Risk stratification in Dutch primary care: a promising approach to manage population health

Girwar, S.M.

CHAPTER 2

CHAPTER 2

A SYSTEMATIC REVIEW OF RISK STRATIFICATION TOOLS INTERNATIONALLY USED IN PRIMARY CARE SETTINGS

Shelley-Ann M. Girwar, Robert Jabroer, Marta Fiocco, Stephen P. Sutch, Mattijs E. Numans, Marc A. Bruijnzeels

Published in Health Sciences Reports. 2021 Jul 23;4(3):e329. doi: 10.1002/hsr2.329

Abstract

Introduction

In our current healthcare situation, burden on healthcare services is increasing, with higher costs and increased utilization. Structured Population Health Management has developed as an approach to balance quality with increasing costs. This approach identifies sub-populations with comparable health risks, to tailor interventions for those that will benefit the most. Worldwide, the use of routine healthcare data extracted from electronic health registries for risk stratification approaches is increasing. Different risk stratification tools are used on different levels of the healthcare continuum. In this systematic literature review, we aimed to explore which tools are used in primary healthcare settings and assess their performance.

Methods

We performed a systematic literature review of studies applying risk stratification tools with health outcomes in primary care populations. Studies in OECD countries published in English language journals were included. Search engines were utilised with keywords e.g. 'primary care', 'risk stratification' and 'model'. Risk stratification tools were compared based on different measures: Area Under the Curve (AUC) and C-statistics for dichotomous outcomes and R2 for continuous outcomes.

Results

The search provided 4718 articles. Specific election criteria such as primary care populations, generic health utilization outcomes, and routinely collected data sources identified 61 articles, reporting on 31 different models. The three most frequently applied models were the Adjusted Clinical Groups (ACG, n=23), the Charlson Comorbidity Index (CCI, n=19) and the Hierarchical Condition Categories (HCC, n=7). Most AUC and C-statistic values above 0.7, with ACG showing slightly improved scores compared to the CCI and HCC (typically between 0.6 and 0.7).

Conclusion

Based on statistical performance, the validity of the ACG was the highest, followed by the CCI and the HCC. The ACG also appeared to be the most flexible, with the use of different international coding systems and measuring a wider variety of health outcomes.

Introduction

For several decades healthcare costs have been rising. This has been attributed to ageing populations and innovative ways of curing and treating diseases, leading to an increased prevalence of chronic illnesses and comorbidities among community dwelling older people (1). Also patients have increased demands regarding increasing choice around the way their healthcare should be organized and have tended to utilize more care. Furthermore, the needs for healthcare are not evenly distributed within populations. In Western countries, the sickest 5% of the population make up for 50% of the total healthcare costs (2). In order to maintain high quality healthcare, resources should be distributed according to the needs of the population instead of the demand. One way of dealing with this is to allocate resources according to the individual care needs in subpopulations. Predicting healthcare utilization and health outcomes based on needs provides opportunities to allocate resources more appropriately. Predictions of health outcomes through risk stratification can be used to tailor proactive clinical care, to install preventive measures, to restructure healthcare and to improve insight for healthcare professionals. In the long run this approach will help improve the quality of care and reduce the costs (3,4).

A way to monitor and predict costly patient outcomes such as hospitalization, high care utilization and emergency department visits, is through the use of structured population health management programs. Population Health Management is an approach that aims to improve the health of a defined group of people and to strive for more equitable distribution of health outcomes within the group. In Population Health Management programs, an important step is to stratify individuals within a specific subpopulation according to the risk of experiencing an adverse event, such as defined undesirable health outcomes or the extent of their healthcare utilization. Stratification analyses are often performed based on the use of routinely collected healthcare data. Typically, the high-risk sub-population usually comprises of a small percentage of the total population. The medium-and low-risk subpopulations are much larger with around 35% of the overall population classified as medium-risk and 60% as low-risk (2). The identification of people classified on their respective risk-estimates is referred to as risk stratification. Preceding risk stratification population segmentation is performed. Segmentation can be performed based on general characteristics such as age, gender and specific diseases, but also on morbidity and healthcare utilization patterns. A discussion of segmentation was outside the scope of this study.

Many methods for risk stratification exist internationally. Current literature regarding risk stratification models prominently focus on stratifying hospital populations, based on readily available hospital data. However, primary care data has a great potential to improve healthcare quality and reduce health costs (5). Especially in countries where primary care registries have nearly 100% coverage of the total population, such as the Netherlands and the United Kingdom (UK), the opportunity arises to assess the whole population by using these routinely collected primary care data. Distribution of risk in a primary care population is different from a hospital or specialized care population. Current literature also mainly focuses on risk stratification models with disease specific outcomes, whereas in this study the focus is on more generic utilization outcomes such as risk on hospitalization, emergency department visits, future high healthcare utilization and high pharmaceutical expenditures.

