Optimised Retrieval Of Primary Care Clinical Prediction Rules From MEDLINE To Establish A Web-Based Register

Claire Keogh[1]*, Emma Wallace[1], Kirsty K O'Brien[1], Paul J Murphy[2], Conor Teljeur[3], Brid McGrath[2], Susan M Smith[3], Niall Doherty[1], Borislav D Dimitrov[1] and Tom Fahey[1]

[1]HRB Centre for Primary Care Research, Department of General Practice,

RCSI Medical School, Royal College of Surgeons in Ireland;

[2]Mercer Library, Royal College of Surgeons in Ireland; and

[3]Department of Public Health and Primary Care, Trinity College, Dublin 2, Ireland

***Corresponding author:**

Dr Claire Keogh, BA (Psych) PhD

HRB Centre for Primary Care Research

Department of General Practice, RCSI Medical School

Royal College of Surgeons in Ireland

Dublin 2, Ireland

Tel: +353 1 4022317

Fax: +353 1 4022764

Email: clairekeogh@rcsi.ie

Abstract

Objective: Identifying clinical prediction rules (CPRs) for primary care from electronic databases is difficult. This study aims to identify a search filter to optimise retrieval of these to establish a register of CPRs for the Cochrane Primary Health Care field.

Study design and setting: Thirty primary care journals were manually searched for CPRs. This was compared to electronic search filters using alternative methodologies: (1) textword searching; (2) proximity searching; (3) inclusion terms using specific phrases and truncation; (4) exclusion terms; and (5) combinations of methodologies.

Results: We manually searched 6344 articles, revealing 41 CPRs. Across the 45 search filters, sensitivities ranged from 12% - 98%, while specificities ranged from 43% - 100%. There was generally a trade-off between the sensitivity and specificity of each filter (i.e. the number of CPRs and total number of articles retrieved). Combining textword searching with the inclusion terms (using specific phrases) resulted in the highest sensitivity (98%) but lower specificity (59%) than other methods. The associated precision (2%) and accuracy (60%) were also low.

Conclusion: The novel use of combing textword searching with inclusion terms was considered the most appropriate for updating a register of primary care CPRs where sensitivity has to be optimised.

Keywords: clinical prediction rules; primary care; medical information retrieval; search filters; proximity searching; evidence-based medicine

Running title: Search strategies for clinical prediction rules in primary care

What is new?

- The optimal sensitive filter results from combining text-word searching with inclusion terms that used specific word phrases. This provided the best balance between sensitivity (98%) and specificity (59%). Using the filter that retrieved the most CPRs resulted in a trade-off in terms of the large total number of articles that had to be searched through.
- This filter will now be used to establish an international register of primary care CPRs that will be made publically available through the Cochrane Primary Health Care field.
- Each filter was combined with various inclusion and exclusion filters. Using the Boolean term 'AND', the inclusion filter acted as a post-search filter that reduced the overall volume of articles to be screened for relevant CPRs.
- There is a real need for indexing CPRs on MEDLINE. This would greatly reduce difficulties in retrieving relevant articles.

1. Background

Clinical medicine has developed an increasing interest in clinical prediction rules (CPRs) [1 - 3]. A CPR may be defined as 'a clinical tool that quantifies the individual contributions that various components of the history, physical examination and basic laboratory results make towards the diagnosis, prognosis or likely response to treatment in a patient' [3]. They are tools that are derived from a single empirical study, as opposed to combining the results of multiple studies on the same topic. Examples of CPRs include the Ottawa ankle rule [4] to determine the need for an X-ray following an ankle injury and the Centor score [5] regarding identification of Group A β haemolytic streptococcal throat infection. Before a CPR can be used in clinical practice, it should pass through a vigorous testing process of five steps: derivation, validation (broad and narrow) and impact analysis (broad and narrow) [6]. Impact analysis determines the impact of the rule upon clinician behaviour, patient outcomes or healthcare quality and is, therefore, considered the most critical test.

As CPRs offer a useful guide for clinicians during diagnosis and prognosis, it would be advantageous for primary care clinicians to have easy access to relevant CPRs. However, this process presents a number of challenges. First, several terms are used interchangeably to describe CPRs including scorecard, algorithm and multivariate model, among others [7]. Second, several terms are used interchangeably in the international literature to describe primary care including 'family practice', 'family medicine' and 'general practice'. Third, commonly used electronic databases have no indexing term for this topic. For example, within MEDLINE, there is no Medical Subject Heading (MeSH) term for CPRs, nor is there a "type of article" limit. Consequently, research has focused on developing electronic search filters to search such databases for CPR articles [1, 2, 8].

In 2003, the Hedges team developed a number of filters to search electronic databases for CPRs [8]. This work generated the Haynes Broad Filter (HBF) and the Haynes Narrow Filter (HNF), both of which are available via the 'Clinical Query' search tool on PubMed MEDLINE. The two search methods offer a trade-off between sensitivity and specificity, with the HBF associated with 96% sensitivity and 79% specificity and the HNF associated with 54% sensitivity and 99% specificity. These filters were developed to identify all CPRs. However, as many of these are based in specialist settings, they are of limited utility to primary care clinicians. The efficacy of either filter in identifying primary care CPR articles has yet to be investigated.

The evidence to date indicates no easy method of accessing CPRs relevant to primary care, thus creating a barrier to implementing their routine use in clinical practice. One way to overcome this is to develop a publicly available electronic register of CPRs specifically for primary care. Efforts are currently underway to create this register in conjunction with the Cochrane Primary Health Care Field. To date there are 239 CPR articles contained on a preliminary version of the register, identified from personal libraries of clinicians and researchers (contact authors for details). Nevertheless, the question remains how best to identify all relevant CPRs to establish as complete a register as possible.

The current study investigates the utility of a number of electronic search filters relative to a manual search at identifying CPRs from a chosen set of journals relevant to primary care published in MEDLINE over a one-year period. Each filter was designed to be tested: (1) independently and (2) in combination with the other filters.

2. Methods

The results from the electronic search filters were compared to a 'reference standard' manual search of 30 journals relevant to primary care for the year 2008. Similar to previous research, the current study conducted a diagnostic test accuracy analysis - the articles retrieved from the electronic search filters were treated as "index test articles" while the manually retrieved articles were treated as the "reference standard articles" for CPRs relevant to primary care [1, 8]. The sensitivity, specificity, precision and accuracy for each of the filters were calculated (Table 1).

2.1. Journal selection

Thirty journals relevant to primary care (Table 2) were purposively chosen through various methods, including: (1) the ISI Web of Knowledge Journal Citation Reports, listed under the category 'medicine, general and internal' and mentioned primary care, family medicine or family practice in their title; (2) the 15 highest-ranked journals according to impact factor ratings in this same category; (3) specialist journals that are known to publish CPRs (based on type of journal/expert opinion); (4) a list of recommendations generated by an information specialist; and (5) an expert consensus meeting, attended by primary care clinicians, academics and information specialists (TF, BDD, SS, KOB, PM and BMcG).

2.2. Reference standard

Articles from each of the 30 journals for the year 2008 were downloaded from PubMed and were screened to exclude articles based on certain publication types: case reports; comments; dictionaries; editorials; and news. The resulting set of articles comprised our 'reference standard' search. Each article was then manually screened in EndNote by title and abstract and classified as CPRs relevant to primary care. For our purposes, a CPR is defined as *"... a*

clinical tool that quantifies the individual contributions that various components of the history, physical examination and basic laboratory results make towards the diagnosis, prognosis, or likely response to treatment in a patient. Clinical prediction rules attempt to formally test, simplify, and increase the accuracy of clinicians' diagnostic and prognostic assessments." [3]. Primary care is defined as "*... normally the point of first medical contact within the health care system, providing open and unlimited access to its users, dealing with all health problems regardless of the age, sex, or any other characteristic of the person concerned (*WONCA)". . An overview of the screening process is presented in Figure 1.

