axentiev A, Virtanen JI, Liu A, et al. Association of Trial Registration With Reporting of Primary Outcomes in Protocols and Publications. JAMA : the journal of the American Medical Association. 2017;318(17):1709-11.
330. Gibson NS, Sohne M, Kruip MJ, Tick LW, Gerdes VE, Bossuyt PM, et al. Further validation and simplification of the Wells clinical decision rule in pulmonary embolism. Thromb Haemost. 2008;99(1):229-34.
331. World Health Organization. Prevention of cardiovascular disease: pocket guidelines for assessment and management of cardiovascular risk:(WH 2007 [Guideline]. Available from: https://www.who.int/cardiovascular_diseases/publications/Prevention_of_Cardiovascular_Disease/en/.
332. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. Circulation. 2002;105(3):310-5.
333. Jackson R, Barham P, Bills J, Birch T, McLennan L, MacMahon S, et al. Management of raised blood pressure in New Zealand: a discussion document. BMJ. 1993;307(6896):107-10.
334. Thomsen TF, Davidsen M, Ibsen H, Jorgensen T, Jensen G, Borch-Johnsen K. A new method for CHD prediction and prevention based on regional risk scores and randomized clinical trials; PRECARD and the Copenhagen Risk Score. J Cardiovasc Risk. 2001;8(5):291-7.
335. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. Diabetes Metab Res Rev. 2000;16(3):164-71.
336. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes care. 2003;26(3):725-31.
337. Kattah JC, Talkad AV, Wang DZ, Hsieh YH, Newman-Toker DE. HINTS to diagnose stroke in the acute vestibular syndrome: three-step bedside oculomotor examination more sensitive than early MRI diffusion-weighted imaging. Stroke. 2009;40(11):3504-10.

338. Apostolakis S, Sullivan RM, Olshansky B, Lip GYH. Factors affecting quality of anticoagulation control among patients with atrial fibrillation on warfarin: the SAMe-TT(2)R(2) score. Chest. 2013;144(5):1555-63.
339. van Belle A, Buller HR, Huisman MV, Huisman PM, Kaasjager K, Kamphuisen PW, et al. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. JAMA : the journal of the American Medical Association. 2006;295(2):172-9.
340. Kline JA, Courtney DM, Than MP, Hogg K, Miller CD, Johnson CL, et al. Accuracy of very low pretest probability estimates for pulmonary embolism using the method of attribute matching compared with the Wells score. Acad Emerg Med. 2010;17(2):133-41.
341. Wells PS, Anderson DR, Rodger M, Forgie M, Kearon C, Dreyer J, et al. Evaluation of D-dimer in the diagnosis of suspected deep-vein thrombosis. The New England journal of medicine. 2003;349(13):1227-35.
342. Constans J, Salmi LR, Sevestre-Pietri MA, Perusat S, Nguon M, Degeilh M, et al. A clinical prediction score for upper extremity deep venous thrombosis. Thromb Haemost. 2008;99(1):202-7.
343. Kline JA, Johnson CL, Pollack CV, Jr., Diercks DB, Hollander JE, Newgard CD, et al. Pretest probability assessment derived from attribute matching. BMC medical informatics and decision making. 2005;5:26.
344. Bösner S, Haasenritter J, Becker A, Karatolios K, Vaucher P, Gencer B, et al. Ruling out coronary artery disease in primary care: development and validation of a simple prediction rule. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2010;182(12):1295-300.
345. Antman EM, Cohen M, Bernink PJ, McCabe CH, Horacek T, Papuchis G, et al. The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. JAMA : the journal of the American Medical Association. 2000;284(7):835-42.
346. Verbeek PR, Vermeulen MJ, Ali FH, Messenger DW, Summers J, Morrison LJ. Derivation of a termination-of-resuscitation guideline for emergency medical technicians using automated external defibrillators. Acad Emerg Med. 2002;9(7):671-8.
347. Ayres K, Conner M, Prestwich A, Hurling R, Cobain M, Lawton R, et al. Exploring the question-behaviour effect: randomized controlled trial of motivational and question-behaviour interventions. Br J Health Psychol. 2013;18(1):31-44.
348. Hess EP, Hollander JE, Schaffer JT, Kline JA, Torres CA, Diercks DB, et al. Shared decision making in patients with low risk chest pain: prospective randomized pragmatic trial. BMJ. 2016;355:i6165.
349. Jørgensen T, Jacobsen RK, Toft U, Aadahl M, Glumer C, Pisinger C. Effect of screening and lifestyle counselling on incidence of ischaemic heart disease in general population: Inter99 randomised trial. BMJ. 2014;348:g3617.
350. Linkins LA, Bates SM, Lang E, Kahn SR, Douketis JD, Julian J, et al. Selective D-dimer testing for diagnosis of a first suspected episode of deep venous thrombosis: a randomized trial. Annals of internal medicine. 2013;158(2):93-100.
351. Than M, Aldous S, Lord SJ, Goodacre S, Frampton CM, Troughton R, et al. A 2-hour diagnostic protocol for possible cardiac chest pain in the

emergency department: a randomized clinical trial. JAMA internal medicine. 2014;174(1):51-8.
352. Tsuyuki RT, Al Hamarneh YN, Jones CA, Hemmelgarn BR. The Effectiveness of Pharmacist Interventions on Cardiovascular Risk: The Multicenter Randomized Controlled RxEACH Trial. J Am Coll Cardiol. 2016;67(24):2846-54.
353. Patel MS, Kurtzman GW, Kannan S, Small DS, Morris A, Honeywell S, Jr., et al. Effect of an Automated Patient Dashboard Using Active Choice and Peer Comparison Performance Feedback to Physicians on Statin Prescribing: The PRESCRIBE Cluster Randomized Clinical Trial. JAMA Netw Open. 2018;1(3):e180818.
354. Holbrook AM, Pullenayegum EM, Troyan SM, Nikitovic M, Crowther MA. Personalized benefit-harm information influences patient decisions regarding warfarin2013. e406-e15 p.
355. Jansen J, Bonner C, McKinn S, Irwig L, Glasziou P, Doust J, et al. General practitioners' use of absolute risk versus individual risk factors in cardiovascular disease prevention: an experimental study. BMJ Open. 2014;4(5):e004812.
356. Scheitel MR, Kessler ME, Shellum JL, Peters SG, Milliner DS, Liu H, et al. Effect of a Novel Clinical Decision Support Tool on the Efficiency and Accuracy of Treatment Recommendations for Cholesterol Management. Appl Clin Inform. 2017;8(1):124-36.
357. Jegan NRA, Kurwitz SA, Kramer LK, Heinzel-Gutenbrunner M, Adarkwah CC, Popert U, et al. The effect of a new lifetime-cardiovascular-risk display on patients' motivation to participate in shared decision-making. BMC Fam Pract. 2018;19(1):84.
358. Dhiman P, Lee H, Kirtley S, Collins GS. A systematic review showed more consideration is needed when conducting nonrandomized studies of interventions. Journal of clinical epidemiology. 2020;117:99-108.
359. Wilson B, Burnett P, Moher D, Altman D, Salman RAS. Completeness of reporting of randomised controlled trials including people with transient ischaemic attack or stroke: A systematic review. European Stroke Journal. 2018;3.
360. Alamri HM, Alharbi F. Quality Assessment of Randomized Clinical Trials Reporting in Endodontic Journals: An Observational Study from 2012 to 2017. J Endod. 2018;44(8):1246-50.
361. Ban J-W, Perera R, Stevens R. Familiarity and use of cardiovascular clinical prediction rules: survey of general practitioners. 2020.
362. Sheridan SL, Crespo E. Does the routine use of global coronary heart disease risk scores translate into clinical benefits or harms? A systematic review of the literature. BMC Health Serv Res. 2008;8:60.
363. Sheridan SL, Viera AJ, Krantz MJ, Ice CL, Steinman LE, Peters KE, et al. The effect of giving global coronary risk information to adults: a systematic review. Arch Intern Med. 2010;170(3):230-9.
364. Usher-Smith JA, Silarova B, Schuit E, Moons KG, Griffin SJ. Impact of provision of cardiovascular disease risk estimates to healthcare professionals and patients: a systematic review. BMJ Open. 2015;5(10):e008717.
365. Karmali KN, Persell SD, Perel P, Lloyd-Jones DM, Berendsen MA, Huffman MD. Risk scoring for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev. 2017;3:CD006887.

366. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev. 2012;11:MR000030.
367. Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. BMJ. 2014;348:g3804.
368. Sarkis-Onofre R, Poletto-Neto V, Cenci MS, Moher D, Pereira-Cenci T. CONSORT endorsement improves the quality of reports of randomized clinical trials in dentistry. Journal of clinical epidemiology.
369. Arra I, Velker V, Sexton T, Rotenberg BW, Boldt RG, Rodrigues G. A CONSORT Clinical Trial Reporting Compliance Audit of the Oncology Randomized Controlled Trial Literature. Cureus. 2013.
370. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. BMJ. 2017;357:j2490.
371. Silman AJ, Macfarlane GJ. Epidemiological studies : a practical guide. 2nd ed. Cambridge: Cambridge University Press; 2002.
372. Gitlin LN. Introducing a new intervention: an overview of research phases and common challenges. Am J Occup Ther. 2013;67(2):177-84.
373. Karlsson LO, Nilsson S, Bang M, Nilsson L, Charitakis E, Janzon M. A clinical decision support tool for improving adherence to guidelines on anticoagulant therapy in patients with atrial fibrillation at risk of stroke: A cluster-randomized trial in a Swedish primary care setting (the CDS-AF study). PLoS medicine. 2018;15(3):e1002528.
374. U.S. Food and Drug Administration. Development & Approval Process | Drugs 2019 [updated 10/28/2019. Available from: https://www.fda.gov/drugs/development-approval-process-drugs.
375. McCartney M. We need better evidence on non-drug interventions for covid-19. BMJ. 2020;370:m3473.
376. Steel N, Ford JA, Newton JN, Davis ACJ, Vos T, Naghavi M, et al. Changes in health in the countries of the UK and 150 English Local Authority areas 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2018;392(10158):1647-61.
377. Goodacre S, Sampson F, Stevenson M, Wailoo A, Sutton A, Thomas S, et al. Measurement of the clinical and cost-effectiveness of non-invasive diagnostic testing strategies for deep vein thrombosis. Health Technol Assess. 2006;10(15):1-168, iii-iv.
378. Thomsen TF, Jorgensen T, Ibsen H, Borch-Johnsen K. Assessment of coronary risk in general practice in relation to the use of guidelines: a survey in Denmark. Preventive medicine. 2001;33(4):300-4.
379. Hobbs FD, Erhardt L. Acceptance of guideline recommendations and perceived implementation of coronary heart disease prevention among primary care physicians in five European countries: the Reassessing European Attitudes about Cardiovascular Treatment (REACT) survey. Fam Pract. 2002;19(6):596-604.
380. De Muylder R, Lorant V, Paulus D, Nackers F, Jeanjean M, Boland B. Obstacles to cardiovascular prevention in general practice. Acta Cardiol. 2004;59(2):119-25.

381. Bonnevie L, Thomsen T, Jørgensen T. The Use of Computerized Decision Support Systems in Preventive Cardiology—Principal Results from the National Precard® Survey in Denmark. European Journal of Cardiovascular Prevention & Rehabilitation. 2005;12(1):52-5.
382. Eichler K, Zoller M, Tschudi P, Steurer J. Barriers to apply cardiovascular prediction rules in primary care: a postal survey. BMC Fam Pract. 2007;8:1.
383. Müller-Riemenschneider F, Holmberg C, Rieckmann N, Kliems H, Rufer V, Müller-Nordhorn J, et al. Barriers to routine risk-score use for healthy primary care patients: survey and qualitative study. Arch Intern Med. 2010;170(8):719-24.
384. Dallongeville J, Banegas JR, Tubach F, Guallar E, Borghi C, De Backer G, et al. Survey of physicians' practices in the control of cardiovascular risk factors: the EURIKA study. European journal of preventive cardiology. 2012;19(3):541-50.
385. Elustondo SG, Aguado PN, de La Rasilla Cooper CG, Manzanet JP, Sendin DS. Cardiovascular risk tables: opinion and degree of use of primary care doctors from Madrid, Spain. J Eval Clin Pract. 2013;19(1):148-52.
386. Sarazin M, Chiappe SG, Kasprzyk M, Mismetti P, Lasserre A. A survey of French general practitioners and a qualitative study on their use and assessment of predictive clinical scores. International journal of general medicine. 2013;6:419-26.
387. Byrne D, O'Connor L, Jennings S, Bennett K, Murphy AW. A Survey of GPs Awareness and Use of Risk Assessment Tools and Cardiovascular Disease Prevention Guidelines. Ir Med J. 2015;108(7):204-7.
388. Heidbuchel H, Dagres N, Antz M, Kuck KH, Lazure P, Murray S, et al. Major knowledge gaps and system barriers to guideline implementation among European physicians treating patients with atrial fibrillation: a European Society of Cardiology international educational needs assessment. Europace. 2018;20(12):1919-28.
389. Pearson SD, Goldman L, Garcia TB, Cook EF, Lee TH. Physician response to a prediction rule for the triage of emergency department patients with chest pain. Journal of general internal medicine. 1994;9(5):241-7.
390. Eaton CB, Galliher JM, McBride PE, Bonham AJ, Kappus JA, Hickner J. Family physician's knowledge, beliefs, and self-reported practice patterns regarding hyperlipidemia: a National Research Network (NRN) survey. J Am Board Fam Med. 2006;19(1):46-53.
391. Shillinglaw B, Viera AJ, Edwards T, Simpson R, Sheridan SL. Use of global coronary heart disease risk assessment in practice: a cross-sectional survey of a sample of U.S. physicians. BMC Health Serv Res. 2012;12:20.
392. Pokharel Y, Steinberg L, Chan W, Akeroyd JM, Jones PH, Nambi V, et al. Case-based educational intervention to assess change in providers' knowledge and attitudes towards the 2013 American College of Cardiology/American Heart Association Cholesterol Management Guideline. Atherosclerosis. 2016;246:115-20.
393. Wu X, Malhotra A, Forman HP, Nunez D, Sanelli P. The Use of High-Risk Criteria in Screening Patients for Blunt Cerebrovascular Injury: A Survey. Acad Radiol. 2017;24(4):456-61.
394. Amroze A, Mazor K, Crawford S, O'Day K, McManus DD, Kapoor A. Survey of confidence in use of stroke and bleeding risk calculators, knowledge