The aim of this study was to perform a systematic literature review to describe and assess the performance of different risk stratification tools with generic health utilization outcomes using routinely collected data, and with possibilities of application to the European context, such as in Dutch primary care. Based on the description of the performance of the tools, we recommend the risk stratification tool best suited for usage in Dutch primary care.

Methods

The PRISMA statements regarding conduction and reporting systematic literature reviews were followed throughout the literature review process (6).

This review was conducted through searches in the search engines Pubmed and Embase. The search-string which contained both keywords and MeSH terms is shown in supplementary table 1. The most important keywords were ‘primary care’, ‘risk stratification’ and ‘model’. EndNote X8.2 was used as the reference manager for the articles. The search-string was produced in collaboration with the Leiden University Medical Center (LUMC) Walaeus library.

The PRISMA flow diagram displays the numbers of included and excluded articles (figure 1).

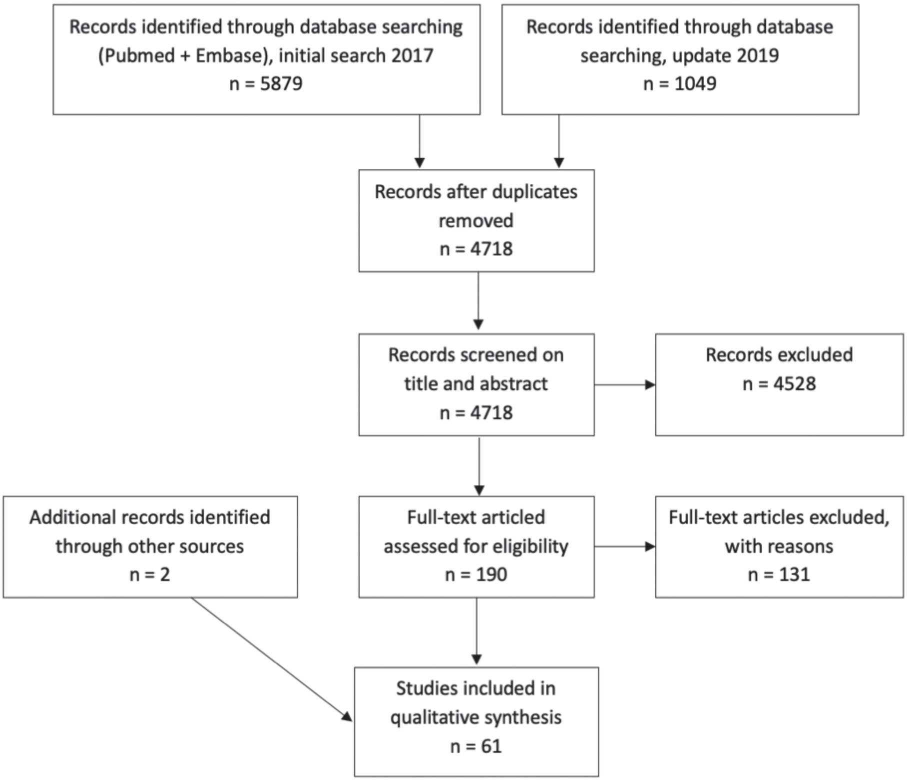

Figure 1: PRISMA flow chart displaying numbers of included and excluded articles

Inclusion criteria

The search characteristics are specified by the **P**opulation, **I**ntervention, **C**ontrol and Outcome (PICO) method. In our research, the *population* is the primary care population. Therefore, we only included articles where models applied to primary care populations are discussed. The *interventions* investigated were the risk stratification approaches and models that are applied to primary care data. *Outcomes* investigated are risks of hospitalization, high healthcare costs, emergency department visits, high pharmaceutical drug expenditure, mortality and other generic health utilization outcomes.

For comparability with a Western-societal environment such as the Dutch situation, only studies performed in countries listed with the Organisation for Economic Co-operation and Development (OECD) (7), were included. Only freely accessible articles in the English language were considered eligible. Articles from January 2007 till August 2019 were reviewed. The inclusion criteria narrowed the search down to a context which was more applicable in a European primary care situation with a gatekeeper's role, such as the Dutch primary care system.

Exclusion criteria

Articles that used risk stratification tools on populations consisting of hospitalized patients, or patients seeking consultation with a specialist (*e.g.* an oncologist or cardiologist) were excluded. These patients were not considered to represent those in a primary care setting. In addition, research looking at specific disease outcome was also excluded, as this review aims at exploring general population outcomes. Articles not freely accessible were excluded as well as articles that were not available in English.