Twelve researchers were involved in the screening process (see Appendix 2 for details on training). Each of the articles (n = 6344) were screened independently by two reviewers. For the *first round* of screening, each pair of researchers was provided with an EndNote file that contained approximately 1000 abstracts of unique articles. Each article was classified as: (1) a CPR relevant to primary care; (2) unsure if a CPR or unsure if relevant to primary care; or (3) neither relevant to primary care nor a CPR. Results were compared between each pair of researchers and any disagreements were resolved by discussion. The articles (n = 100) in the unsure category were those that could not be readily classified as CPRs relevant to primary care on the basis of the title and abstract alone. These articles were further subjected to a *second round* of screening by a panel of four researchers with the most experience in classifying CPRs (TF, BDD, SS, KOB). Each article was classified as: (1) a CPR relevant to primary care; (2) a CPR in a specialist setting; or (3) not a CPR; or (4) full text required. In total, 24 full texts were retrieved and subjected to a *third round* of screening. Each full text article was independently screened by two experts. Articles were classified according to the same categories used previously. Results were compared and any disagreements were

resolved by discussion. This set of articles, designated as *"CPRs relevant to primary care"* were then considered the 'reference standard' subset.

2.3. Electronic filter search

Each of the electronic filters was run in MEDLINE. Each filter was run (1) independently and (2) in combination with other filters. The filters were run in either PubMed or Ebsco host for the selected 30 journal titles for the year 2008 (see Appendix 1 for search strings). Each was limited to humans and irrelevant publication types were excluded (case reports, comments, dictionaries, editorials and news).

2.4. Haynes filters

The Haynes Broad Filter (HBF) and the Haynes Narrow Filter (HNF) include textwords and Medical Subject Headings (MeSH) search terms and have been described in detail elsewhere [8]. Previous research has validated the use of textword searching in identifying all clinically relevant CPRs (not just those relevant to primary care) and reported similar levels of sensitivity and specificity, despite using various sizes of journal sets [1, 8].

2.5. McGrath/Murphy filters

The McGrath/Murphy Broad Filter (MMBF) and McGrath/Murphy Narrow Filter (MMNF) were developed in-house by two library information specialists (PM and BMcG – see Appendix 1). These filters involved proximity searching and were run on EBSCO host (as opposed to PubMed) which is able to facilitate this type of searching. To develop a proximity filter, a sample of CPR studies was selected from the 30 primary care journals. A content analysis of the MEDLINE records for this sample of articles was performed, leading to the derivation of a search string that specified proximate words or terms in significant

relationship within MEDLINE records. For example, phrases such as 'prediction models' or 'decision rules' in this or in any word order were matched using proximity operators, thus enabling proximate word matching and a potentially more accurate search string relative to that developed previously [1, 8]. For example, the EBSCO host search string "predict*N3 rule*" translates into "match the truncated word root 'predict' and find it nearby and in any order within three words of the truncated word root 'rule'."

2.6. Teljeur/Murphy filters

An alternative search methodology was developed (CT and PM) which resulted in three novel filters. These filters were derived from a content analysis of both CPR and non-CPR articles, using statistical packages that determine the word frequency in articles (R 2.9.1 and Microsoft Excel; CT). To determine the words that could be used to identify CPR articles, a content analysis of the titles and abstracts of 239 articles on the existing preliminary CPR register was performed. An equivalent content analysis was conducted on a reference set of non-CPR articles, defined as articles drawn from a random selection of 10 days during the first three months of 2008, as indexed by PubMed (n = 6447). To allow for equal comparison between the two sets of articles, all non-CPR articles had to be written in English and were limited to human content. The results from the two sets of articles were compared to determine which words were more common in CPR articles than in non-CPR articles.

Words that were at least twice as likely to appear in CPR articles as non-CPR articles were considered for the two inclusion filters. As many of the relevant words appeared in pairs, it was decided that the filter should include word pairs rather than individual words. The word pairs identified are considered to be universal to CPRs as none of the pairs are specific to primary care. The first inclusion filter was based in a set of 26 word pairs, the

Teljeur/Murphy Inclusion Filter 26 items (TMIF-26). This filter retrieved articles that contained specific phrases that appeared in a specific order. For example, an article was only retrieved if it contained the entire phrase "clinical prediction" in this order, but not if it contained the single words "clinical" or "prediction".

The second inclusion filter was based on a set of 22 word pairs, the Teljeur/Murphy Inclusion Filter 22 items (TMIF-22). This filter restricted searching for relevant word pairs to the title and abstract. Furthermore, an increased number of search words were truncated, for example, predict*. The process of truncation, in this example, finds words that begin with the root term 'predict', such as 'prediction' or 'predictive' etc. As a result, the overall number of search terms is reduced from 26 to 22, given that, for example, predict* AND rule* (TIF-22) retrieves "prediction rule", as well as "predictive rule" included in the previous search (TIF-26).

A set of 30 exclusion search terms was also identified and comprised the Teljeur/Murphy Exclusion Filter (TMEF). Exclusion words were identified as words with a relatively high frequency in the non-CPR set but not appearing in the CPR set. The initial list was further reduced to terms deemed highly improbable to appear in a CPR relevant to primary care. Thus, the exclusion terms removed articles from non-relevant settings (e.g. genetics-based research).

2.7. Combination of filters

The seven individual filters (HBF, HNF, MMBF, MMNF, TMIF-26, TMIF-22 and TMEF) were combined with the Teljeur/Murphy inclusion and exclusion filters (TMIF-26, TMIF-22 and/or TMEF). Firstly, the individual filters were combined with the inclusion terms using

the Boolean search term ‘AND’ (e.g. HBF AND TMIF-26). In this way, the inclusion filters acted as a post-search filter, designed to reduce the overall volume of articles to be screened for relevant CPRs. Secondly, the individual filters were combined with the inclusion filters using the Boolean search term ‘OR’ (e.g. HBF OR TMIF-26). In this way, the inclusion filters acted as an attempt to increase the number of relevant CPRs retrieved, despite potentially increasing the overall volume of articles to be screened. Finally, the individual filters were combined with the exclusion filters using the Boolean search term ‘NOT’ (e.g. HBF NOT TMIF-26). In this way, the exclusion terms also acted as a post-search filter to reduce the overall volume of articles to be screened.

3. Results

3.1. Reference standard

The reference standard manual search of the 30 journals identified 6344 articles for further analysis. Forty one of these articles were classified as “CPRs relevant to primary care” (Figure 1). These comprised the 'reference standard' subset (see Appendix 3 for details).

3.2. Overview of electronic filters

To calculate the test accuracy scores, the manual search was used as the reference standard and the electronic filters were used as the index tests. Results are presented in Table 3. All electronic filters retrieved a smaller number of articles than the manual search. The broad filters (HBF and MMBF) retrieved a higher total number of articles and more CPRs relevant to primary care than the two narrow filters (HNF and MMNF). Of the two inclusion filters, the TMIF-26 resulted in a higher yield of articles overall, as well a higher number of relevant CPRs relative to the TMIF-22. The exclusion filter (TMEF) retrieved the highest number of articles overall but a relatively low number of relevant CPRs. Each of the broad and narrow filters was also combined with each of TMIF-26, TMIF-22 and TMEF. When combined with the Boolean search term ‘AND’, the addition of the inclusion filters decreased the total number of articles retrieved by each filter, but also frequently decreased the number of relevant CPRs. In contrast, combining the search filters with the Boolean search term ‘OR’, the addition of the inclusion filters increased the total number of articles retrieved by each filter but also increased the number of relevant CPRs retrieved. Combining the broad and narrow filters with the exclusion filter decreased the total number of articles retrieved, as well as decreasing the number of relevant CPRs retrieved. Combining each of the inclusion filters with the exclusion filter also decreased both the total number of articles and the relevant CPRs retrieved by each filter.