of anticoagulants, and comfort with prescription of anticoagulation in challenging scenarios: SUPPORT-AF II study. J Thromb Thrombolysis. 2019;48(4):629-37.
395. Imms A, Quinn S, Nelson M. General practitioners' use of cardiovascular risk calculators. Aust Fam Physician. 2010;39(1-2):57-60.
396. Ferguson C, Inglis SC, Newton PJ, Middleton S, Macdonald PS, Davidson PM. Education and practice gaps on atrial fibrillation and anticoagulation: a survey of cardiovascular nurses. BMC Med Educ. 2016;16:9.
397. Ferguson C, Hickman LD, Phillips J, Newton PJ, Inglis SC, Lam L, et al. An mHealth intervention to improve nurses' atrial fibrillation and anticoagulation knowledge and practice: the EVICOAG study. Eur J Cardiovasc Nurs. 2019;18(1):7-15.
398. Kesieme EB, Arekhandia BJ, Inuwa IM, Akpayak IC, Ekpe EE, Olawoye OA, et al. Knowledge and practice of prophylaxis of deep venous thrombosis: A survey among Nigerian surgeons. Niger J Clin Pract. 2016;19(2):170-4.
399. Sposito AC, Ramires JAF, Jukema JW, Molina JC, Silva PMD, Ghadanfar MM, et al. Physicians attitudes and adherence to use of risk scores for primary prevention of cardiovascular disease: cross- sectional survey in three world regions. Current Medical Research and Opinion, 2009, Vol25(5), p1171-1178. 2009;25(5):1171-8.
400. Dykema J, Jones NR, Piche T, Stevenson J. Surveying clinicians by web: current issues in design and administration. Eval Health Prof. 2013;36(3):352-81.
401. Couper MP. Web surveys - A review of issues and approaches. Public Opin Q. 2000;64(4):464-94.
402. Doctors.net.uk 2017 [cited 2019 October 18]. Available from: https://www.doctors.net.uk.
403. van Steenkiste B, Weijden TVD, Stoffers HEJH, Grol R. Barriers to implementing cardiovascular risk tables in routine general practice. Scandinavian Journal of Primary Health Care, 2004, Vol22(1), p32-37. 2004;22(1):32-7.
404. Brown B, Cheraghi-Sohi S, Jaki T, Su TL, Buchan I, Sperrin M. Understanding clinical prediction models as 'innovations': a mixed methods study in UK family practice. BMC medical informatics and decision making. 2016;16:106.
405. British Cardiac Society, British Hypertension Society, Diabetes U. K., Heart U. K., Primary Care Cardiovascular Society, Stroke Association. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart. 2005;91 Suppl 5:v1-52.
406. Deanfield J, Sattar N, Simpson I, Wood D, Bradbury K, Fox K, et al. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). Heart. 2014;100(2):ii1-ii67.
407. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357.
408. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. European Heart Journal. 2016;37(29):2315-81.
409. Jackson R. Updated New Zealand cardiovascular disease risk- benefit prediction guide. BMJ. 2000;320(7236):709.

410. Pylypchuk R, Wells S, Kerr A, Poppe K, Riddell T, Harwood M, et al. Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. The Lancet. 2018;391(10133):1897-907.
411. Ministry of Health. Cardiovascular Disease Risk Assessment and Management for Primary Care 2018 [cited 2019 October 18]. Available from: https://www.health.govt.nz/publication/cardiovascular-disease-risk-assessment-and-management-primary-care.
412. World Health Organization and International Society of Hypertension. WHO/ISH cardiovascular risk prediction charts 2007 [cited 2019 October 18]. Available from: http://www.who.int/cardiovascular_diseases/guidelines/Chart_predictions/en/.
413. World Health Organization. Prevention of cardiovascular disease : guidelines for assessment and management of total cardiovascular risk: World Health Organization; 2007 [cited 2019 October 18]. Available from: https://apps.who.int/iris/handle/10665/43685.
414. Merwick Á, Albers GW, Amarenco P, Arsava EM, Ay H, Calvet D, et al. Addition of brain and carotid imaging to the ABCD2 score to identify patients at early risk of stroke after transient ischaemic attack: a multicentre observational study. The Lancet Neurology. 2010;9(11):1060-9.
415. National Institute for Health and Care Excellence. Stroke and transient ischaemic attack in over 16s: diagnosis and initial management 2019 [cited 2019 October 18]. Available from: https://www.nice.org.uk/guidance/ng128.
416. Johnston SC, Gress DR, Browner WS, Sidney S. Short- term Prognosis After Emergency Department Diagnosis of TIA. JAMA : the journal of the American Medical Association. 2000;284(22):2901-6.
417. National Institute for Health and Care Excellence. Atrial Fibrillation: management 2014 [updated August 2014; cited 2019 October 18]. Available from: https://www.nice.org.uk/guidance/cg180.
418. Scottish Intercollegiate Guidelines Network. Antithrombotics: indications and management 2012 [updated June 2013; cited 2019 October 18]. Available from: https://www.sign.ac.uk/sign-129-antithrombotics-indications-and-management.html.
419. Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J. 2016;37(38):2893-962.
420. National Institute for Health and Care Excellence. Venous thromboembolic diseases: diagnosis, management and thrombophilia testing 2012 [updated November 2015; cited 2019 October 18]. Available from: https://www.nice.org.uk/guidance/cg144.
421. Scottish Intercollegiate Guidelines Network. Prevention and management of venous thromboembolism 2010 [updated October 2014; cited 2019 October 18]. Available from: https://www.sign.ac.uk/sign-122-prevention-and-management-of-venous-thromboembolism.html.
422. Le Gal G, Righini M, Roy PM, Sanchez O, Aujesky D, Bounameaux H, et al. Prediction of pulmonary embolism in the emergency department: the revised Geneva score. Annals of internal medicine. 2006;144(3):165-71.
423. Konstantinides SV, Torbicki A, Agnelli G, Danchin N, Fitzmaurice D, Galie N, et al. 2014 ESC guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J. 2014;35(43):3033-69, 69a-69k.

424. Williams B, Mancia G, Spiering W, Rosei EA, Azizi M, Burnier M, et al. 2018 ESC/ ESH Guidelines for the management of arterial hypertension. European Heart Journal. 2018;39(33):3021-104.
425. Catapano AL, Graham I, De Backer G, Wiklund O, Chapman MJ, Drexel H, et al. 2016 ESC/ EAS Guidelines for the Management of Dyslipidaemias. European Heart Journal. 2016;37(39):2999-3058.
426. Altman DG. Practical statistics for medical research. Boca Raton, Fl.: Chapman & Hall/CRC; 1999.
427. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310(6973):170.
428. Perneger TV. What's wrong with Bonferroni adjustments. BMJ. 1998;316(7139):1236-8.
429. Morgan JF. p Value fetishism and use of the Bonferroni adjustment. Evid Based Ment Health. 2007;10(2):34-5.
430. Kothari R, Hall K, Brott T, Broderick J. Early stroke recognition: developing an out-of-hospital NIH Stroke Scale. Acad Emerg Med. 1997;4(10):986-90.
431. Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke. 1989;20(7):864-70.
432. Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. BMJ. 2013;346:f2573.
433. Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E. FRAX and the assessment of fracture probability in men and women from the UK. Osteoporos Int. 2008;19(4):385-97.
434. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Medical decision making : an international journal of the Society for Medical Decision Making. 1981;1(3):239-46.
435. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. BMJ Open. 2015;5(3):e007825.
436. National Institute for Health and Care Excellence. Lipid modification: Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease 2008 [Available from: https://www.nice.org.uk/guidance/cg67.
437. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. BMJ. 2012;344:e4181.
438. National Institute for Health and Care Excellence. Fever in under 5s: assessment and initial management 2019 [cited 2020 22 November]. Available from: https://www.nice.org.uk/guidance/ng143/chapter/Recommendations#clinical-assessment-of-children-with-fever.
439. Jones TL, Baxter MA, Khanduja V. A quick guide to survey research. Ann R Coll Surg Engl. 2013;95(1):5-7.
440. Albu JB, Sohler N, Li R, Li X, Young E, Gregg EW, et al. An Interrupted Time Series Analysis to Determine the Effect of an Electronic Health Record-

Based Intervention on Appropriate Screening for Type 2 Diabetes in Urban Primary Care Clinics in New York City. Diabetes care. 2017;40(8):1058-64.
441. Lay-Flurrie SL, Sheppard JP, Stevens RJ, Mallen C, Heneghan C, Hobbs FDR, et al. Impact of Changes to National Hypertension Guidelines on Hypertension Management and Outcomes in the United Kingdom. Hypertension. 2020;75(2):356-64.
442. Runyon MS, Richman PB, Kline JA. Emergency Medicine Practitioner Knowledge and Use of Decision Rules for the Evaluation of Patients with Suspected Pulmonary Embolism: Variations by Practice Setting and Training Level. Academic Emergency Medicine. 2007;14(1):53-7.
443. Graham ID, Stiell IG, Laupacis A, O'Connor AM, Wells GA. Emergency physicians' attitudes toward and use of clinical decision rules for radiography. Acad Emerg Med. 1998;5(2):134-40.
444. Graham ID, Stiell IG, Laupacis A, McAuley L, Howell M, Clancy M, et al. Awareness and use of the Ottawa ankle and knee rules in 5 countries: can publication alone be enough to change practice? Ann Emerg Med. 2001;37(3):259-66.
445. Boutis K, Constantine E, Schuh S, Pecaric M, Stephens D, Narayanan UG. Pediatric Emergency Physician Opinions on Ankle Radiograph Clinical Decision Rules. Academic Emergency Medicine. 2010;17(7):709-17.
446. Dowling SK, Wishart I. Use of the Ottawa Ankle Rules in children: a survey of physicians' practice patterns. Cjem. 2011;13(5):333-8; E44-6.
447. Beutel BG, Trehan SK, Shalvoy RM, Mello MJ. The Ottawa knee rule: examining use in an academic emergency department. West J Emerg Med. 2012;13(4):366-72.
448. Haskins R, Osmotherly PG, Southgate E, Rivett DA. Physiotherapists' knowledge, attitudes and practices regarding clinical prediction rules for low back pain. Manual Therapy. 2014;19(2):142-51.
449. Eagles D, Stiell IG, Clement CM, Brehaut J, Taljaard M, Kelly AM, et al. International survey of emergency physicians' awareness and use of the Canadian Cervical-Spine Rule and the Canadian Computed Tomography Head Rule. Acad Emerg Med. 2008;15(12):1256-61.
450. Ballard DW, Rauchwerger AS, Reed ME, Vinson DR, Mark DG, Offerman SR, et al. Emergency Physicians' Knowledge and Attitudes of Clinical Decision Support in the Electronic Health Record: A Survey- based Study. Academic Emergency Medicine. 2013;20(4):352-60.
451. Edwards QT, Maradiegue A, Seibert D, Saunders-Goldson S, Humphreys S. Breast cancer risk elements and nurse practitioners' knowledge, use, and perceived comfort level of breast cancer risk assessment. J Am Acad Nurse Pract. 2009;21(5):270-7.
452. Kappen TH, van Loon K, Kappen MA, van Wolfswinkel L, Vergouwe Y, van Klei WA, et al. Barriers and facilitators perceived by physicians when using prediction models in practice. Journal of clinical epidemiology. 2016;70:136-45.
453. van der Steen JT, Albers G, Licht-Strunk E, Muller MT, Ribbe MW. A validated risk score to estimate mortality risk in patients with dementia and pneumonia: barriers to clinical impact. International psychogeriatrics / IPA. 2011;23(1):31-43.
454. The Institute of Translational Health Sciences. Primary Care Research Network (WPRN) 2007 [updated 2020; cited 2020 December 5]. Available from:

https://www.iths.org/investigators/find-collaborators/primary-care-research-network/.
455. Smeets M, Zervas S, Leben H, Vermandere M, Janssens S, Mullens W, et al. General practitioners' perceptions about their role in current and future heart failure care: an exploratory qualitative study. BMC Health Serv Res. 2019;19(1):432.
456. Bissett SM, Stone KM, Rapley T, Preshaw PM. An exploratory qualitative interview study about collaboration between medicine and dentistry in relation to diabetes management. BMJ Open. 2013;3(2).
457. Hampson N, Ruane S. The value of pharmacists in general practice: perspectives of general practitioners-an exploratory interview study. Int J Clin Pharm. 2019;41(2):496-503.
458. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2.
459. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Schünemann HJ. Rating Quality of Evidence and Strength of Recommendations: GRADE: What Is "Quality of Evidence" and Why Is It Important to Clinicians? BMJ: British Medical Journal. 2008;336(7651):995-8.
460. Gabbay J, le May A. Evidence based guidelines or collectively constructed "mindlines?" Ethnographic study of knowledge management in primary care. BMJ. 2004;329(7473):1013.
461. Green ML. Evidence-based medicine training in internal medicine residency programs a national survey. Journal of general internal medicine. 2000;15(2):129-33.
462. Kuhn GJ, Wyer PC, Cordell WH, Rowe BH, Society for Academic Emergency Medicine Evidence-based Medicine Interest G. A survey to determine the prevalence and characteristics of training in Evidence-Based Medicine in emergency medicine residency programs. J Emerg Med. 2005;28(3):353-9.
463. Bednarczyk J, Pauls M, Fridfinnson J, Weldon E. Characteristics of evidence-based medicine training in Royal College of Physicians and Surgeons of Canada emergency medicine residencies - a national survey of program directors. BMC Med Educ. 2014;14:57.
464. Aiyer M, Hemmer P, Meyer L, Albritton TA, Levine S, Reddy S. Evidence-based medicine in internal medicine clerkships: a national survey. South Med J. 2002;95(12):1389-95.
465. Blanco MA, Capello CF, Dorsch JL, Perry G, Zanetti ML. A survey study of evidence-based medicine training in US and Canadian medical schools. J Med Libr Assoc. 2014;102(3):160-8.
466. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, Jr., et al. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol. 2019.
467. 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.