The initial search, conducted in December 2017, yielded 5879 articles. In September 2019, an update of the search was conducted, resulting in an additional 1049 articles. After removing duplicates according to the manual of the Free University (VU) library (8), 4718 articles remained. Articles were screened on both title and abstract, based on the criteria mentioned earlier. 78% of the screening, based on title and abstract, was performed by two researchers independently (R.J. & S.G.). Their results were compared, and in the case of disagreement (2%), the articles were discussed until consensus was achieved. The main causes for disagreement concerned indistinct and misunderstood study populations and model outcomes. As the percentage of disagreement was low, the remaining 22% of the titles and abstracts where only screened by one researcher. After screening on title and abstract, 190 articles remained to be screened on their full text. Screening of all 190 full papers was performed by the same two researchers independently and results were compared. Again, in case of disagreement (21%), the article was discussed until consensus was achieved. After exclusion of 131 articles, including 17 titles which were either not freely accessible or where no English versions of the full papers were available, 59 studies remained to be included in this review. Two further articles were added through the snowball method, resulting in 61 articles.

Assessing performance of models

The different models were compared on three aspects: frequency of use, statistical diagnostic validity, and performance in primary care.

For each identified risk stratification model, **the frequency of use** of the model was presented, taking into account all included studies.

For the assessment of the **statistical diagnostic validity**, reviewed studies were divided into *application, validation* and *comparison* studies. In the *application* studies, risk stratification tools were applied for purposes other than assessing their statistic diagnostic validity. Therefore, *application* studies did not present any statistical diagnostic measures of the risk stratification tools. In the *validation* studies and in most of the *comparison* studies, statistical diagnostic measures of the applied risk stratification tools were provided. Area Under the Curve (AUC) and C-statistics for models with dichotomous outcomes and R^2 values for models with continues outcomes were used to validate risk stratification tools. Models with AUC or C-statistic values between 0.5 and 0.6 were classified as performing *poorly*, values between 0.6 and 0.7 were considered *sufficient* and values above 0.7 were considered *good* (9). Ten of the reviewed papers, the *comparison* studies, compared more than one risk stratification tool in the same study population with the same record data, enabling a more appropriate comparison between risk stratification tools. Most of the comparison studies presented statistical diagnostic values, as they are mostly also validation studies.

For **performance in primary care**, we assessed the type of routinely collected data that is used as input of the model. Models using input data available in Dutch primary care health records were assumed to have a good potential performance in Dutch primary care.

Results

A total of 31 risk stratification models were identified in the literature. The three most frequently applied tools, taking into account all included studies, concern the Adjusted Clinical Groups (ACG), the Charlson Comorbidity Index (CCI) and the Hierarchical Condition Categories (HCC). These three main risk stratification tools are presented in table 1, with predicted outcomes and diagnostic values. Assessment of these tools, their diagnostic validity and applicability in primary care are described in order. The remaining 28 risk stratification tools can be found in supplementary table 2.

Table 1: Overview of the three most frequently identified risk stratification models with their characteristics and diagnostic properties for different outcomes.

	First author, year	Adjusted Clinical Group (ACG)	Charlson Comorbidity Index (CCI)	Hierarchical Condition Categories (HCC)
Categories		ACG-categories (1-93), Resource Utilization Bands (RUBs), Expended Diagnosis Clusters (EDC) count	Six categories based on chronic condition count	Score based on aggregated conditions (70 categories)
Total number of studies in which the model was applied		**n=23**	**n=19**	**n=7**
Diagnostic properties for different outcomes:				
Hospitalization	Haas, 2013,[4]	C=0.73	C=0.68	C=0.67
	Lemke, 2012,[12]	AUC=0.80	AUC=0.78	
(number of hospitalizations)	Shadmi, 2011,[16]	R^2=0.24	R^2=0.11	
(unplanned hospitalization)	Maltenfort, 2019,[11]	AUC=0.82		
	Inouye, 2008,[20]		C=0.72	
	Ou, 2011,[21]		C=0.61	
	Mosley, 2009,[25]			AUC=0.64
Emergency department visits	Haas, 2013,[4]	C=0.67	C=0.59	C=0.58
	Ou, 2011,[21]		C=0.63	
	Wallace, 2016,[22]		C=0.58	
Costs				
(Top 10% total costs)	Haas, 2013,[4]	C=0.76	C=0.70	C=0.70
(total costs)	Brilleman, 2014,[14]	R^2=0.41	R^2=0.34	
(pharmaceutical costs)	Aguado, 2008,[10]	R^2=0.39		
(total costs)	Sicras-Mainar, 2013,[13]	R^2=0.37		
(total costs)	Charlson, 2008,[18]		R^2=0.22	
(total costs)	Charlson, 2014,[19]		R^2=0.20	
(high total costs)	Ou, 2011,[21]		C=0.64	