3.3. The individual filters

The HBF was associated with higher sensitivity (76%) and lower specificity (81%) than the HNF (29% and 99%, respectively). The MMBF was associated with higher sensitivity (56%) and lower specificity (96%) than the MMNF (39% and 99%, respectively). The two inclusion filters reported different results, with the TMIF-26 resulting in higher sensitivity (95%) and lower specificity (62%) than the TMIF-22 (83% and 90%, respectively). The exclusion filter (TMEF via PubMed) results in the lowest sensitivity and specificity (59% and 43%, respectively). Note that the exclusion filter was also tested in Ebsco host and retrieved a total of 7464 articles. This latter result was not further analysed as the yield of articles was higher than that of the original hand search method.

3.4. Combination of filters

The individual Haynes Filters were combined with the Teljeur/Murphy inclusion or exclusion filters. For the HBF combinations with the inclusion filter using 'AND' and combinations with the exclusion filters reduced the sensitivity (by 3% - 44%) and increased the specificity (by 7% - 14%). However, combinations with the inclusion filter using 'OR' actually increased the sensitivity (by 9% - 22%) and decreased the specificity (by 5% - 30%) For the HNF, combination with the inclusion filters using 'AND' and combinations with the exclusion filters reduced the sensitivity (by 2% - 12%) and maintained the specificity (99%). In contrast, the combinations with the inclusion filter using 'OR' and the exclusion filters increased the sensitivity (by 3% - 69%) and decreased the specificity (by 4% - 40%). Combinations between the individual McGrath/Murphy Filters and the Teljeur/Murphy inclusion or exclusion filters also changed their initial outcome. For the MMBF, the various combinations of inclusion filters with 'AND' and the exclusion filters maintained/reduced the

sensitivity (up to 17%) and increased the specificity (by 1% - 3%). However, the combinations with the inclusion filter using 'OR' and the exclusion filters increased the sensitivity (by 32%) and decreased the specificity (by 6% - 12%). For the MMNF, the combination of inclusion filters with 'AND' and the exclusion filters maintained/reduced the sensitivity (up to 27%) and maintained/reduced the specificity (up to 1%). In contrast, the combinations with the inclusion filter using 'OR' and the exclusion filters increased the sensitivity (by 44% - 46%) and decreased the specificity (by 7%-13%).

Finally, for the Teljeur/Murphy Filters, combining the inclusion filters with the exclusion filter served to decrease the sensitivity (by 32% - 39%) and increase the specificity (by 4% - 19%) relative to running each inclusion filter independently.

3.5. Precision and accuracy

The level of precision across all filters was low, ranging from 1% to 37%. In general, the reported level of accuracy across all electronic filters was high, with all filters >76%, with the exception of the TMIF-26 (62%), HNF OR TMIF-26 (60%) and the TMEF (44%). This suggests that the majority of articles retrieved from the 30 journals were correctly classified by most of the filters.

3.6. Search strategy for register

For purposes of identifying a search strategy to best identify CPRs relevant to primary care and establish the register, the 'HBF OR TMIF-26' filter and the 'HNF OR TMIF-26' filter offered the best sensitivity (98%), thus providing access to the highest yield of relevant CPRs out of the 45 filters. Given that the two filters resulted in the same sensitivity, the optimal filter for our needs would provide the best balance between sensitivity and specificity. In the

current study, the ‘HNF OR TMIF-26’ filter offers better specificity (59%) than the ‘HBF OR TMIF-26’ filter (51%). However, the high sensitivity resulted in a trade-off for low specificity (59%), precision (2%) and accuracy (60%). Nevertheless, restricting the search of the MEDLINE database to the 30 journals relevant to primary care and application of the ‘HNF ORTMIF-26’ to this journal set resulted in significantly fewer articles to be screened for CPRs when compared to all of the articles published by PubMed every year (see Figure 2).

4. Discussion

CPRs are a valuable tool in supporting primary care clinicians in making evidence based decisions. Our work takes a pragmatic approach in analyzing a number of electronic filters to identify CPRs relevant to this setting. The work presented here contributes to generating an international register of CPRs relevant to primary care, which will be made publically available through collaboration with the Cochrane Primary Health Care field.

4.1. Main results

The electronic filters presented here offer a number of novel search methodologies to identify CPRs, thus building on the previous published work. In order to establish a register of primary care CPRs many of the new search methods were more advantageous than the previous Haynes filters. The use of proximity searching detailed in the McGrath/Murphy Filters offered mixed results. Specifically, the broad filter (MMBF) resulted in lower sensitivity and higher specificity relative to the HBF. In contrast, the narrow filter (MMNF) resulted in higher sensitivity and similar high specificity relative to the HNF. The use of inclusion filters, using either specific phrase searching (TMIF-26) or searching the title and abstract and truncating search terms (TMIF-22), resulted in some of the highest sensitivities. Interestingly, the combination of either of the inclusion filters with the Boolean search term 'AND' to the HBF roughly halved the total number of articles retrieved by the HBF, without significantly impacting on the number of relevant CPRs retrieved, suggesting the utility of the inclusion filters as a post-search method to remove irrelevant articles from already established search methodologies (e.g. HBF or HNF). In contrast, combinations of the inclusion filters with the Boolean search term 'OR' increased the total yield of relevant CPR articles but also increased the total number of articles retrieved by the filter.

The exclusion filter (TMEF) did not operate as expected. When executed via Ebsco host, the TMEF retrieved more articles than the hand search method. When executed via PubMed, the TMEF resulted in relatively low levels of sensitivity and specificity. Indeed, the combination of the TMEF and other filters in PubMed, for example the HBF and TMEF, actually removed relevant CPRs that had been retrieved by the individual filter (i.e. HBF alone). Although it may be possible to execute searches based on exclusion terms using a different search interface or using different search terms, the TMEF presented here does not offer a suitable method for updating the register.

4.2. Context of previous research

In the current study, the HBF and HNF were associated with lower levels of sensitivity and higher specificity than previous published research indicated. This effect can most likely be attributed to the use of a smaller set of journals, each purposively selected according to our specific criteria. Despite the loss in sensitivity, the benefit of working with this smaller set of journals is evident from the smaller number of articles retrieved by each filter relative to previous research [8], making the process of identifying primary care related CPRs less arduous.

4.3. Limitations

One potential limitation is the low level of precision reported across all filters (range 1 – 37%). This is understandable given the relatively small number of primary care CPRs to the total number of articles in the 30 journals in 2008. Notably, our low levels of precision are consistent with those reported previously as being inherent in all search strategies in the area of CPRs and not necessarily specific for our current study [8].

It should also be noted that the reference standard manual search for the year 2008 identified 41 CPR articles relevant to primary care. Previous research on retrieving treatment, diagnosis and prognosis studies indicates that a minimum of 99 relevant articles should be identified in the hand search process in order to create new search strategies for the MEDLINE database [50]. The current work falls short of this proposed standard.

The current work may be open to some criticism for the pre-screening step executed during the manual search of the 30 journals relevant to primary care. Specifically, prior to the first round of screening by the reviewers, certain publication types were intentionally excluded (case reports, comments, dictionaries, editorials and news). This pre-screening approach was previously used by some researchers but not by others when developing search strategies to retrieve clinically relevant literature [8, 51]. Although this pre-screening step was not considered to have any impact on the reported levels of sensitivity, it may have resulted in an over-estimation of specificity and precision [51]. For example, including these articles in our optimal search filter for this context would decrease the specificity by 20% (from 59% to 39%) and the precision by 1% (from 2% to 1%). This would result in about an extra 1300 articles to search through. Although it remains possible that CPRs could be mentioned in comments or in editorials, a pragmatic decision was made to omit these publication types because it is highly unlikely that any new studies would be presented in this format only, without the original article indexed and retrieved from the retrospective and/or prospective application of the search filter to the MEDLINE database.