468. University of Nottingham. QRISK – a new cardiovascular risk score to identify patients at high risk of cardiovascular disease for prevention 2014 [Available from: https://impact.ref.ac.uk/casestudies/CaseStudy.aspx?Id=27328#.
469. Lee TC, Shah NU, Haack A, Baxter SL. Clinical Implementation of Predictive Models Embedded within Electronic Health Record Systems: A Systematic Review. Informatics (MDPI). 2020;7(3).
470. Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. BMJ. 2004;329(7473):1017.
471. Maggio LA. Educating physicians in evidence based medicine: current practices and curricular strategies. Perspect Med Educ. 2016;5(6):358-61.
472. Kumaravel B, Hearn JH, Jahangiri L, Pollard R, Stocker CJ, Nunan D. A systematic review and taxonomy of tools for evaluating evidence-based medicine teaching in medical education. Syst Rev. 2020;9(1):91.
473. Albarqouni L, Hoffmann T, Glasziou P. Evidence-based practice educational intervention studies: a systematic review of what is taught and how it is measured. BMC Med Educ. 2018;18(1):177.
474. Glasziou P, Haynes B. The paths from research to improved health outcomes. Evid Based Nurs. 2005;8(2):36-8.
475. Hartling L, Hamm MP, Milne A, Vandermeer B, Santaguida PL, Ansari M, et al. Testing the Risk of Bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. Journal of clinical epidemiology. 2013;66(9):973-81.

Appendices

Chapter 2.

Appendix 2.1. Participant invitation letter for the survey of authors.



Department for Continuing Education
Rewley House
1 Wellington Square
Oxford
Oxfordshire
OX1 2JA
United Kingdom

Participant Information Sheet (Participant Invitation Letter)

PO Box 912
East Olympia, WA 98540
USA
(360) 970-3953 jong-wook.ban@conted.ox.ac.uk

Name of recipient

Department

Institution

Street address

State/Province, Postal code

Country

Saturday, June 20, 2015

Dear Author,

My name is Jong-Wook Ban, and I am a doctoral student at the University of Oxford. **I am writing to invite you to participate in a study:** "Assessment of needs for a new cardiovascular prediction rule in derivations: a review of derivation studies included in the International Register of Clinical Prediction Rules for Primary Care and a survey of authors". I am interested in understanding whether and how existing prediction rules are considered while developing new cardiovascular prediction rules.

You have been invited to take part in the survey because your paper titled *"Title of study"* published in the *"Name of the journal"* in *"Publication year"* met the inclusion criteria: (1) the study derived a new clinical prediction rule, and (2) the prediction rule is included in the cardiovascular domain of the International Register of Clinical Prediction Rules for Primary Care.

The survey questionnaire consists of only three questions, and it should take less than 2 minutes to complete. Simply return the completed survey using the pre-stamped envelope included in the packet. The completion of this survey implies your consent to participate.

You may ask any questions about the study before deciding whether or not to participate. Whether or not you participate is your choice, and if you decide to take part, you may choose to withdraw yourself and your data from the study at any time without any penalty, and without giving a reason, by advising the researchers of your decision. My contact information is provided above.

Storage of and access to collected data

The information collected from the survey will be stored in a secure online database provided by the University of Oxford, and will not include any information that can identify you as a survey participant. Only I and my direct supervisors will have access to the data. After the publication of the study, the data will be preserved in the Oxford Research Archives (ORD-Data) for three years.

Publication of research

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives researchers easy access to the full text of freely available theses, thereby increasing the likely impact and use of that research.

If you agree to participate in this project, the research will be written up as a part of a thesis.

On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will be published with open access. This study may also be published in a peer-reviewed journal.

Ethics review and how to raise a concern

This project has been reviewed by and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (CUREC reference number: MSIDREC-C1-2015-161). If you have a concern about any aspect of this project, please speak to me (the USA, 1-360-970-3953) and I will do my best to answer your query. I should acknowledge your concern within ten working days and give you an indication of how I intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford (using the contact details below) who will seek to resolve the matter in a reasonably expeditious manner:

Chair, Medical Sciences Inter-Divisional Research Ethics Committee;

E-mail: ethics@medsci.ox.ac.uk

Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

Your participation will be very much appreciated, and it will contribute greatly to the understanding of how cardiovascular prediction rules are created. **In case you are not able to participate in the survey, we would be very grateful if you could provide an explanation as to why and return it using the stamped return envelope.**

Thank you very much for reading this letter.

Sincerely,



Jong-Wook Ban MD, MSc

Appendix 2.2. Online questionnaire for the survey of authors.



Author survey of derivation studies of cardiovascular clinical prediction rules

Online questionnaire

- Please enter the unique alphanumeric code provided in the invitation:

- Please confirm that you have read and understood the information about the survey in the attachment by checking the box and pressing the “Start survey” button.

  o I read and understood the information provided.

1. At the time of derivation, were you aware of any existing prediction rules that addressed the same clinical problem?

   a. Yes
   b. No

2. How did you become aware of existing prediction rules that addressed the same clinical problem (Please select ALL that apply)?

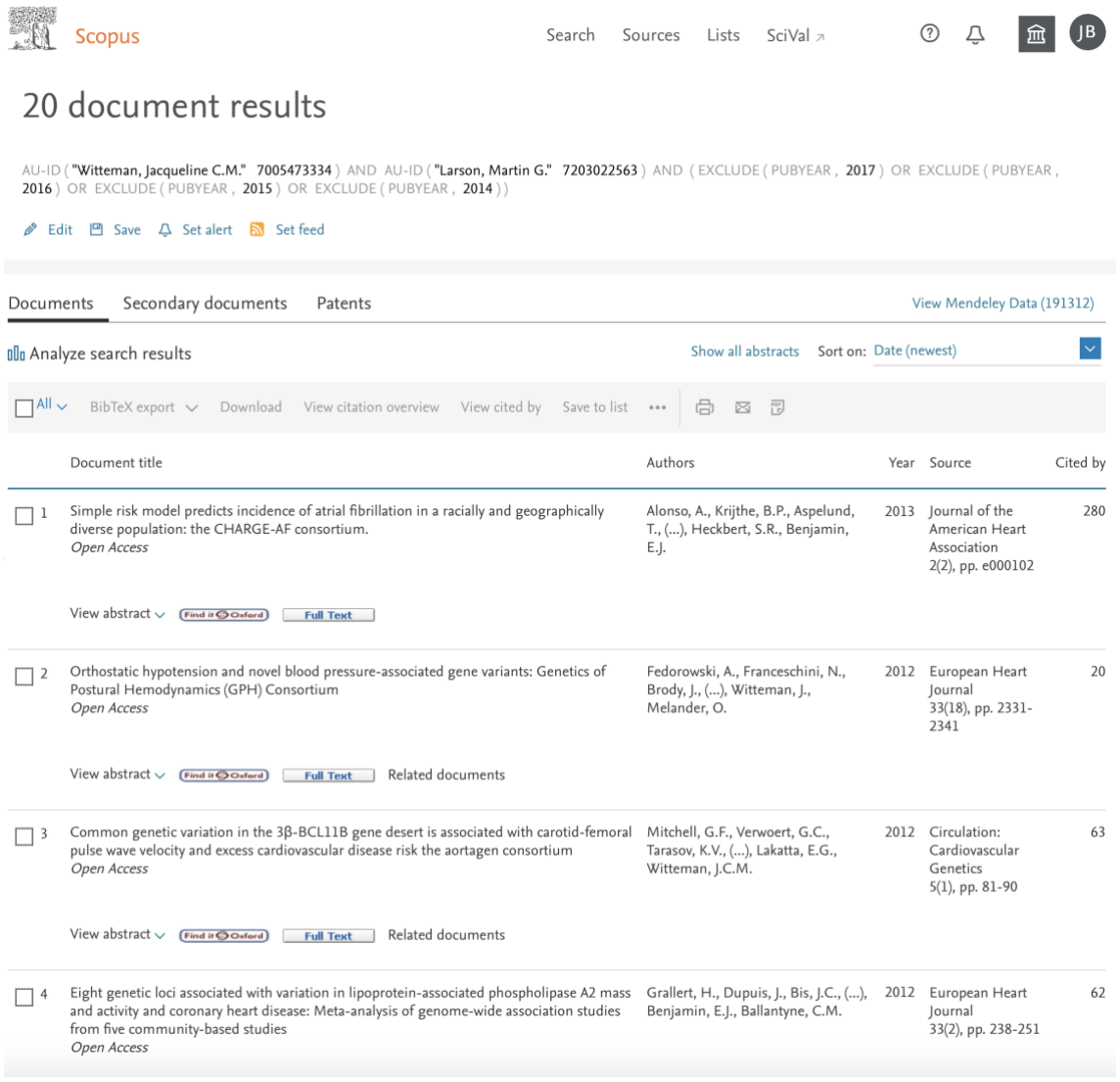
   a. Conducted a full systematic review
   b. Conducted a limited systematic review
   c. Searched my team’s collection of references
   d. Consulted experts in the clinical area
   e. Did not conduct a search

3. How important do you think it is to cite existing prediction rules of the same problem when deriving a new prediction rule?

   a. Very important
   b. Somewhat Important
   c. Somewhat unimportant
   d. Unimportant

Chapter 3.

Appendix 3.1. Investigating prior co-authorship using the advanced search option in Scopus.



Scopus Search Sources Lists SciVal

20 document results

AU-ID ("Witteman, Jacqueline C.M." 7005473334) AND AU-ID ("Larson, Martin G." 7203022563) AND (EXCLUDE (PUBYEAR , 2017) OR EXCLUDE (PUBYEAR , 2016) OR EXCLUDE (PUBYEAR , 2015) OR EXCLUDE (PUBYEAR , 2014))

Edit Save Set alert Set feed

Documents Secondary documents Patents View Mendeley Data (191312)

Analyze search results Show all abstracts Sort on: Date (newest)

All BibTeX export Download View citation overview View cited by Save to list

	Document title	Authors	Year	Source	Cited by
1	Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *Open Access*	Alonso, A., Krijthe, B.P., Aspelund, T., (...), Heckbert, S.R., Benjamin, E.J.	2013	Journal of the American Heart Association 2(2), pp. e000102	280
	View abstract Find it Oxford Full Text				
2	Orthostatic hypotension and novel blood pressure-associated gene variants: Genetics of Postural Hemodynamics (GPH) Consortium *Open Access*	Fedorowski, A., Franceschini, N., Brody, J., (...), Witteman, J., Melander, O.	2012	European Heart Journal 33(18), pp. 2331-2341	20
	View abstract Find it Oxford Full Text Related documents				
3	Common genetic variation in the 3β-BCL11B gene desert is associated with carotid-femoral pulse wave velocity and excess cardiovascular disease risk the aortagen consortium *Open Access*	Mitchell, G.F., Verwoert, G.C., Tarasov, K.V., (...), Lakatta, E.G., Witteman, J.C.M.	2012	Circulation: Cardiovascular Genetics 5(1), pp. 81-90	63
	View abstract Find it Oxford Full Text Related documents				
4	Eight genetic loci associated with variation in lipoprotein-associated phospholipase A2 mass and activity and coronary heart disease: Meta-analysis of genome-wide association studies from five community-based studies *Open Access*	Grallert, H., Dupuis, J., Bis, J.C., (...), Benjamin, E.J., Ballantyne, C.M.	2012	European Heart Journal 33(2), pp. 238-251	62

Chapter 4.

Appendix 4.1. Journal that published studies included in chapter 4.

Number of studies	Journals that published external validation study of Framingham Wilson CHD risk rule
3	Medicina Clinica
2	American Journal of Cardiology, Diabetic Care, European Heart Journal, Methods of Information in Medicine, PLoS ONE
1	American Heart Journal, Annals of Family Medicine, Annals of Internal Medicine, Archives of Internal Medicine, Atención Primaria, BMC Cardiovascular Disorders, Cardiology, Circulation Journal, Diabetic Medicine, European Journal of Preventive Cardiology, HIV Clinical Trials, Hypertension Research, International Journal of Epidemiology, Journal of Atherosclerosis and Thrombosis, Journal of Epidemiology and Community Health, Kidney International, Liver International, Liver Transplantation, Medical Journal of Australia, Preventing Chronic Disease, Renal Failure, Revista Colombiana de Cardiología, Revista Española de Cardiología, Revista Española de Salud Pública, Salud Pública de México, Scandinavian Journal of Primary Health Care

Number of studies	Journals that published studies with a performance measure of Framingham Wilson CHD risk rule
14	Journal of the American Medical Association
5	Journal of the American College of Cardiology
3	Circulation, European Heart Journal, European Journal of Preventive Cardiology, Heart
2	American Heart Journal, American Journal of Cardiology, Atherosclerosis
1	American Journal of Medicine, Annals of Internal Medicine, Arteriosclerosis Thrombosis Vascular Biology, British Journal of General Practice, Circulation-Cardiovascular Genetics, Diabetes, Diabetic Care, Diabetic Medicine, Frontiers in Genetics, Genetic Epidemiology, Growth Factors, International Journal of Cardiovascular Imaging, JACC-Cardiovascular Imaging, Journal of Cardiovascular Magnetic Resonance, Journal of Human Hypertension, Kidney International, Nefrologia, PLoS ONE, Preventive Medicine Reports, Revista Española de Cardiología, Scientific Reports

Appendix 4.2. Studies that derived a new cardiovascular CPR using the poor performance of Wilson CHD risk rule in the external validation study as a justification.

External validation study	Derivation study of a new cardiovascular CPR
Cañón-Barroso et al. (1)	Su X, Xu Y, Tan Z, Wang X, Yang P, Su Y, et al. Prediction for cardiovascular diseases based on laboratory data: An analysis of random forest model. J Clin Lab Anal. 2020;34(9):e23421.
Chien et al. (2)	Jamthikar A, Gupta D, Khanna NN, Saba L, Laird JR, Suri JS. Cardiovascular/stroke risk prevention: A new machine learning framework integrating carotid ultrasound image-based phenotypes and its harmonics with conventional risk factors. Indian Heart J. 2020;72(4):258-64.
Herrera et al. (3)	De Socio GV, Pucci G, Baldelli F, Schillaci G. Observed versus predicted cardiovascular events and all-cause death in HIV infection: a longitudinal cohort study. BMC Infect Dis. 2017;17(1):414.
Comín et al. (4)	Ramos R, Baena-Diez JM, Quesada M, Solanas P, Subirana I, Sala J, et al. Derivation and validation of REASON: a risk score identifying candidates to screen for peripheral arterial disease using ankle brachial index. Atherosclerosis. 2011;214(2):474-9.
Davis et al. (5)	Davis WA, Knuiman MW, Davis TM. An Australian cardiovascular risk equation for type 2 diabetes: the Fremantle Diabetes Study. Intern Med J. 2010;40(4):286-92.
DeFilippis et al. (6)	Cho IJ, Sung JM, Chang HJ, Chung N, Kim HC. Incremental Value of Repeated Risk Factor Measurements for Cardiovascular Disease Prediction in Middle-Aged Korean Adults: Results From the NHIS-HEALS (National Health Insurance System-National Health Screening Cohort). Circ Cardiovasc Qual Outcomes. 2017;10(11).
Ducloux et al. (7)	Israni AK, Snyder JJ, Skeans MA, Peng Y, Maclean JR, Weinhandl ED, et al. Predicting coronary heart disease after kidney transplantation: Patient Outcomes in Renal Transplantation (PORT) Study. Am J Transplant. 2010;10(2):338-53. Pita-Fernandez S, Pertega-Diaz S, Valdes-Canedo F, Seijo-Bestilleiro R, Seoane-Pillado T, Fernandez-Rivera C, et al. Incidence of cardiovascular events after kidney transplantation and cardiovascular risk scores: study protocol. BMC Cardiovasc Disord. 2011;11:2.
Empana et al. (8)	Balkau B, Hu G, Qiao Q, Tuomilehto J, Borch-Johnsen K, Pyorala K, et al. Prediction of the risk of cardiovascular mortality using a score that includes glucose as a risk factor. The DECODE Study. Diabetologia. 2004;47(12):2118-28.

External validation study	Derivation study of a new cardiovascular CPR
	Barzi F, Patel A, Gu D, Sritara P, Lam TH, Rodgers A, et al. Cardiovascular risk prediction tools for populations in Asia. J Epidemiol Community Health. 2007;61(2):115-21. Cournot M, Taraszkiewicz D, Cambou JP, Galinier M, Boccalon H, Hanaire-Broutin H, et al. Additional prognostic value of physical examination, exercise testing, and arterial ultrasonography for coronary risk assessment in primary prevention. American heart journal. 2009;158(5):845-51. Empana JP, Tafflet M, Escolano S, Vergnaux 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. 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.
Guckelberger et al. (9)	Moody WE, Holloway B, Arumugam P, Gill S, Wahid YS, Boivin CM, et al. Prognostic value of coronary risk factors, exercise capacity and single photon emission computed tomography in liver transplantation candidates: A 5-year follow-up study. J Nucl Cardiol. 2020.
Marrugat et al. (10)	Empana JP, Tafflet M, Escolano S, Vergnaux 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. Huerta JM, Tormo MJ, Gavrila D, Navarro C. Cardiovascular risk estimated after 13 years of follow-up in a low-incidence Mediterranean region with high-prevalence of cardiovascular risk factors. BMC Public Health. 2010;10:640.
Van der Heijden et al. (11)	Magri CJ, Debono R, Calleja N, Galea J, Fava S. Prognostic indicators and generation of novel risk equations for estimation of 10-year and 20-year mortality following acute coronary syndrome. Postgrad Med J. 2017;93(1099):245-9. Price AH, Weir CJ, Welsh P, McLachlan S, Strachan MWJ, Sattar N, et al. Comparison of non-traditional biomarkers, and combinations of biomarkers, for vascular risk prediction in people with type 2 diabetes: The Edinburgh Type 2 Diabetes Study. Atherosclerosis. 2017;264:67-73.

Chapter 5.

Appendix 5.1. Search strategy to identify systematic reviews that summarised design, methods, or reporting of clinical prediction rule studies.

1. Forecasting/
2. Models, Statistical/
3. Prognosis/
4. Multivariate Analysis/
5. 1 or 2 or 3 or 4
6. Biomedical Research/
7. Research Design/
8. Research Report/
9. "method*".m_titl.
10. "report*".m_titl.
11. 6 or 7 or 8 or 9 or 10
12. 5 and 11
13. limit 12 to “reviews (best balance of sensitivity and specificity)”

Appendix 5.2. Electronic database search strategies for impact studies of cardiovascular clinical prediction rules introduced by guidelines.

Australian absolute cardiovascular risk chart (12)	National Institute for Health and Care Excellence (NICE) Framingham equation (13)	National Institute for Health and Care Excellence (NICE) guideline for management of atrial fibrillation (14)	World Health Organization (WHO) 2007 for prevention of cardiovascular disease (15)
1. National Vascular Disease Prevention Alliance.mp.	1. National Institute for Clinical Excellence.mp.	1. National Institute for Clinical Excellence.mp.	1. World Health Organization/
2. NVDPA.mp.	2. National Institute for Health and Clinical Excellence.mp.	2. National Institute for Health and Clinical Excellence.mp.	2. International Society of Hypertension.mp.
3. (Australia$ adj3 guid$).mp.	3. National Institute for Health and Care Excellence.mp.	3. National Institute for Health and Care Excellence.mp.	3. (WHO adj3 guid$).mp.
4. 1 or 2 or 3	4. (NICE adj3 guid$).mp.	4. (NICE adj3 guid$).mp.	4. (ISH adj3 guid$).mp.
5. Cardiovascular disease$.mp.	5. 1 or 2 or 3 or 4	5. 1 or 2 or 3 or 4	5. 1 or 2 or 3 or 4
6. Cardiovascular risk$.mp.	6. Cardiovascular disease$.mp.	6. Atrial fibrillation.mp.	6. Cardiovascular disease$.mp.
7. CVD risk$.mp.	7. Cardiovascular risk$.mp.	7. A-fib.mp.	7. Cardiovascular risk$.mp.
8. CV risk$.mp.	8. CVD risk$.mp.	8. 6 or 7	8. CVD risk$.mp.
9. 5 or 6 or 7 or 8	9. CV risk$.mp.	9. limit 8 to yr="2013-2018"	9. CV risk$.mp.
10. limit 10 to yr="2013-2018"	10. 6 or 7 or 8 or 9		10. 6 or 7 or 8 or 9 limit 10 to
	11. limit 10 to yr="2013-2018"		11. yr="2013-2018"

Appendix 5.3. Instruments for reporting assessment.

5.3.1. Criteria for assessing the report of randomised controlled trials. [a, b, c]

	Section / topic	Criteria
Title and abstract		
1	Title and abstract	• Title identified the study as a randomised trial. • Abstract provided a structured summary of trial design, methods, results, and conclusions.
Introduction		
2	Background and objectives	• Scientific background and rationale for the trial were explained. • Specific objectives or hypotheses were presented.
Methods		
3	Trial design	• The trial design was described using terms such as parallel, crossover, or factorial with an allocation ratio (e.g. 1:1).
4	Participants	• Eligibility criteria for participants were provided. • Settings and locations (where the data were collected) were described.
5	Interventions	• For each group, the different components of the interventions were described with sufficient details for replication, including how and when they were actually administered. • The description included how the interventions were standardised • The description included how adherence of care providers with the protocol was assessed or enhanced.
6	Outcomes	• Primary (and secondary) outcome measures were defined with sufficient details for others to use, including how and when they were assessed.
7	Sample size	• How sample size was determined was explained.
8	Randomisation: sequence generation	• The method used to generate the random allocation sequence was reported. • Type of randomisation was described either by indicating that a simple randomisation was used or by describing details of any restriction applied (such as blocking and block size).
9	Randomisation: allocation concealment	• The mechanism used to implement the random allocation sequence was reported by describing any steps taken to conceal the sequence until interventions were assigned (e.g. sequentially numbered containers).
10	Implementation	• Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions were reported.
11	Blinding	• It was reported who was blinded after assignment to interventions (e.g. participants, care providers, those assessing outcomes) and how.
12	Statistical methods	• Statistical methods used to compare groups for primary (and secondary) outcomes were described.
Results		
13	Participant flow	• For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome were reported • For each group, losses and exclusions after randomisation were reported with reasons • The number of care providers or centres performing the intervention in each group and the number of patients treated by each care provider or in each centre were reported

		• Details of the experimental treatment and comparator as they were implemented was reported (e.g. Number of participants attended all planned treatment sessions).
14	Recruitment	• Dates defining the periods of recruitment and follow-up were reported.
15	Baseline data	• A table was provided to present baseline demographic and clinical characteristics for each group.
16	Numbers analysed	• For each group, the number of participants (denominator) included in each analysis was presented. • It was described whether the analysis was by the original assigned groups.
17	Outcome and estimation	• Results for each group, and the estimated effect size and its precision (such as 95% confidence interval) were presented for each primary (and secondary) outcome.
18	Harms	• Important harm or unintended effect was reported for each group.
Discussion		
19	Limitations	• Trial limitations were described by addressing sources of potential bias, imprecision, and, if relevant, the multiplicity of analyses. • In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centres in each group.
20	Generalizability	• Generalizability (external validity, applicability) of the trial findings was discussed according to the intervention, comparators, patients, care providers, or centres involved in the trial.
21	Interpretation	• The interpretation was consistent with the results and considered other relevant evidence.

[a] Criteria are based on relevant items from the CONSORT 2010 Statement (16) and an extension of the CONSORT statement for trials of nonpharmacological treatments (17).

[b] The following items from the CONSORT 2010 Statement (16) were excluded:

Item 3b: Important changes to methods after trial commencement (such as eligibility criteria), with reasons.

Item 6b: Any changes to trial outcomes after the trial commenced, with reasons.

Item 7b: When applicable, explanation of any interim analyses and stopping guidelines.

Item 11b: If relevant, description of the similarity of interventions.

Item 12b: Methods for additional analyses, such as subgroup analyses and adjusted analyses.

Item 14b: Why the trial ended or was stopped.

Item 17b: For binary outcomes, the presentation of both absolute and relative effect sizes is recommended.

Item 18: Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory.

Item 23: Registration number and name of the trial registry.

Item 24: Where the full trial protocol can be accessed, if available.

Item 25: Sources of funding and other support (such as a supply of drugs), the role of funders.

[c] The following items from an extension of the CONSORT statement for trials of nonpharmacological treatments (17) were excluded:

Item 3: When applicable, eligibility criteria for centres and those performing the interventions.

Item 7: When applicable, details of whether and how the clustering by care providers or centres was addressed.

Item 8: When applicable, how care providers were allocated to each trial group.

Item 11a: Whether or not those administering co-interventions were blinded to group assignment.

Item 11b: If blinded, method of blinding and description of the similarity of interventions.

Item 12: When applicable, details of whether and how the clustering by care providers or centres was addressed.

Item 15: When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centres (volume) in each group.