4.4. Future research

It would be unnecessary to run these filters in MEDLINE, if MeSH terms for CPRs or "type of article" designation as a CPR existed. The availability of such indexing terms would

significantly reduce the number of articles to review when searching for CPRs on any given topic. However, even if such indexing were available, the identification of relevant CPRs would remain difficult given the problems associated with searching for 'primary care' specific articles. Generating and maintaining a publicly available international register currently appears to be the only viable way to allow an ease of access to articles on CPRs relevant to primary care.

From the perspective of developing a register of primary care CPRs, the novel methodological approach used in combining the Haynes Narrow Filter with the Teljeur/Murphy Inclusion Filter-26 item (HNF ORTMIF-26) provided the search method where sensitivity was optimised and where the best balance between sensitivity and specificity was achieved. Although the high sensitivity is associated with a trade-off for low levels of specificity, precision and accuracy, it is a necessary sacrifice given the aim of the current work is to retrieve as many primary care CPRs as possible. The high sensitivity achieved with the 'NHF OR TMIF-26' filter resulted in 40 of the 41 relevant CPRs being retrieved. The one missing article was not retrieved by any of the alternative filters. The discordance between the various search filters highlights the difficulty in developing optimal search strings, a problem generic to this type of research.

The 'HNF OR TMIF-26' filter will now be used to retrospectively and prospectively search the MEDLINE database. Although this filter will contribute to the best regular update of a CPR register, it is apparent that this method is insufficient to retrieve all relevant CPRs for the register. This work will need to be supplemented by searching additional databases (e.g., Embase, LILACS and Cochrane), searching references of relevant articles, conducting PubMed searches for each of the most prolific authors in the field of primary care diagnostic

research, contacting key authors in the field for published and non-published articles, searching non-priority journals and other secondary resources, for example the Rational Clinical Examination series [52]. Once established, each article included in the register will be indexed according to the associated level of evidence, the clinical domains, the methodological quality, as well as the clinical setting and patient populations. This will allow for easy navigation through the database to identify subject specific content and help determine the relevancy of a particular CPR.

Furthermore, we recognise that the filters will need to be periodically updated. For example, removing articles containing genetic search terms through the exclusion filter may be counter-productive in the future, if CPRs relevant to primary care begin to incorporate genetic measures as diagnostic tools. This highlights the evolving nature of search filters, and is consistent with previous research which indicates the need to update filters, for example due to changes that occur in MEDLINE indexing [51].

The novel search filters provided mixed results when compared with the previous filters produced by Haynes and colleagues. However, the performance of each of these filters was considered only in the context of finding CPRs relevant to primary care. Future work might seek to validate the novel strategies presented here, for example, by applying the filters to a wider range of clinical settings or alternatively smaller, subject-specific areas.

Section 4.5 Clinical relevance

The current work will help establish a publicly available international register of CPRs. With the increasing computerisation of medical practices, clinicians will be able to access the contents of the register during consultation, allowing on-site access to up-to-date research.

Nevertheless, CPRs are not designed to replace clinical skills and experience and should only be viewed as flexible tools to assist in clinical decision-making. Indeed, CPRs have been criticised for being incomplete in terms of variables included in certain rules, the range of clinical domains covered and incorporating patient and/or physician preferences [53]. Furthermore, application of CPRs to clinical practice should be evidence based. It is recommended that all rules go through the steps of validation and impact analyses before being implemented in routine daily practice [6, 54]. Yet reviews of the literature indicate that relative to the number of derived CPRs, few rules have been validated and fewer still have resulted in impact analysis [3, 6]. As a result, the impact of many CPRs on physician behaviour, patient outcome and/or cost remains largely unknown.

Part of the problem is that conducting independent studies in each clinical area is time consuming, resource demanding and expensive. Indeed, restricting research to independent studies limits the generalisability of the rule in terms of time, setting and specific patient population [53, 54]. One proposed solution is to globally establish large multipurpose and standardised routine databases from daily care practices. International collaboration would then allow for easy updating and validation of CPRs [53]. Once established, the register can identify areas for further research.

5. Conclusion

Without systematic indexing of articles as CPRs in electronic databases such as MEDLINE, locating CPRs for primary care will be challenging. Until a register is fully developed and publicly available, the optimal search filter, in many ways, depends on the specific needs and amount of time available to the clinician/researcher. For example, if the aim is to obtain a quick overview of an area with minimum input, combinations of the novel approaches

offered by the McGrath/Murphy Narrow Filter and the inclusion and exclusion filters appear most efficient. For our current purpose, the aim is to use a highly sensitive search filter to maximise retrieval of relevant CPRs, despite the trade-off for specificity, precision and accuracy. As such, the Haynes Narrow Filter combined with the Teljeur/Murphy Inclusion Filter -26 items represents the most valuable option to establish and update an international register of CPRs relevant to primary care.

6. Acknowledgements

The authors would like to thank Dr Grainne Cousins, Dr Wai-Sun Chan, Dr Rose Galvin, Dr Tim Hinchey and Dr Lucy Hederman for their helpful comments during the preparation of this manuscript. We also acknowledge the help and assistance provided by the medical students Ciara Dillon, Fran O'Reilly and Robert Ohle for the manual search and review of the identified articles for this study. This work was supported by the Health Research Board (HRB) of Ireland through the HRB Centre for Primary Care Research under Grant HRC/2007/1.

7. Conflict of Interest Statement

No competing interests have been declared.

8. References

[1] Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. J Am Med Inform Assoc 2001;8:391-7.

[2] Holland JL, Wilczynski NL, Haynes RB and The Hedges Team. Optimal search strategies for identifying sound clinical prediction studies in EMBASE. BMC Med Inform Decis Mak 2005;5:11.

[3] 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. JAMA 2000;284:79-84.

[4] Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowall I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med 1992;21:384-390.

[5] Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Med Decis Mak 1981;1:239-46.

[6] Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med 2006;144:201-209.

[7] Rothwell PM. Prognostic Models. Pract Neurol 2008;8:242-253.

[8] Wong SS.-L., Wilczynski NL, Haynes RB, Ramkissoonsingh R. Developing optimal search strategies for detecting sound clinical prediction studies in MEDLINE. AMIA Annu Symp Proc 2003;728-732.

[9] The Young Infants Clinical Signs Working Group. Clinical signs that predict severe illness in children under age 2 months: a muticentre study. Lancet 2008;371:135-42.

[10] Birnbaum A, Esses D, Bijur P, Wollowitz A, Gallagher EJ. Failure to validate the San Francisco Syncope Rule in an independent emergency department population. Ann Emerg Med 2008;52:151-9

[11] Bont J, Hak E, Hoes AW, Macfarlane JT, Verheij TJM. Predicting death in elderly patients with community-acquired pneumonia: a prospective validation study reevaluating the CRB-65 severity assessment tool. Arch Intern Med 2008;168:1465-8.

[12] Boyd M, Koziol-McLain J, Yates K, Kerse N, McLean C, Pilcher C et al. Emergency department case-finding for high-risk older adults: the Brief Risk Identification for Geriatric Health Tool (BRIGHT). Acad Emerg Med 2008;15:598-606.

[13] Briggs A, Spencer M, Wang H, Mannino D, Sin D. Development and validation of a prognostic index for health outcomes in chronic obstructive pulmonary disease. Arch Intern Med 2008;168:71-9.

[14] Bruyninckx R, Aertgeerts B, Bruyninckx P, Buntinx F. Signs and symptoms in diagnosing acute myocardial infarction and acute coronary syndrome: a diagnostic meta-analysis. Br J Gen Pract 2008; DOI: 10.3399/bjgp08X277014.

[15] Cameron AJ, Magliano DJ, Zimmet PZ, Welborn TA, Colagiuris S, Tonkin AM et al. The metabolic syndrome as a tool for predicting future diabetes: the AusDiab study. J Intern Med 2008;264:177-86.