5.3.2. Criteria for assessing the report of cluster randomised trials. [a, b, c, d]

	Section / topic	Criteria
Title and abstract		
1	Title and abstract	• Title identified the study as a cluster randomised trial • Abstract provided a structured summary of trial design, methods, results, and conclusions
Introduction		
2	Background and objectives	• Scientific background, the rationale for the trial, and the rationale for using cluster design were explained. • Specific objectives or hypotheses were presented, and it was clear whether they pertain to cluster level, individual participant level, or both.
Methods		
3	Trial design	• The trial design was described using terms such as parallel, crossover, or factorial with an allocation ratio (e.g. 1:1). • The cluster was defined, and how the design features apply to the clusters were described.
4	Participants	• Eligibility criteria for participants and clusters were provided. • Settings and locations (where the data were collected) were described.
5	Interventions	• For each group, the different components of the interventions were described with sufficient details for replication, including how and when they were actually administered. • The description included how the interventions were standardised. • The description included how adherence of care providers with the protocol was assessed or enhanced. • It was described whether interventions pertain to the cluster level, the individual participant level, or both.
6	Outcomes	• Primary (and secondary) outcome measures were defined with sufficient details for others to use, including how and when they were assessed. • It was described whether outcome measures pertain to the cluster level, the individual participant level, or both.
7	Sample size	• How sample size was determined was explained with a number of cluster(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or k), and an indication of its uncertainty.
8	Randomisation: sequence generation	• The method used to generate the random allocation sequence was reported. • Type of randomisation was described either by indicating that a simple randomisation was used or by describing details of any restriction applied (such as blocking and block size).
9	Randomisation: allocation concealment	• The mechanism used to implement the random allocation sequence was reported by describing any steps taken to conceal the sequence until interventions were assigned (e.g. sequentially numbered containers). • It was specified that allocation was based on clusters rather than individuals and described whether allocation

		concealment (if any) was at the cluster level, the individual participant level, or both.
10	Implementation	• Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions were reported. • The mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling) was described. • It was described from whom consent was sought (representatives of the cluster, individual cluster members, or both) and whether consent was sought before or after randomisation.
11	Blinding	• It was reported who was blinded after assignment to interventions (e.g. participants, care providers, those assessing outcomes) and how.
12	Statistical methods	• Statistical methods used to compare groups for primary (and secondary) outcomes and how clustering was taken into account were described.
Results		
13	Participant flow	• For each group, the numbers of both participants and clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome were reported. • For each group, losses and exclusions after randomisation were reported for both clusters and individual cluster members, with reasons. • The number of care providers or centres performing the intervention in each group and the number of patients treated by each care provider or in each centre were reported. • Details of the experimental treatment and comparator as they were implemented was reported (e.g. the number of participants attended all planned treatment sessions).
14	Recruitment	• Dates defining the periods of recruitment and follow-up were reported.
15	Baseline data	• A table was provided to present baseline characteristics for the individual and cluster levels for each group.
16	Numbers analysed	• For each group, the number of participants (denominator) and clusters included in each analysis was presented. • It was described whether the analysis was by the original assigned groups.
17	Outcome and estimation	• Results for each group, and the estimated effect size and its precision (such as 95% confidence interval) were presented for each primary (and secondary) outcome. • A coefficient of intracluster correlation (ICC or k) was reported for each primary outcome
18	Harms	• Important harm or unintended effect was reported for each group.
Discussion		
19	Limitations	• Trial limitations were described by addressing sources of potential bias, imprecision, and, if relevant, the multiplicity of analyses.

		• In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centres in each group.
20	Generalizability	• Generalizability (external validity, applicability) of the trial findings to clusters and/or individual participants was discussed according to the intervention, comparators, patients, care providers, or centres involved in the trial.
21	Interpretation	• The interpretation was consistent with results and considered other relevant evidence.

[a] Criteria are based on relevant items from the CONSORT 2010 Statement (16), an extension of the CONSORT statement for trials of nonpharmacological treatments (17), and an extension of the CONSORT statement to cluster randomised trials (18).

[b] The following items from the CONSORT 2010 Statement (16) were excluded:

Item 3b: Important changes to methods after trial commencement (such as eligibility criteria), with reasons.

Item 6b: Any changes to trial outcomes after the trial commenced, with reasons.

Item 7b: When applicable, explanation of any interim analyses and stopping guidelines.

Item 11b: If relevant, description of the similarity of interventions.

Item 12b: Methods for additional analyses, such as subgroup analyses and adjusted analyses.

Item 14b: Why the trial ended or was stopped.

Item 17b: For binary outcomes, the presentation of both absolute and relative effect sizes is recommended.

Item 18: Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory.

Item 23: Registration number and name of the trial registry.

Item 24: Where the full trial protocol can be accessed, if available.

Item 25: Sources of funding and other support (such as a supply of drugs), the role of funders.

[c] The following items from an extension of the CONSORT statement for trials of nonpharmacological treatments (17) were excluded:

Item 3: When applicable, eligibility criteria for centres and those performing the interventions.

Item 7: When applicable, details of whether and how the clustering by care providers or centres was addressed

Item 8: When applicable, how care providers were allocated to each trial group.
Item 11a: Whether or not those administering co-interventions were blinded to group assignment.

Item 11b: If blinded, method of blinding and description of the similarity of interventions.

Item 12: When applicable, details of whether and how the clustering by care providers or centres was addressed.

Item 15: When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centres (volume) in each group.

[d] The following item from an extension of the CONSORT statement to cluster randomised trials (18) were excluded:

Item 8b: Details of stratification or matching if used.

5.3.3. Criteria for assessing the report of non-randomised trials. [a, b]

	Section / topic	Criteria
Title and abstract		
1	Title and abstract	• Information on how units were allocated to interventions (e.g. nonrandomised trial) was provided. • Structured abstract was used. • Information on the target population or study sample was provided.
Introduction		
2	Background	• Scientific background and rationale for the trial were explained.
Methods		
3	Participants	• Eligibility criteria for participants, including criteria at different levels in the recruitment/sampling plan (e.g., cities, clinics, subjects), were reported. • Recruitment setting and method of recruitment (e.g. referral, self-selection, sampling method) were described. • Settings and locations (where the data were collected) were described.
4	Interventions	• Interventions intended for each study condition, and how and when they were actually administered were described in detail, specifically including content, delivery method, unit of delivery, deliverer, setting, exposure quantity and duration, time span, and activities to increase compliance or adherence (e.g., incentives).
5	Objectives [c]	• Specific objectives or hypotheses were described.
6	Outcomes	• Primary (and secondary) outcome measures were defined with a description of any methods used to collect data.
7	Sample size	• How sample size was determined was explained.
8	Assignment method	• Unit of assignment (the unit being assigned to study condition, e.g., individual, group, community) was defined. • The method used to assign units to study conditions was reported.
9	Blinding	• It was described whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to study condition assignment; if so, a statement was included regarding how the blinding was accomplished and how it was assessed.
10	Unit of analysis	• Description of the smallest unit that is being analysed to assess intervention effects (e.g., individual, group, or community) was provided.
11	Statistical methods	• Statistical methods used to compare groups for primary (and secondary) outcomes were described.
Results		
12	Participant flow	• The flow of participants through each stage of the study was presented by reporting the number of participants in enrollment, assignment, allocation and intervention exposure, follow-up, and analysis. • Protocol deviations from the study as planned were reported with reasons.
13	Recruitment	• Dates defining the periods of recruitment and follow-up were reported.
14	Baseline data	• Baseline demographic and clinical characteristics of participants in each study condition were provided. • Baseline comparisons of those lost to follow-up and those retained, overall and by study condition were reported. • Comparison between the study population at baseline and target population of interest was described.

15	Baseline equivalence	• Data on study group equivalence at baseline (and statistical methods used to control for baseline differences either in methods or results section) were reported.
16	Numbers analysed	• The number of participants (denominator) included in each analysis for each study condition was reported. • It was described whether the analysis strategy was “intention to treat” or, if not, how non-compliers were treated in the analyses.
17	Outcomes and estimation	• For each primary and secondary outcome, a summary of results for each study condition, and the estimated effect size and a confidence interval to indicate the precision were presented.
18	Adverse events	• Important adverse events or unintended effect was reported for each group.
Discussion		
19	Interpretation (limitation)	• Results were interpreted by taking into account study hypotheses, sources of potential bias, imprecision of measures, multiplicative analyses, and other limitations or weaknesses of the study. • Results were discussed by taking into account the mechanism by which the intervention was intended to work (causal pathways) or alternative mechanisms or explanations. • Fidelity of implementation (success of and barriers to implementing the intervention) was discussed. • Research, programmatic, or policy implications were discussed.
20	Generalizability	• Generalizability (external validity) of the trial findings was discussed by taking into account the study population, the characteristics of the intervention, length of follow-up, incentives, compliance rates, specific sites/settings involved in the study, or other contextual issues.
21	Interpretation (Overall evidence)	• The general interpretation of the results was discussed in the context of current evidence.

[a] Criteria are based on relevant items from the TREND statement (19).

[b] The following items from the TREND statement (19) were excluded:

Item 2 second bullet: Theories used in designing behavioural interventions.

Item 6 third bullet: Information on validated instruments such as psychometric and biometric properties.

Item 8 third bullet: Inclusion of aspects employed to help minimise potential bias induced due to non-randomisation (e.g., matching).

Item 10 second bullet: If the unit of analysis differs from the unit of assignment, the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis).

Item 11 second bullet: Statistical methods used for additional analyses, such as subgroup analyses and adjusted analysis.

Item 11 third bullet: Methods for imputing missing data, if used.

Item fourth bullet: Statistical software or programs used.

Item 17 second bullet: Inclusion of null and negative findings.

Item 17 third bullet: Inclusion of results from testing pre-specified causal pathways through which the intervention was intended to operate, if any.

Item 18 first bullet: Summary of other analyses performed, including subgroup or restricted analyses, indicating which are pre-specified or exploratory.

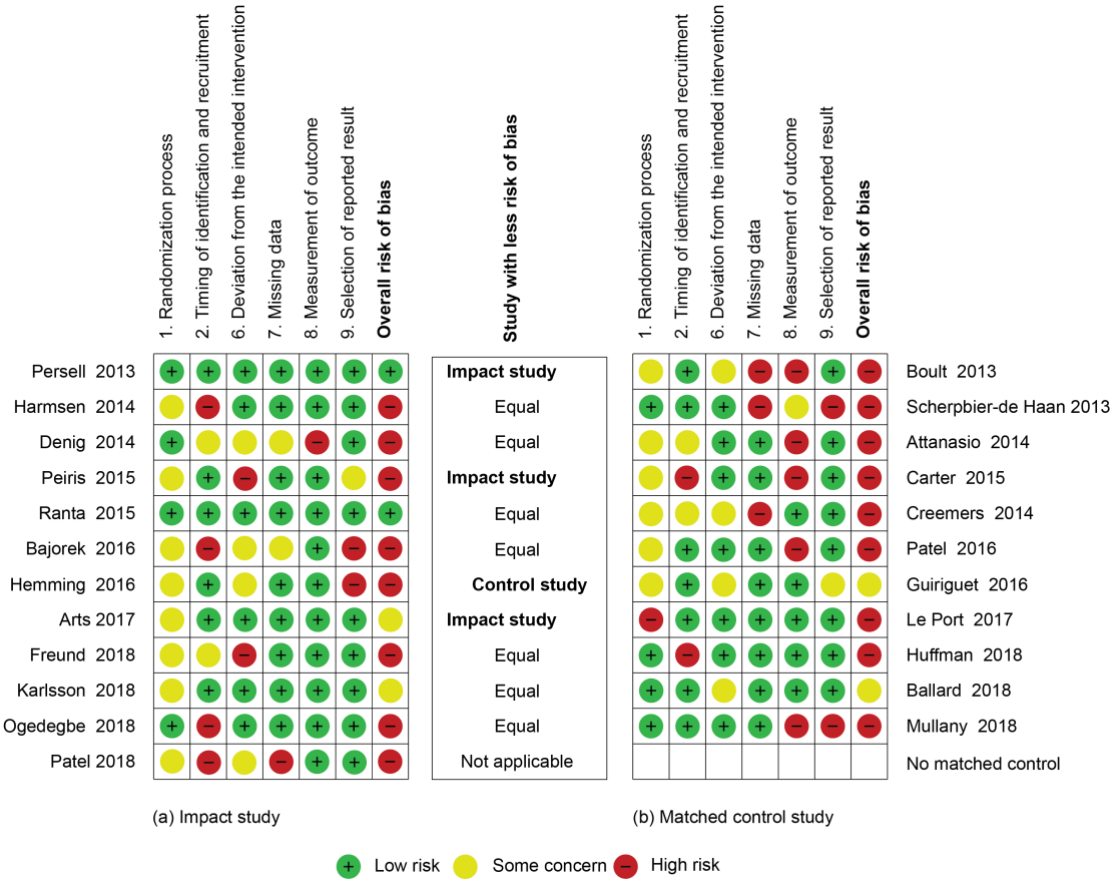
[c] We considered this item was reported properly when it was completely reported either in the introduction or method section.

Appendix 5.4. Risk of bias assessments.

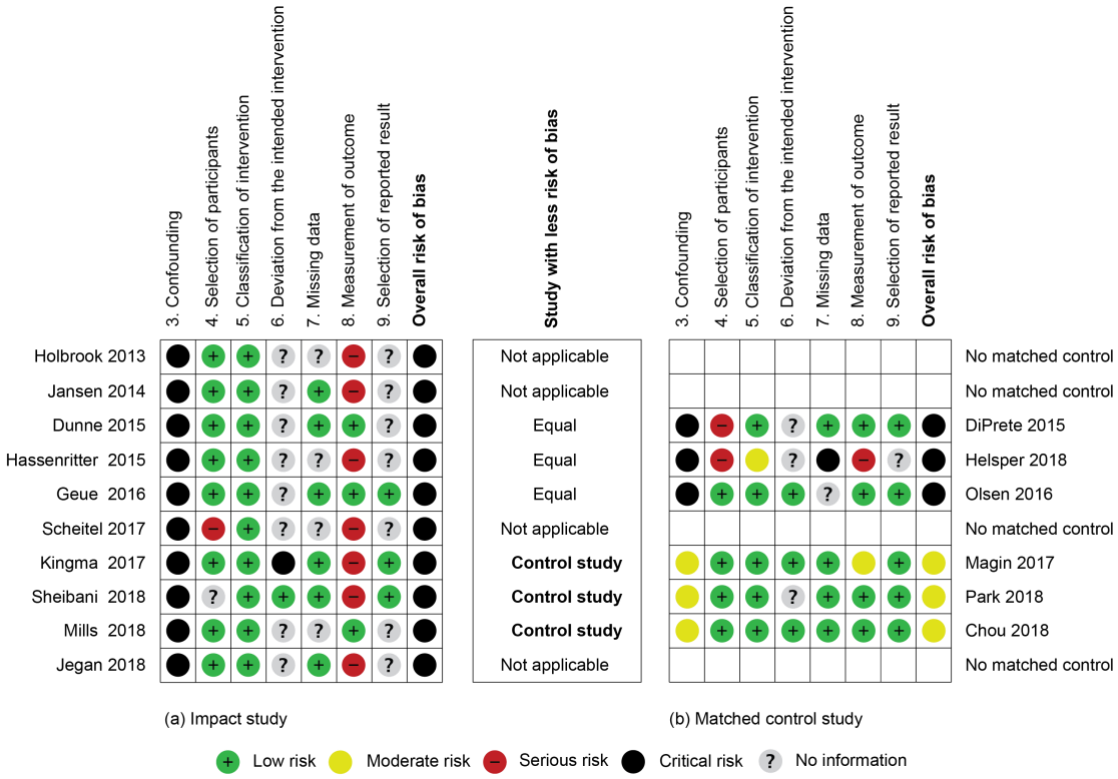
5.4.1. Risk-of-bias judgement for impact studies of cardiovascular clinical prediction rules and matched control studies of nonpharmacologic intervention with randomised control study design.



5.4.2. Risk-of-bias judgement for impact studies of cardiovascular clinical prediction rules and matched control studies of nonpharmacologic intervention with cluster randomised trial study design.



5.4.3. Risk-of-bias judgement for impact studies of cardiovascular clinical prediction rules and matched control studies of nonpharmacologic intervention with non-randomised experimental study design.



Appendix 5.5. Reporting assessments.

5.5.1. Randomised controlled trials that complied with reporting criteria.

CONSORT domain	Randomised controlled trial	
	Impact study, n = 23	Control study, n = 23
1. Title and abstract	18 (78.3)	16 (69.6)
2. Background and objectives	21 (91.3)	21 (91.3)
3. Trial design	0 (0.0)	2 (8.7)
4. Participants	15 (65.2)	11 (47.8)
5. Intervention	9 (39.1)	6 (27.1)
6. Outcomes	21 (91.3)	18 (78.3)
7. Sample size	19 (82.6)	21 (91.3)
8. Randomization: sequence generation	9 (39.1)	9 (39.1)
9. Randomization: allocation concealment mechanism	7 (30.4)	10 (43.5)
10. Randomization: implementation	2 (8.7)	6 (27.1)
11. Blinding	7 (30.4)	9 (39.1)
12. Statistical methods	23 (100.0)	23 (100.0)
13. Participant flow	4 (17.4)	2 (8.7)
14. Recruitment	17 (73.9)	18 (78.3)
15. Baseline data	23 (100.0)	21 (91.3)
16. Numbers analysed	8 (34.8)	8 (34.8)
17. Outcomes and estimation	16 (69.6)	15 (65.2)
18. Harms / Adverse events	7 (30.4)	11 (47.8)
19. Limitations	1 (4.3)	2 (8.7)
20. Generalizability	17 (73.9)	12 (52.2)
21. Interpretation	19 (82.6)	19 (82.6)
Mean number of studies that complied with reporting criteria	12.5 (54.5)	12.4 (53.8)

Values are numbers (%) of studies.

5.5.2. Cluster randomised trials that complied with reporting criteria.

CONSORT domain	Cluster randomised trial	
	Impact study, n = 12	Control study, n = 11
1. Title and abstract	10 (83.3)	9 (81.8)
2. Background and objectives	0 (0.0)	2 (18.2)
3. Trial design	2 (16.7)	2 (18.2)
4. Participants	5 (41.7)	9 (81.8)
5. Intervention	5 (41.7)	3 (27.3)
6. Outcomes	9 (75.0)	7 (63.6)
7. Sample size	6 (50.0)	4 (36.4)
8. Randomization: sequence generation	5 (41.7)	4 (36.4)
9. Randomization: allocation concealment mechanism	2 (16.7)	2 (18.2)
10. Randomization: implementation	1 (8.3)	0 (0.0)
11. Blinding	1 (8.3)	4 (36.4)
12. Statistical methods	12 (100.0)	11 (100.0)
13. Participant flow	3 (25.0)	4 (36.4)
14. Recruitment	9 (75.0)	9 (81.8)
15. Baseline data	8 (66.7)	7 (63.6)
16. Numbers analysed	3 (25.0)	3 (27.3)
17. Outcomes and estimation	8 (66.7)	3 (27.3)
18. Harms / Adverse events	5 (41.7)	6 (18.2)
19. Limitations	0 (0.0)	0 (0.0)
20. Generalizability	3 (25.0)	4 (36.4)
21. Interpretation	7 (58.3)	9 (81.8)
Mean number of studies that complied with reporting criteria	5.0 (41.3)	4.9 (44.2)

Values are numbers (%) of studies.

5.5.3. Non-randomised experiments that complied with reporting criteria.

TREND domain	Non-randomised experiments	
	Impact study, n = 10	Control study, n = 6
1. Title and abstract	2 (20.0)	4 (66.7)
2. Background	9 (90.0)	6 (100.0)
3. Participants	2 (20.0)	3 (50.0)
4. Intervention	4 (40.0)	4 (66.7)
5. Objectives	10 (100.0)	6 (100.0)
6. Outcomes	8 (80.0)	6 (100.0)
7. Sample size	5 (50.0)	1 (16.7)
8. Assignment method	6 (60.0)	3 (50.0)
9. Blinding	0 (0.0)	0 (0.0)
10. Unit of analysis	7 (70.0)	5 (83.3)
11. Statistical methods	9 (90.0)	5 (83.3)
12. Participant flow	0 (0.0)	0 (0.0)
13. Recruitment	7 (70.0)	6 (100.0)
14. Baseline data	1 (10.0)	1 (16.7)
15. Baseline equivalence	3 (30.0)	0 (0.0)
16. Numbers analysed	0 (0.0)	1 (16.7)
17. Outcomes and estimation	7 (70.0)	4 (66.7)
18. Adverse events	2 (20.0)	1 (16.7)
19. Interpretation (limitation)	0 (0.0)	0 (0.0)
20. Generalizability	2 (20.0)	1 (16.7)
21. Interpretation (Overall evidence)	9 (90.0)	6 (100.0)
Mean number of studies that complied with reporting criteria	4.4 (44.3)	3.0 (50.0)

Values are numbers (%) of studies.

Chapter 6.

Appendix 6.1. Electronic database search strategy to find studies that assessed the familiarity or use of cardiovascular CPRs.

	Search terms
1	(chd risk assessment$ or 2 cvd risk assessment$ or heart disease risk assessment$ or coronary disease risk assessment$ or cardiovascular disease risk assessment$ or cardiovascular risk assessment$ or cv risk assessment$ or cardiovascular disease$ risk assessment$ or coronary risk assessment$ or coronary risk scor$ or heart disease risk scor$ or chd risk scor$ or cardiovascular risk scor$ or cardiovascular disease$ risk scor$ or cvd risk scor$ or cv risk scor$).mp.
2	cardiovascular diseases/ or coronary disease/ or cardiovascular disease$.mp. or heart disease$.mp. or coronary disease$.mp. or cardiovascular risk?.mp. or coronary risk?.mp. or exp hypertension/ or exp hyperlipidemia/
3	risk function.mp. or Risk Assessment/mt or risk functions.mp. or risk equation$.mp. or risk chart?.mp. or (risk adj3 tool$).mp. or risk assessment function?.mp. or risk assessor.mp. or risk appraisal$.mp. or risk calculation$.mp. or risk calculator$.mp. or risk factor$ calculator$.mp. or risk factor$ calculation$.mp. or risk engine$.mp. or risk equation$.mp. or risk table$.mp. or risk threshold$.mp. or risk disc?.mp. or risk disk?.mp. or risk scoring method?.mp. or scoring scheme?.mp. or risk scoring system?.mp. or risk prediction?.mp. or predictive instrument?.mp. or project$ risk?.mp. or cdss.mp.
4	2 and 3
5	1 or 4
6	(new zealand chart$ or sheffield table$ or procam or General Rule to Enable Atheroma Treatment or dundee guideline$ or shaper scor$ or (brhs adj3 score$) or (brhs adj3 risk$) or copenhagen risk or precard or (framingham adj1 (function or functions)) or (framingham adj2 risk) or framingham equation or framingham model$ or (busselton adj2 risk$) or (busselton adj2 score$) or erica risk score$ or framingham scor$ or dundee scor$ or brhs scor$ or British Regional Heart study risk scor$ or brhs risk scor$ or dundee risk scor$ or framingham guideline$ or framingham risk? or new zealand table$ or ncep guideline? or smac guideline? or copenhagen risk?).mp.
7	5 or 6
8	exp decision support techniques/ or Diagnosis, Computer-Assisted/ or Decision Support Systems,Clinical/ or algorithms/ or algorithm?.mp. or algorythm?.mp. or decision support?.mp. or predictive model?.mp. or treatment decision?.mp. or scoring method$.mp. or (prediction$ adj3 method$).mp.
9	Risk Factors/ or exp Risk Assessment/ or (risk? adj1 assess$).mp. or risk factor?.mp.
10	2 and 8 and 9
11	7 or 10
12	stroke.mp. or exp Stroke/ or cerebrovascular.mp. or exp Cerebrovascular Circulation/
13	3 and 12
14	8 and 9 and 12
15	13 or 14
16	11 or 15

17	(ABCD2 or ABCD3 or CHADS2 or CHA2DS2-VASc or HAS-BLED or Wells scor$ or Wells criteria).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, or1ganism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
18	16 or 17
19	Health care surveys/ OR exp health surveys OR (question$ or survey$).ti,ab.
20	(Adopt$ or Use$ or Usage or Using or Uptake or Utilise $ or Aware$ or Knowledge$ or Familiar$).ti.
21	18 and 19 and 20

Appendix 6.2. Online questionnaire for general practitioners' (GPs) awareness and use of cardiovascular clinical prediction rules (CPRs) in the United Kingdom.



General practitioners' awareness and use of cardiovascular clinical prediction rules

Online questionnaire

Researchers from the Departments of Primary Care Health Sciences and Continuing Education at the University of Oxford invite you to participate in a survey. The aim of this study is to evaluate the awareness and use of cardiovascular clinical prediction rules among general practitioners (GPs) in the UK. You have been invited to participate as you are a member of Doctors.net.uk, opted to receive e-mails about research participation, and are a GP currently practising in the UK. The survey will take about 5 minutes to complete, and all eligible members completing the survey will receive 1,000 eSR points which are equivalent to £5. You will only be able to complete the survey once.

Please note that your participation is entirely voluntary, and you may withdraw at any point for any reason. Your answers will be completely anonymous, and we will use all reasonable endeavours to keep them confidential. Doctors.net.uk is the data controller with respect to your personal data and, as such, will determine how your personal data is used. Doctors.net.uk will share only fully anonymised data with the University of Oxford, for the purposes of research. Anonymised data will be stored in a password-protected file and may be used in academic publications. Research data will be stored for a minimum of three years after publication. Although Doctors.net.uk will collect IP addresses for the purpose of detecting fraudulent entries, they will be retained for no longer than necessary during the study period and will not be passed to other parties including researchers involved in this study.

This survey is a part of DPhil project. The principal researcher is Jong-Wook Ban (jong-wook.ban@conted.ox.uk), who is attached to the Centre for Evidence-Based Medicine at the University of Oxford. This project is being completed under the supervision of Dr Rafael Perera (rafael.perera@phc.ox.ac.uk). This project has been reviewed by and received ethics clearance through, the University of Oxford Central University Research Ethics Committee [R63716/RE001].

If you have a concern about any aspect of this project, please contact Jong-Wook Ban or his supervisor who will do their best to answer your query. The researcher should acknowledge your concern within ten working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant Chair of the Research Ethics Committee at the University of Oxford:

Chair, Medical Sciences Inter-Divisional Research Ethics Committee; E-mail: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD OR

The chair will seek to resolve the matter in a reasonably expeditious manner.