[16] Donnan PT, Dorwald DWT, Mutch B, Morris AD. Development and validation of a model for predicting emergency admissions over the next year (PEONY): a UK historical cohort study. Arch Intern Med 2008;168:1416-22.

[17] Ebell MH. Diagnosis of appendicitis: part 1. History and physical examination. Am Fam Physician 2008;77:828-30.

[18] Ford AC, Talley NJ, Veldhuyzen van Zanten SJO, Vakil NB, Simel DL, Moayyedi P. Will the history and physical examination help establish that irritable bowel syndrome is causing this patient's lower gastrointestinal tract symptoms? JAMA 2008;300:1793-1805.

[19] Fowkes G and The Ankle Brachial Index Collaboration. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. JAMA 2008;300:197-208.

[20] Henschke N, Maher CG, Rofshauge KM. A systematic review identifies five "red flags" to screen for vertebral fracture in patients with low back pain. J Clin Epidemiol 2008;61:110-118.

[21] Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Shiekh A et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008;336:a332.

[22] Hsieh M, Auble TE, Tealy DM. Validation of the Acute Heart Failure Index. Ann Emerg Med 2008;51:37-44.

[23] Inouye SK, Zhang Y, Jones RN, Shi P, Cupples LA, Calderon HN et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. Med Care 2008;46:726-31.

[24] Inzitari M, Pozzi C, Ferrucci L, Chiarantini D, Rinaldi LA, Baccini M et al. Subtle neurological abnormalities as risk factors for cognitive and functional decline, cerebrovascular events, and mortality in older community-dwelling adults. Arch Intern Med 2008; 168:1270-6.

[25] Kastelein M, Wagemakers HPA, Luijsterburg PAJ, Verhaar JAN, Koes BW, Bierma-Zeinstra SMA. Assessing medical collateral ligament knee lesions in general practice. Am J Med 2008;121:982-88.

[26] Klok FA, Mos ICM, Nijkeuter M, Righini M, Perrier A, Le Gal G et al. Simplification of the revised Geneva score for assessing clinical probability of pulmonary embolism. Arch Intern Med 2008;168:2131-6.

[27] Kriston L, Hölzel L, Weiser AK, Berner MM, Härter M. Meta-analysis: are 3 questions enough to detect unhealthy alcohol use? Ann Intern Med 2008; 149:89-88.

[28] Kshirsagar AV, Bang H, Bomback AS, Vupputuri S, Shoham DA, Keen LM et al. A simple algorithm to predict incident kidney disease. Arch Intern Med 2008;168:2466-73.

[29] Liu H, Palge NM, Goldzweig CL, Wong E, Zhou A, Suttorp MJ et al. Screening for osteoporosis in men: a systematic review for an American College of Physicians guideline. Ann Intern Med 2008; 148:685-701.

[30] Madhok V, Falk G, Rogers A, Struthers AD, Sullivan FM, Fahey T. The accuracy of symptoms, signs and diagnostic tests in the diagnosis of left ventricular dysfunction in primary care: a diagnostic accuracy systematic review. BMC Fam Pract 2008;9:56.

[31] McGinn T, O'Connor-Moore N, Alfandre D, Gardenier D, Wisnivesky J. Validation of a hepatitis C screening tool in primary care. Arch Intern Med 2008,168:2009-13.

[32] Medbø A, Melbye H. What role may symptoms play in the diagnosis of airflow limitation? A study in an elderly population. Scand J Prim Health Care 2008;26:92-8.

[33] Parikh N, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D et al. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. Ann Intern Med 2008;148:102-110.

[34] Pezzotti P, Scalmana S, Mastromattei A, Di Lallo D. The accuracy of the MMSE in detecting cognitive impairment when administered by general practitioners: a prospective observational study. BMC Fam Pract 2008;9:29.

[35] Quinn J, McDermott D, Kramer N, Yeh C, Kohn MA, Stiell I et al. Death after emergency department visits for syncope: how common and can it be predicted? Ann Emerg Med 2008;51:585-590.

[36] Rahman M, Simmons RK, Harding AH, Wareham NJ, Griffin SJ. A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study. Fam Pract 2008;25:191-6.

[37] Richman PB, Simmons RK, Harding AH, Wareham NJ, Griffin SJ. Independent evaluation of an out-of hospital termination of resuscitation (TOR) clinical decision rule. Acad Emerg Med 2008;15:517-21.

[38] Rodger MA, Kahn SR, Well PS, Anderson DA, Chagnon I, Le Gal G et al. Identifying unprovoked thromboembolism patients at low risk for recurrence who can discontinue anticoagulant therapy. CMAJ 2008;179:417-26.

[39] Silvis ML, Clinch CR, Tillet JS. Clinical inquiries. What is the best way to evaluate an acute traumatic knee injury? J Fam Pract 2008;57:116-8.

[40] Simmons RK, Sharp S, Boekholdt M, Sargeant LA, Khaw KT, Warehams NJ et al. Evaluation of the Framingham risk score in the European Prospective Investigation of

Cancer-Norfolk cohort: does adding glycated haemoglobin improve the prediction of coronary heart disease events? Arch Intern Med 2008;168:1209-16.

[41] Suigoka T, Hayashino Y, Konno S, Kikuchi S, Fukuhara S. Predictive value of self-reported patient information for the identification of lumbar spinal stenosis. Fam Pract 2008;25:237-44.

[42] Tazakarji B, Modi S, Lee TJ, Cable K. Clinical inquiries. When should you admit a patient with suspected CAP? J Fam Pract 2008;57:195-7.

[43] Toll DB, Oudega R, Vergouwe Y, Moons KGM, Hoes AW. A new diagnostic rule for deep vein thrombosis: safety and efficiency in clinically relevant subgroups. Fam Pract 2008:doi:10.1093/fampra/cmm075

[44] van Voorhees BW, Paunesku D, Gollan J, Kuwabara S, Reinecke M, Basu A. Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA). Ann Fam Med 2008;6:503-11.

[45] Venmans LMAJ, Bont J, Gorter KJ, Verheij TJM, Rutten GEHM, Hak E. Prediction of complicated lower respiratory tract infections in older patients with diabetes. Br J Gen Pract, 2008;58:564-8.

[46] Yang X, Yee So W, Tong PCY, Ma RCW, Kong APS, Lam CWK. Development and validation of an all-cause mortality risk score in type 2 diabetes. Arch Intern Med 2008;168:451-7.

[47] Young J, De Sutter A, Merenstein D, van Essen GA, Kaiser L, Varonen H et al. Antibiotics for adults with clinically diagnosed acute rhinosinusitis: a meta-analysis of individual patient data. Lancet 2008;371:908-14.

[48] Zehtabchi S, Wright B. Does this emergency department patient with headache require neuroimaging? Ann Emerg Med 2008;51:324-7.

[49] Zethelius B, Berglund L, Sundström J, Ingelsson E, Basu S, Larsson A et al. Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. N Engl J Med 2008;358:2107-16.

[50] Yao X, Wilczynski NL, Walter SD, Haynes RB. Sample size determination for bibliographic retrieval studies. BMC Med Inform Decis Mak 2008;8:43

[51] Haynes RB, Wilczynski N, McKibbon A, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994;1:447-58.

[52] Simel DL, Rennie D, editors. *The rational clinical examination: evidence-based clinical diagnosis*. U.S.A.: McGraw-Hill Medical; 2009.

[53] Knottnerus JA. Diagnostic prediction rules: Principles, requirements, and pitfalls. Prim Care 1995; 22:341-363

[54] Toll DB, Janssen KJM, Vergouwe Y, Moons KGM. Validation, updating and impact of clinical prediction rules: A review. J Clin Epidemiol 2008; 61: 1085-1094.