TC1 Informed consent

Please confirm that you are 18 years or over, have read and understood this information, and agree to participate

- ☐ Yes
- ☐ No THANK AND CLOSE

Section I: Eligibility

S1 Speciality confirmation: Are you a...

- GP
- Hospital Doctor THANK AND CLOSE

S2 GP type: Which of the following best describes your role in the practice?

- GP Partner/Principal
- Salaried GP
- Retainer GP
- Sessional GP
- GP Registrar/In training THANK AND CLOSE
- Locum GP
- Other THANK AND CLOSE

S3 Region: Where are you currently practising?

- North West SHA

- North East SHA
- Yorkshire & Humber SHA
- East Midlands SHA
- West Midlands SHA
- East of England SHA
- London SHA
- South East SHA
- South West SHA
- Scotland
- Wales
- Northern Ireland
- Retired THANK AND CLOSE
- Not working in the UK THANK AND CLOSE

Section II: CPRs for Cardiovascular Disease Risk- RANDOMISE THE ORDER OF THE SECTIONS

CD1 Which of the following clinical prediction rules for Cardiovascular Disease Risk have you heard of?

	Integrated in EHR	Heard of it	Not heard of it	Not sure
QRISK2 or QRISK3				
JBS3 calculator or QRISK Lifetime				
ASSIGN				
UKPDS				
Systematic Coronary Risk Evaluation (SCORE) risk charts or Heartscore				
PROCAM				
Any Framingham risk score				
ACC/AHA pooled cohort equation				
New Zealand tables				
NZ Primary prevention equations				

WHO/ISH risk prediction charts				

CD2 Please name any other clinical prediction rules you use for Cardiovascular Disease Risk? *Or select "I do not use any other clinical prediction rules for cardiovascular disease."*

☐ I do not use any other clinical prediction rules for cardiovascular disease risk.

CD3 How often do you use the following clinical prediction rules for Cardiovascular Disease Risk?

	In most or all relevant cases	Occasionally	Rarely	Never
QRISK2 or QRISK3				
JBS3 calculator or QRISK Lifetime				
ASSIGN				
UKPDS				
Systematic Coronary Risk Evaluation (SCORE) risk charts or Heartscore				
PROCAM				
Any Framingham risk score				
ACC/AHA pooled cohort equation				
New Zealand tables				
NZ Primary prevention equations				
WHO/ISH risk prediction charts				

CD4 For what reasons do you use a clinical prediction rule for Cardiovascular Disease Risk? *Please tick all that apply or select "I do not use clinical prediction rule for cardiovascular disease risk."*

- Aid diagnosis [e.g. rule in or rule out]
- Assessing severity
- To guide therapy

- To guide referral
- Comply with clinical guidelines / QOF
- Automatically generated by practice software
- Inform or educate patients
- Other [please specify]

☐ I do not use clinical prediction rule for cardiovascular disease risk EXCLUSIVE

Section III: CPRs for Stroke and Venous Thromboembolism - RANDOMISE THE ORDER OF THE SECTIONS

SV1 Which of the following clinical prediction rules for Stroke and Venous Thromboembolism have you heard of?

	Integrated in EHR	Heard of it	Not heard of it	Not sure
$ABCD^2$, $ABCD^3$ or $ABCD^3$-I				
California score				
$CHADS^2$ or CHA_2DS_2-VASc				
HAS-BLED				
Any Wells score for DVT				
Any Wells score for PE				
Geneva or revised Geneva score				
Pulmonary Embolism Severity Index (PESI) or simplified Pulmonary Embolism Severity Index (sPESI)				

SV2 Please name any other clinical prediction rules for Stroke and Venous Thromboembolism you use? *Or select "I do not use any other clinical prediction rule for Stroke and Venous Thromboembolism."*

☐ I do not use any other clinical prediction rule for Stroke and Venous Thromboembolism

SV3 How often do you use the following clinical prediction rule for Stroke and Venous Thromboembolism?

	In most or all relevant cases	Occasionally	Rarely	Never
$ABCD^2$, $ABCD^3$ or $ABCD^3$-I				
California score				
$CHADS^2$ or CHA_2DS_2-VASc				
HAS-BLED				
Any Wells score for DVT				
Any Wells score for PE				
Geneva or revised Geneva score				
Pulmonary Embolism Severity Index (PESI) or simplified Pulmonary Embolism Severity Index (sPESI)				

SV4 For what reasons do you use a clinical prediction rule for Stroke and Venous Thromboembolism? *Please tick all that apply or select "I do not use clinical prediction rule for Stroke and Venous Thromboembolism."*

- Aid diagnosis [e.g. rule in or rule out]
- Assessing severity
- To guide therapy
- To guide referral
- Comply with clinical guidelines / QOF
- Automatically generated by practice software
- Inform or educate patients
- Other [please specify]

☐ I do not use clinical prediction rule for Stroke and Venous Thromboembolism EXCLUSIVE

Any other clinical prediction rules that you use in clinical practice, including scores, charts, algorithms and computerised algorithms:

Section IV: Demographic questions

D1 Overall practice size: Please write in the number of GPs working in your practice (including yourself). *Please write in the number.* Validate to a minimum of 1

☐ I do not know.

D2 Practice Location: Is your practice based in a...

- Rural area
- Urban area
- Semi-rural area
- Suburban area

D3 Gender: Are you...

- Male
- Female
- Other
- Prefer not to say

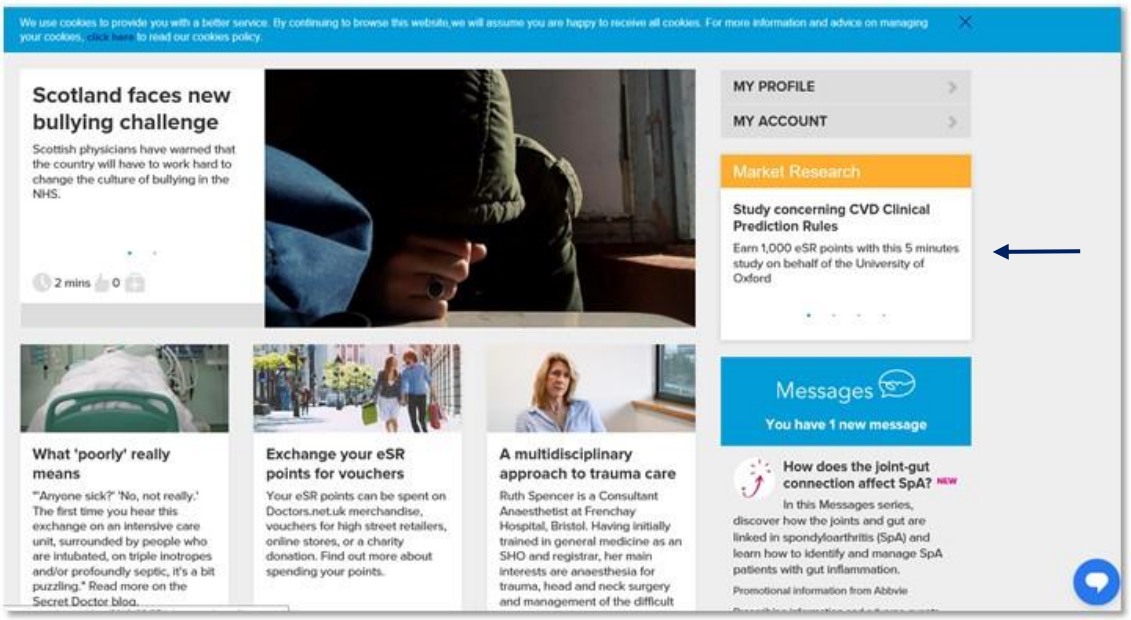
D4 Year of qualification: When did you qualify as a doctor? *Please write in the year you qualified.* Range of 1950 to 2019

D5 Research or teaching: Are you currently involved in either teaching students/trainees or research?

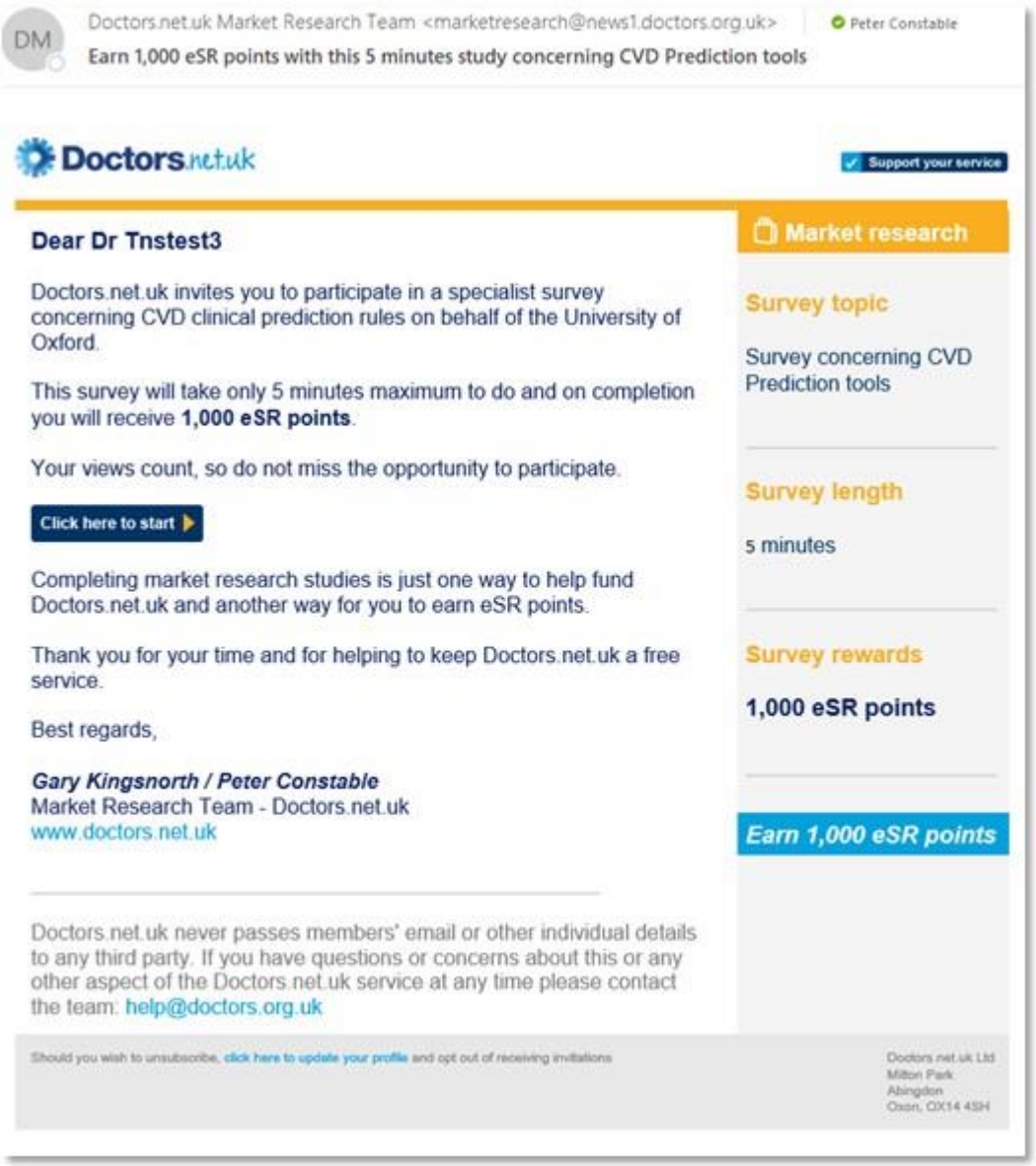
- Both teaching and research
- Teaching only
- Research only
- Neither

Appendix 6.3. Materials used to recruit general practitioners (GPs).

6.3.1. Advertisement on GP members' homepage of doctors.net.uk (arrow)



6.3.2. E-mail to GP members who opted to receive information about research participation



DM

Doctors.net.uk Market Research Team <marketresearch@news1.doctors.org.uk>

Peter Constable

Earn 1,000 eSR points with this 5 minutes study concerning CVD Prediction tools

Doctors.net.uk

Support your service

Dear Dr Tnstest3

Doctors.net.uk invites you to participate in a specialist survey concerning CVD clinical prediction rules on behalf of the University of Oxford.

This survey will take only 5 minutes maximum to do and on completion you will receive **1,000 eSR points**.

Your views count, so do not miss the opportunity to participate.

Click here to start

Completing market research studies is just one way to help fund Doctors.net.uk and another way for you to earn eSR points.

Thank you for your time and for helping to keep Doctors.net.uk a free service.

Best regards,

Gary Kingsnorth / Peter Constable
Market Research Team - Doctors.net.uk
www.doctors.net.uk

Doctors.net.uk never passes members' email or other individual details to any third party. If you have questions or concerns about this or any other aspect of the Doctors.net.uk service at any time please contact the team: help@doctors.org.uk

Market research

Survey topic

Survey concerning CVD Prediction tools

Survey length

5 minutes

Survey rewards

1,000 eSR points

Earn 1,000 eSR points

Should you wish to unsubscribe, click here to update your profile and opt out of receiving invitations

Doctors.net.uk Ltd
Milton Park
Abingdon
Oxon, OX14 4SH

Chapter 7.

Appendix 7.1. Search strategy to identify studies that evaluated the influence of research evidence on the uptake of clinical prediction rules.