Table 1

Definitions of the statistical terms used to calculate the performance of each search filter

Statistical term	**Definition**
Sensitivity	The number of articles classified by the filter as relevant to CPRs in primary care from all those retrieved as such by the 'reference standard' manual search
Specificity	The number of articles classified as irrelevant by the filter from all those that were identified as irrelevant by the 'reference standard' manual search
Precision	The number of articles classified as relevant by both the electronic filter and 'reference standard' manual search (true positive) over the total number of articles classified as relevant by the electronic filter only (true positive and false positive)
Accuracy	The number of articles that were correctly classified by both the electronic filter and 'reference standard' manual search either as relevant or irrelevant (true positive and true negative) over the total number of searched articles

Table 2

Thirty journal titles selected for inclusion as relevant CPRs in primary care

Journal titles	
Academic Emergency Medicine	Family Medicine
American Family Physician	Family Practice
American Journal of Medicine	Journal of American Medical Association
Annals of Emergency Medicine	Journal of the American Board of Family Medicine
Annals of Family Medicine	Journal of Clinical Epidemiology
Annals of Internal Medicine	Journal of Family Practice
Annals of Medicine	Journal of Internal Medicine
Annual Review of Medicine	Lancet
Archives of Internal Medicine	Medical Care
BMC Family Practice	Medical Decision Making
British Medical Journal	Medicine
British Journal of General Practice	New England Journal of Medicine
Canadian Family Physician	Public Library of Science Medicine
Canadian Medical Association Journal	Primary Care
Cochrane Database Systematic Reviews	Scandinavian Journal of Primary Health Care

Table 3

Results from the diagnostic test accuracy for each electronic search filter compared to the reference standard search of 30 journal articles

Filter name	N articles retrieved	N CPRs retrieved	Sensitivity (95% CI) as %	Specificity (95% CI) as %	Precision (95% CI) as %	Accuracy (95% CI) as %
Haynes Broad Filter						
HBF	1251	31	76 (62-89)	81(80-82)	2 (2-3)	81 (80-82)
HBF AND TMIF-26	731	30	73(60-87)	89 (88-90)	4 (3-6)	89 (88-90)
HBF OR TMIF-26	3139	40	98 (93-102)	51 (50-52)	1 (1-2)	52 (50-52)
HBF AND TMIF-22	520	30	73 (60-87)	92 (92-93)	6 (4-8)	92 (91-93)
HBF OR TMIF-22	1548	35	85 (75-96)	76 (75-77)	2 (2-3)	76 (75-77)
HBF NOT TMEF	750	17	41 (26-57)	88 (88-89)	2 (1-3)	88 (87-89)
HBF AND TMIF-26 NOT TMEF	403	17	41(26-57)	94 (93-94)	4 (2-6)	94 (93-94)
HBF OR TMIF-26 NOT TMEF	1433	15	37 (22-51)	78 (76-79)	1 (1-2)	77 (76-78)
HBF AND TMIF-22 NOT TMEF	304	17	41 (26-57)	95 (95-96)	6 (3-8)	95 (95-96)
HBF OR TMIF-22 NOT TMEF	613	13	32 (17-46)	90 (90-91)	2 (1-3)	90 (89-91)
Haynes Narrow Filter						
HNF	89	12	29 (15-43)	99 (99-99)	13 (6-21)	98 (98-99)
HNF AND TMIF-26	68	11	27 (13-40)	99 (99-99)	16 (7-25)	99 (98-99)
HNF OR TMIF-26	2594	40	98 (93-102)	59 (58-61)	2 (1-2)	60 (59-61)
HNF AND TMIF-22	61	12	29 (15-43)	99 (99-99)	20 (10-30)	99 (99-99)
HNF OR TMIF-22	783	34	83 (71-94)	88 (97-90)	4 (3-6)	88 (97-89)
HNF NOT TMEF	58	8	20 (7-32)	99 (99-99)	14 (5-23)	99 (98-99)
HNF AND TMIF-26 NOT TMEF	41	7	17 (6-29)	99 (99-100)	17 (6-29)	99 (99-99)
HNF OR TMIF-26 NOT TMEF	1252	15	37 (22-51)	80 (79-81)	1 (1-2)	80 (79-81)
HNF AND TMIF-22 NOT TMEF	42	8	20 (7-32)	99 (99-100)	19 (7-31)	99 (99-99)
HNF OR TMIF-22 NOT TMEF	41	13	32 (17-46)	95 (95-96)	4 (2-6)	95 (94-95)
McGrath/Murphy Broad Filter						
MMBF	264	23	56 (41-72)	96 (96-97)	9 (5-12)	96 (95-96)
MMBF AND TMIF-26	95	21	51 (36-67)	99 (99-99)	22 (14-30)	99 (98-99)
MMBF OR TMIF-26	735	36	88 (78-98)	89 (88-90)	5 (3-6)	89 (88-90)

MMBF AND TMIF-22	56	16	39 (24-54)	99 (99-100)	29 (17-40)	99 (99-99)
MMBF OR TMIF-22	1028	36	88 (78-98)	84 (83-85)	4 (2-5)	84 (83-85)
MMBF NOT TMEF	231	23	56 (41-71)	97 (96-97)	10 (6-14)	96 (96-97)
MMBF AND TMIF-26 NOT TMEF	89	21	51 (36-67)	99 (99-99)	24 (15-32)	99 (98-99)
MMBF OR TMIF-26 NOT TMEF	663	36	88 (78-98)	90 (89-91)	5 (4-7)	90 (89-91)
MMBF AND TMIF-22 NOT TMEF	50	16	39 (24-54)	99 (99-100)	32 (19-45)	99 (99-99)
MMBF OR TMIF-22 NOT TMEF	871	36	88 (78-98)	87 (86-88)	4 (3-5)	87 (86-88)
McGrath/Murphy Narrow Filter						
MMNF	63	16	39 (24-54)	99 (99-99)	25 (15-36)	99 (99-99)
MMNF AND TMIF-26	41	15	37 (22-51)	100 (99-100)	37 (22-51)	99 (99-99)
MMNF OR TMIF-26	597	34	83 (71-94)	91 (90-92)	6 (4-8)	91 (90-92)
MMNF AND TMIF-22	16	5	12 (2-22)	100 (100-100)	31 (9-54)	99 (99-99)
MMNF OR TMIF-22	939	35	85 (75-96)	86 (85-87)	4 (3-5)	86 (85-87)
MMNF NOT TMEF	57	16	39 (22-54)	99 (99-100)	28 (16-40)	99 (99-99)
MMNF AND TMIF-26 NOT TMEF	40	15	37 (22-51)	100 (99-100)	38 (23-53)	99 (99-99)
MMNF OR TMIF-26 NOT TMEF	543	34	83 (71-94)	92 (91-93)	6 (4-8)	92 (91-93)
MMNF AND TMIF-22 NOT TMEF	15	5	12 (2-22)	100 (100-100)	33 (9-57)	99 (99-99)
MMNF OR TMIF-22 NOT TMEF	793	35	85 (75-96)	88 (87-89)	4 (3-6)	88 (87-89)
Teljeur/Murphy Inclusion and Exclusion Filters (PubMed)						
TMIF-26	2432	39	95 (89-102)	62 (61-63)	2 (1-2)	62 (61-63)
TMIF-22	693	34	83 (71-94)	90 (89-90)	5 (3-7)	90 (89-90)
TMEF*	3589	24	59 (43-74)	43 (42-45)	1 (0.4-1)	44 (42-45)
TMIF-26 NOT TMEF	1241	23	56 (41-71)	81 (80-82)	2 (1-3)	81 (80-81)
TMIF-22 NOT TMEF	407	21	51 (36-67)	94 (93-94)	5 (3-7)	94 (93-94)

Note. The manual reference standard search resulted in a total number of 6344 articles being searched, 41 of which were CPRs relevant to primary care.* The TMEF was also executed via Ebsco host. N articles retrieved = 7464.

HBF = Haynes Broad Filter; HNF = Haynes Narrow Filter; MMBF = McGrath/Murphy Broad Filter; MMNF = McGrath/Murphy Narrow Filter; TMIF-26 = Teljeur/Murphy Inclusion Filter 26 item; TMIF-22 = Teljeur/Murphy Inclusion Filter 22 item; TMEF = Teljeur Murphy Exclusion Filter

AND, OR and NOT are Boolean search terms. AND retrieves articles that include all the search terms. OR retrieves articles that include at least one of the search terms. NOT excludes the retrieval of articles that contain the search terms.

Figure 1

Manual search ‘reference standard’ set of CPRs relevant to primary care for year 2008.

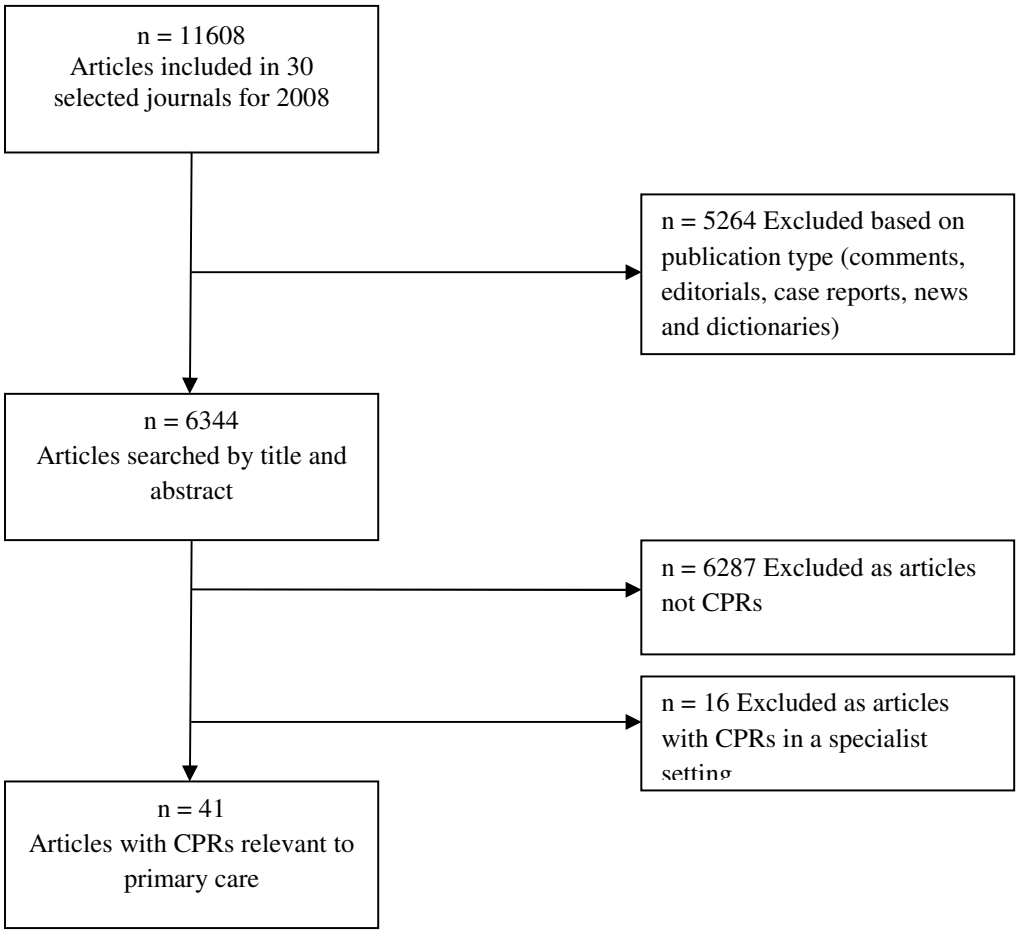
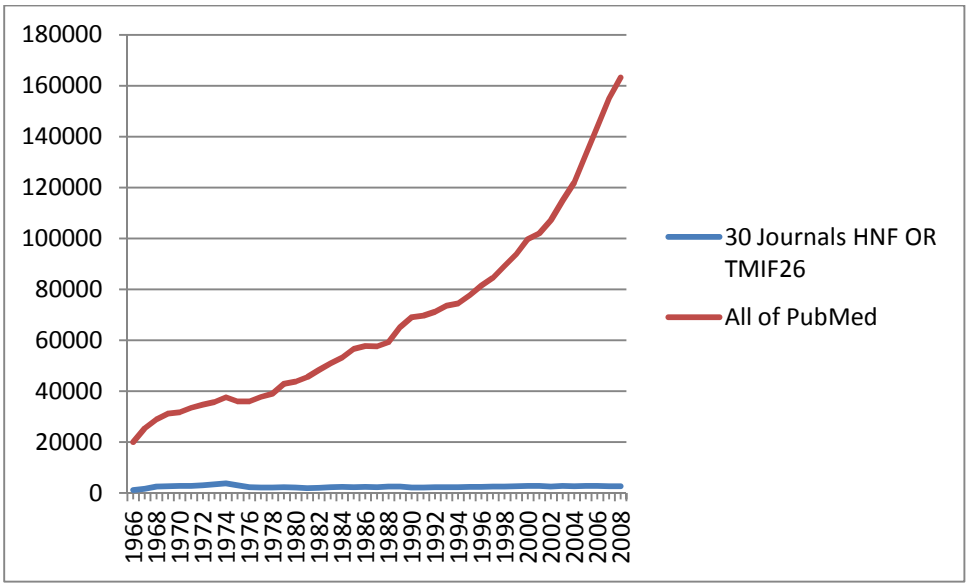


Figure 2

Total number of articles retrieved by the Haynes Narrow Filter combined with the Teljeur/Murphy Inclusion Filter 26 item (HNF OR TMIF-26) versus total number of articles published in PubMed from 1966-2008



APPENDICES

Appendix 1

Search strings for each of the electronic search strategies for the 30 journal titles

Database	Filter name	Filter search string
PubMed	Haynes Broad Filter (HBF)	(predict*[tiab] OR predictive value of tests[mh] OR scor*[tiab] OR observ*[tiab] OR observer variation[mh])
PubMed	Haynes Narrow Filter (HNF)	(validation[tiab] OR validate[tiab])
EBSCO host	McGrath/Murphy Broad Filter (MMBF)	((predict* N3 rule* OR predict* N3 model OR predict* N3 models) OR (decision* N3 rule*) OR (TX validat*))
EBSCO host	McGrath/Murphy Narrow Filter (MMNF)	((predict* N3 rule* OR predict* N3 model OR predict* N3 models) OR (decision* N3 rule*))
PubMed	Teljeur/Murphy Inclusion Filter 26 item (TMIF-26)	"clinical prediction" OR "clinical model*" OR "clinical score*" OR "decision rule*" OR "diagnostic accuracy" OR "diagnostic rule*" OR "diagnostic score*" OR "diagnostic value" OR "predictive outcome*" OR "predictive rule*" OR "predictive score*" OR "predictive value" OR "predictive risk*" OR "prediction outcome*" OR "prediction rule*" OR "prediction score*" OR "prediction value*" OR "prediction risk*" OR "risk assessment" OR "risk score*" OR "validation decision*" OR "validation rule*" OR "validation score*" OR (derivation AND validation) OR (sensitivity AND specificity) OR (symptoms AND signs)
PubMed	Teljeur/Murphy Inclusion Filter 22 item (TMIF-22)	(clinical[tiab] AND predict*[tiab]) OR (clinical[tiab] AND model*[tiab]) OR (clinical[tiab] AND score*[tiab]) OR (decision [tiab] AND rule*[tiab]) OR (derive*[tiab] AND validat*[tiab]) OR (diagnos*[tiab] AND accura*[tiab]) OR (diagnos*[tiab] AND rule*[tiab]) OR (diagnos*[tiab] AND score*[tiab]) OR (diagnos*[tiab] AND value[tiab]) OR (predict*[tiab] AND outcome*[tiab]) OR (predict*[tiab] AND rule*[tiab] OR (predict*[tiab] AND score*[tiab]) OR (predict*[tiab] AND validat*[tiab]) OR (predict*[tiab] AND value*[tiab]) OR (risk*[tiab] AND assessment*[tiab]) OR (risk[tiab] AND score*[tiab]) OR (sensitivity[tiab] AND specificity[tiab]) OR (symptoms[tiab] AND signs[tiab]) OR (validat*[tiab] AND decision*[tiab]) OR (validat*[tiab] AND rule*[tiab]) OR (validat*[tiab] AND score*[tiab]) OR (predict*[tiab] AND risk*[tiab])
PubMed	Teljeur/Murphy Exclusion Filter (TMEF)	(allele OR amino OR animal OR apoptosis OR chromosome OR congenital OR dental OR dna OR endogenous OR endothelial OR epithelial OR mammalian OR mice OR molecule OR molecular OR mouse OR mutate OR mutation OR necrosis OR pathogenesis OR phosphorylation OR polymorphism OR receptor OR signal OR species OR tissue OR tumor OR tumour OR tyrosine OR vitro)