1. Forecasting/
2. Models, Statistical/
3. Prognosis/
4. Multivariate Analysis/
5. 1 or 2 or 3 or 4
6. (Adopt$ or Use$ or Usage or Using or Uptake or Utilise $ or Aware$ or Knowledge$ or Familiar$).ti.
7. Barrier$.mp.
8. Facilitator$.mp.
9. "Attitude of Health Personnel"/
10. 7 or 8 or 9
11. 5 and 6 and 10

Appendix 7.2. Recruitment letter

Department of Primary Care Health Sciences
& Department of Continuing Education

Radcliffe Observatory Quarter
Woodstock Road, Oxford
OX2 6GG

Rewley House
1 Wellington Square, Oxford
OX1 2JA

Veronika Williams, Honorary Researcher
Jong-Wook Ban, DPhil candidate in Evidence-Based Health Care
Email: jong-wook.ban@conted.ox.ac.uk

Dear Primary care provider,

We, Jong-Wook Ban and Veronika Williams, are researchers at the Department of Primary Care Health Sciences and Department of Continuing Education, the University of Oxford, would like to invite you to take part in a study titled "Influence of evidence on primary care providers' uptake of cardiovascular clinical prediction rules: A qualitative interview study".

Clinical prediction rules (CPRs) are tools that help doctors and patients with diagnostic and prognostic decisions. Examples of CPRs for cardiovascular conditions include Atherosclerotic Cardiovascular Disease (ASCVD) risk calculator and Wells score for deep vein thrombosis. Although many cardiovascular CPRs have been developed, only a few are used by primary care providers (PCPs). Previously, we reported various flaws in cardiovascular CPR research such as developing CPRs without justifying their needs. Now, we would like to interview PCPs to ask how these flaws influenced their use of cardiovascular CPR.

We are looking for PCPs currently practising in the USA to take part in an interview that will last no more than 30 minutes. You will be asked to share your experiences about your decisions to use (or not to use) certain cardiovascular CPRs and about your views on cardiovascular CPR research. If you are interested, please complete an online screening questionnaire (https://form.jotform.com/92198332414255) that will help us select 20 PCPs for an interview. If you complete the questionnaire or take part in the interview, you will be compensated for your time.

If you are interested but would like more information please contact Jong-Wook Ban at the Department of Primary Care Health Sciences/Department of Continuing Education, University of Oxford at jong-wook.ban@conted.ox.ac.uk. There is no obligation to take part.

Many thanks and best wishes,



Jong-Wook Ban, MD, MSc

Appendix 7.3. Online questionnaire for primary care providers



Influence of evidence on primary care providers' uptake of cardiovascular clinical prediction rules: A qualitative interview study

Screening questionnaire

Ethics review approval reference: R69695/RE001

General Information

Researchers from the Departments of Primary Care Health Sciences and Continuing Education at the University of Oxford invite you to participate in a qualitative interview study. The aim of this study is to explore the primary care providers' (PCPs) perceptions of the evidence and process for developing cardiovascular clinical prediction rules (CPRs), and how PCPs' perceptions influence their decisions to use CPRs in practice.

In preparation, we would like you to take a short screening questionnaire to (1) to express your interest in taking part the interview, (2) to provide information about yourself as a PCP and your practice, and (3) to briefly describe your experiences with cardiovascular CPR and CPR research. We will use the information to invite a sample of PCPs for interview. The questionnaire will take no more than 5 minutes to complete, and you will receive a $5.00 Amazon eGift card for your time.

Do I have to take part?

Please note that your participation is entirely voluntary, and you may withdraw at any point for any reason. However, you will only receive the eGift card after completing the survey in full.

How will my data be used?

Although your answers to the questionnaire will not be anonymous, we will use all reasonable endeavours to keep them confidential. Your data will be stored in a password-protected file at the University of Oxford's server and may be used in academic publications. Research data will be stored for a minimum of three years after the publication. Any personal data (e.g. contact information) will be destroyed once the study is completed and they are no longer needed.

Who will have access to my data?

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data are used in the study. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/node/673106.

This study is a part of DPhil project. The principal researcher is Jong-Wook Ban (jong-wook.ban@conted.ox.uk), who is attached to the Centre for Evidence-Based Medicine at the University of Oxford. This study is being completed under the supervision of Dr Rafael Perera (rafael.perera@phc.ox.ac.uk).

This study has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee [Ethics Approval Reference: R69695/RE001].

Who do I contact if I have a concern about the study or I wish to complain?

If you have a concern about any aspect of this project, please contact Jong-Wook Ban or his supervisor who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant Chair of the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

E-mail: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford 0X1 2JD, United Kingdom

Informed consent

Please confirm that you are 18 years or over, have read and understood the above information, and agree to participate with the understanding that the data (including your personal data) you submit will be processed accordingly, please check the relevant box below to get started.

☐ Yes, I agree to take part

Demographics

1. Specialty confirmation: My speciality is...

☐ Family Practice

☐ Internal Medicine

☐ Geriatrics

☐ General practice

☐ Other:

2. Gender: I am...

☐ Male

☐ Female

☐ Other

☐ Prefer not to say

3. Duration of practice: In which year did you start practising as a primary care provider (PCP)?

[ ] (e.g. 1997)

4. PCP type: Which of the following best describes your role in the practice?

☐ Physician (MD or DO)

☐ Advanced Registered Nurse Practitioner (ARNP)

☐ Physician Assistant (PA)

☐ Physician in training (e.g. Resident)

☐ Other:

5. Teaching or research involvement: Have you been involved in either teaching students/trainees or research in the past 12 months?

☐ Both teaching and research

☐ Teaching only

☐ Research only

☐ Neither

Practice characteristics

1. Overall practice size: Please write in the number of PCPs working in your practice (including yourself).

[ ] (Enter 0 if you do not know)

2. Practice setting: My practice is based in a...

☐ Rural area

☐ Semi-rural area

☐ Suburban area

☐ Urban area

3. Practice location: Where are you currently practising?

☐ Western Washington

☐ Eastern Washington

☐ Wyoming

☐ Montana

☐ Alaska

☐ Idaho

☐ Other:

4. Integration of clinical prediction rule (CPR) in electronic health record (EHR) software: Please indicate whether any cardiovascular CPR is integrated in your electronic health record software.

☐ Practice does not have EHR software

☐ No cardiovascular CPR is integrated in EHR software

☐ At least one cardiovascular CPR is integrated in EHR software

Please name the electronic health record software your practice uses:

Use of cardiovascular clinical prediction rules

1. Use of cardiovascular clinical prediction rules (CPRs): Do you use any cardiovascular CPRs?

☐ Never

☐ Rarely

☐ Occasionally

☐ In most or all relevant cases

2. Research evidence: What research evidence did you consider before making your decision to use or not to use the cardiovascular CPRs?

☐ I did not consider any research evidence.

☐ I considered the following research evidence.

Please list all research evidence you considered:

3. Please discuss why you chose to consider each particular evidence or chose not to consider any research evidence.

4. Please describe how this research evidence influenced your decision to use or not to use the cardiovascular CPRs.

Participation in the interview study

1. Confirmation of interest: We are conducting an interview study to explore the PCPs' perceptions of the evidence and process for developing cardiovascular CPRs, and how PCPs' perceptions influence their decision to use CPRs in practice. You will receive a $50.00 Amazon eGift Care when invited for and complete a telephone or online interview. Would you like to be contacted for the interview?

☐ Yes, I confirm that I am interested in participating and would like to be contacted.

☐ No

2. Contact information: Please provide the following information so that we can send an eGift Card for completing the questionnaire and more information about the interview, including informed consent.

Name:

E-mail:

Phone Number:

Appendix 7.4. Topic guide for interviews with primary care providers (PCPs)

Research question/topic	Key concept	Interview question	Probe/follow-up question
Introduction	A. Introduction and context	Greeting and introduction. Briefly go over the interview purpose, scope, and process. Do you have any questions about the interview? Could you please confirm giving me permission to interview you and audio-recording the interview?	Confirm written consent. Obtain verbal consent when needed.
Transition	B. Clinical prediction rule (CPR) use	Could you please start by briefly telling me about you and your practice? In the online questionnaire, you indicated that you use cardiovascular CPRs rarely/occasionally/in most or all relevant cases. Can you tell me which CPRs you use and how you use them?	Ask a small follow-up question to build up a rapport. There are many similar CPRs for the problem. How did you decide that the CPR was the one to use in your practice? Are there any guideline-recommended cardiovascular CPRs that you are not using? Why?
What are PCPs' perceptions of and experiences with cardiovascular CPR research?	C. Familiarity and perception of cardiovascular CPR research	What does it mean to you when you hear the term "cardiovascular CPR research"? How would you define or describe it? You also indicated that you considered research evidence before deciding to use the CPR in the online questionnaire. What specific research	

		evidence did you consider? How did you know that using this CPR will accurately predict the risk of your patients? How did you know that using it will improve your practice?	
	D. Experience of using cardiovascular CPR research	If you recall reading any research paper about cardiovascular CPR, can you tell me about your experience with the paper? How did you come across it? How easy was it to read and understand the paper? How did you determine that the results of the paper were trustworthy?	What were the challenges?
	E. Views on cardiovascular CPR research	What are your thoughts/views on cardiovascular CPR research? Do you think PCPs should consider research evidence when he or she decides to use the CPR? Why?	
How do these perceptions and experiences influence their use of cardiovascular CPRs in practice?	F. Decision making	Did research evidence influence whether and how you use the CPR? How did these your experiences and thoughts on CPR research influence your decision to use the CPR?	What are the important things that will make you choose (or not choose) certain CPRs?
Closing		Thank the PCP for participating.	

Appendix 7.5. Demographic data of primary care providers (PCPs) interviewed.

Alphanumeric code	Year started	Speciality, Provider type	Practice location, setting	Practice size	Use of cardiovascular clinical prediction rule (CPR)
PR88	2019	Family Practice, Physician	Western Washington, Suburban	12	In most or all relevant cases
SY54	2019	Internal Medicine, Physician in training	Eastern Washington, Suburban	0	In most or all relevant cases
FA84	2018	Internal Medicine, Physician in training	Eastern Washington, Urban	60	Rarely
BG35	2017	Family Practice, Physician	Western Washington, Urban	50	In most or all relevant cases
GO38	2012	Internal Medicine, Physician	Eastern Washington, Suburban	36	Rarely
ZY58 [a]	2010	Family Practice, Physician	Idaho, Suburban	29	In most or all relevant cases
IS78	2010	Internal Medicine, Physician	Eastern Washington, Urban	35	Occasionally
FX91 [a]	2007	Family Practice, Physician	Western Washington, Urban	48	Occasionally
TF63	2007	Family Practice, Physician assistant	Idaho, Rural	7	Rarely
KJ34 [a]	1999	Family Practice, Physician	Western Washington, Urban	6	Occasionally
IN63 [b]	1996	Family Practice, Physician	Western Washington, Semi-rural	33	In most or all relevant cases
JA36	1991	Internal Medicine, Physician	Western Washington, Suburban	70	Occasionally
PX33 [a]	1984	Family Practice, Physician	Western Washington, Urban	30	In most or all relevant cases
QB54 [c]	1978	Family Practice, Physician	Idaho, Semi-rural	6	Occasionally
GR91 [a]	1978	Family Practice, Physician	Western Washington, Suburban	30	Occasionally

[a] PCP with research experience
[b] PCP who indicated that he did not consider research evidence when choosing cardiovascular CPR in the online questionnaire
[c] PCP was interviewed, but the interview was not recorded

References

1. 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.
2. Chien KL, Lin HJ, Su TC, Chen YY, Chen PC. Comparing the Consistency and Performance of Various Coronary Heart Disease Prediction Models for Primary Prevention Using a National Representative Cohort in Taiwan. Circ J. 2018;82(7):1805-12.
3. 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.
4. 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.
5. 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.
6. DeFilippis AP, Young R, Carrubba CJ, McEvoy JW, Budoff MJ, Blumenthal RS, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. Annals of internal medicine. 2015;162(4):266-75.
7. Ducloux D, Kazory A, Chalopin JM. Predicting coronary heart disease in renal transplant recipients: a prospective study. Kidney Int. 2004;66(1):441-7.
8. 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.
9. 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.
10. 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.
11. van der Heijden AA, Ortegon 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.
12. National Vascular Disease Prevention Alliance. Guidelines for the management of Absolute cardiovascular disease risk 2012 [Guideline]. Available from: https://www.heartfoundation.org.au/images/uploads/publications/Absolute-CVD-Risk-Full-Guidelines.pdf.
13. National Institute for Health and Care Excellence. Lipid modification: Cardiovascular risk assessment and the modification of blood lipids for the

primary and secondary prevention of cardiovascular disease 2008 [Available from: https://www.nice.org.uk/guidance/cg67.
14. National Institute for Health and Care Excellence. Atrial fibrillation: The management of atrial fibrillation 2006 [Available from: https://www.nice.org.uk/guidance/cg36.
15. World Health Organization. Prevention of cardiovascular disease: pocket guidelines for assessment and management of cardiovascular risk:(WH 2007 [Guideline]. Available from: https://www.who.int/cardiovascular_diseases/publications/Prevention_of_Cardiovascular_Disease/en/.
16. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
17. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, Group C. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Annals of internal medicine. 2008;148(4):295-309.
18. Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012;345:e5661.
19. Des Jarlais DC, Lyles C, Crepaz N, Group T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004;94(3):361-6.