Note: For each filter the following limits were applied: (1) Articles were limited to humans; and (2) irrelevant publication types were excluded (case reports, comments, dictionaries, editorials, and news).

Appendix 2

Details of training provided to reviewers

Prior to screening articles for each of the electronic filters, all reviewers were required to attend a training session. Each reviewer formed part of a pair that comprised of a clinician and a non-clinical researcher. Clear and precise definitions of (1) clinical prediction rules (CPRs) and (2) primary care were provided. Reviewers were encouraged to voice any uncertainties about either definition. Each reviewer was then provided with an identical EndNote file. Each file contained a specially selected set of 100 journal articles from the year 2007, 5 of which were known to be CPRs. All articles were required to be classified as following: (1) not a CPR; (2) definitely a CPR relevant to primary care; or (3) unsure (either unsure the article was a CPR or unsure of relevance to primary care). If reviewers were certain that an article was a CPR but not relevant to primary care, it was also placed in the unsure category. Results indicated good agreement between reviewers for the 'sure' category, with all reviewers identifying at least 4 of the 5 CPRs. Reviewers were required to make their search as sensitive as possible, and as such were encouraged to include articles in the 'unsure' category for later discussion. Results indicated that those with the least experience at classifying CPRs placed more articles in the 'unsure' category, relative to those with the most experience in this regard. All articles placed in this category were discussed openly between the group and a consensus was reached.

Appendix 3

Articles identified as clinical prediction rules during the manual search (reference standard) search of the 30 journals for the year 2008

	First author	Article title	Journal title
1	Young infants clinical signs study group [9]	Clinical signs that predict severe illness in children under age 2 months: a muticentre study	Lancet
2	Birnbaum [10]	Failure to validate the San Francisco Syncope Rule in an independent emergency department population	Annals of Emergency Medicine
3	Bont [11]	Predicting death in elderly patients with community-acquired pneumonia: a prospective validation study reevaluating the CRB-65 severity assessment tool	Archives of Internal Medicine
4	Boyd [12]	Emergency department case-finding for high-risk older adults: the Brief Risk Identification for Geriatric Health Tool (BRIGHT)	Academic Emergency Medicine
5	Briggs [13]	Development and validation of a prognostic index for health outcomes in chronic obstructive pulmonary disease	Archives of Internal Medicine
6	Bruyninckx [14]	Signs and symptoms in diagnosing acute myocardial infarction and acute coronary syndrome: a diagnostic meta-analysis	British Journal of General Practice
7	Cameron [15]	The metabolic syndrome as a tool for predicting future diabetes: the AusDiab study	Journal of Internal Medicine
8	Donnan [16]	Development and validation of a model for predicting emergency admissions over the next year (PEONY): a UK historical cohort study	Archives of Internal Medicine
9	Ebell [17]	Diagnosis of appendicitis: part 1. History and physical examination	American Family Physician
10	Ford [18]	Will the history and physical examination help establish that irritable bowel syndrome is causing this patient's lower gastrointestinal tract symptoms?	Journal of American Medical Association
11	Fowkes [19]	Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis	Journal of American Medical Association
12	Henschke [20]	A systematic review identifies five "red flags" to screen for vertebral fracture in patients with low back pain	Journal of Clinical Epidemiology
13	Hippisley-Cox [21]	Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2	British Medical Journal

14	Hsieh [22]	Validation of the Acute Heart Failure Index	Annals of Emergency Medicine
15	Inouye [23]	Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model	Medical Care
16	Inzitari [24]	Subtle neurological abnormalities as risk factors for cognitive and functional decline, cerebrovascular events, and mortality in older community-dwelling adults	Archives of Internal Medicine
17	Kastelein [25]	Assessing medical collateral ligament knee lesions in general practice	American Journal of Medicine
18	Klok [26]	Simplification of the revised Geneva score for assessing clinical probability of pulmonary embolism	Archives of Internal Medicine
19	Kriston [27]	Meta-analysis: are 3 questions enough to detect unhealthy alcohol use?	Annals of Internal Medicine
20	Kshirsagar [28]	A simple algorithm to predict incident kidney disease	Archives of Internal Medicine
21	Liu [29]	Screening for osteoporosis in men: a systematic review for an American College of Physicians guideline	Annals of Internal Medicine
22	Madhok [30]	The accuracy of symptoms, signs and diagnostic tests in the diagnosis of left ventricular dysfunction in primary care: a diagnostic accuracy systematic review	BMC Family Practice
23	McGinn [31]	Validation of a hepatitis C screening tool in primary care	Archives of Internal Medicine
24	Medbø [32]	What role may symptoms play in the diagnosis of airflow limitation? A study in an elderly population	Scandinavian Journal of Primary Health Care
25	Parikh [33]	A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study	Annals of Internal Medicine
26	Pezzotti [34]	The accuracy of the MMSE in detecting cognitive impairment when administered by general practitioners: a prospective observational study	BMC Family Practice
27	Quinn [35]	Death after emergency department visits for syncope: how common and can it be predicted?	Annals of Emergency Medicine
28	Rahman [36]	A simple risk score identifies individuals at high risk of developing Type 2 diabetes: a prospective cohort study	Family Practice
29	Richman [37]	Independent evaluation of an out-of hospital termination of resuscitation (TOR) clinical decision rule	Academic Emergency Medicine
30	Rodger [38]	Identifying unprovoked thromboembolism patients at low risk for recurrence who can	Canadian Medical Association Journal

		discontinue anticoagulant therapy	
31	Silvis [39]	Clinical inquiries. What is the best way to evaluate an acute traumatic knee injury?	Journal of Family Practice
32	Simmons [40]	Evaluation of the Framingham risk score in the European Prospective Investigation of Cancer-Norfolk cohort: does adding glycated haemoglobin improve the prediction of coronary heart disease events?	Archives of Internal Medicine
33	Sugioka [41]	Predictive value of self-reported patient information for the identification of lumbar spinal stenosis	Family Practice
34	Tazakarji [42]	Clinical inquiries. When should you admit a patient with suspected CAP?	Journal of Family Practice
35	Toll [43]	A new diagnostic rule for deep vein thrombosis: safety and efficiency in clinically relevant subgroups	Family Practice
36	van Voorhees [44]	Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA)	Annals of Family Medicine
37	Venmans [45]	Prediction of complicated lower respiratory tract infections in older patients with diabetes	British Journal of General Practice
38	Yang [46]	Development and validation of an all-cause mortality risk score in type 2 diabetes	Archives of Internal Medicine
39	Young [47]	Antibiotics for adults with clinically diagnosed acute rhinosinusitis: a meta-analysis of individual patient data	Lancet
40	Zehtabchi [48]	Does this emergency department patient with headache require neuroimaging?	Annals of Emergency Medicine
41	Zethelius [50]	Use of multiple biomarkers to improve the prediction of death from cardiovascular causes	New England Journal of Medicine