Table 1: Continued

	First author, year	Adjusted Clinical Group (ACG)	Charlson Comorbidity Index (CCI)	Hierarchical Condition Categories (HCC)
Utilization of different healthcare services				
(GP visits)	Brilleman, 2013,[15]	R^2=0. 37	R^2=0.26	
(primary care visits)	Shadmi, 2011,[16]	R^2=0.54	R^2=0.18	
(specialist visits)	Shadmi, 2011,[16]	R^2=0.45	R^2=0.13	
(number of diagnostic imaging tests)	Shadmi, 2011,[16]	R^2=0.37	R^2=0.15	
(visits)	Sicras-Mainar, 2013,[13]	R^2=0.42		
(number of diagnoses / reasons for visit)	Sicras-Mainar, 2013,[13]	R^2=0.77		
(high outpatient visits)	Ou, 2011,[21]		C=0.63	
Input data for the model		Age, gender, diagnostic codes, pharmaceutical information, healthcare costs	Presence or absence of chronic conditions based on diagnosis codes; weighted	ICD-9 of ICD-10 diagnosis codes

AUC= Area Under the ROC Curve; C= C-statistic; R2=R square

In the grey fields diagnostic values according to the comparison studies are represented. The dark grey fields present values from comparison studies in which all three most frequently used risk stratification models are compared with each other. The studies associated with the light grey concern comparison studies comparing only two out of the three main models.

Adjusted Clinical Groups: 23 studies

The ACG is the most frequently applied risk stratification tool in our review. The ACG system is a risk stratification model designed by the Johns Hopkins University. The model was originally developed to predict and measure multimorbidity in a population. The ACG system is a measure of comorbidity and can predict utilization costs, hospitalization and emergency department visits. The model is able to use patients' data from Electronic Health Records (EHRs), insurance claims, disease registries and health status surveys(10). Minimal input data for the model are healthcare diagnoses in a specific time interval, gender and age, to which the ACG classifies people to one of 93 ACG categories. These categories represent expected healthcare utilization. In addition, different probabilities for future utilization of healthcare services are calculated. This information can be used by healthcare professionals to make informed clinical and administrative decisions (4).

Of the 23 ACG studies, eight provided statistical diagnostic values for the accuracy of the model, calculated for different outcomes. For ***prediction of hospitalization***, the model is diagnostically assessed three times with AUC and C statistic values between 0.73 and 0.82 (4,11,12). The diagnostic accuracy can be classified as *good*.

In one study a C-value of 0.67 is presented for ***prediction of emergency department visitation***, which classifies as *Sufficient*, and a C-value of 0.76 for ***prediction of high total costs***, again classifying as *good* (4). Three other studies presented R^2 values between 0.37 and 0.41 for explaining the variation of healthcare ***costs*** by the ACG model (10,13,14). ***Variations in high utilization of different healthcare services***, such as primary care visits, specialists' visits and numbers of diagnostic imaging tests, diagnoses and hospitalizations, are discussed in three studies, with R^2 values ranging from 0.24 to 0.77 (13,15,16).

ACG is highly suitable for application in primary care populations, as using *International Classification of Primary Care* (ICPC) codes as input is possible (10). ICPC codes are used to classify complaints and diagnoses of patients in many primary care settings, such as in the Netherlands. This information is stored in EHRs. The model uses other input variables such as age, gender, pharmaceutical information and previous visitation, stored in the EHR as well.

Charlson Comorbidity Index: 19 studies

The CCI is the second-most studied risk stratification model. The CCI was developed by Charlson and colleagues in 1987 and was originally an age-comorbidity index that predicted a relative risk of death within a year for hospital admitted cancer patients (17). Since that time, many adjustments have been made and in addition to mortality predictions the model is now used to predict hospitalization, emergency department visitation, future healthcare utilization and morbidity in wider populations. The system categorizes the population into six categories, based on the presence of comorbidities and chronic conditions, of which a weighted sum is provided (from zero conditions as category one to five or more conditions as category six) (18,19). The model investigates the effect of multimorbidity and predicts several outcomes. Variations of the CCI exist and the validity on predictions have been consistently investigated (4).

From the 18 studies in which the CCI or a modification was used, 10 provided statistical diagnostic values. AUC and C-values range from 0.61 to 0.78 for the prediction of future hospitalization (4,12,20,21), which correspond to an accuracy of Sufficient and Good. For emergency department visitation C-statistics between 0.58 and 0.63 are provided (4,21,22) (poor to sufficient) and for total costs, R^2 values were between 0.20 and 0.34 (14,18,19). For healthcare utilization of different healthcare services R^2 values were between 0.13 and 0.26 (15,16,23).

Input variables for the CCI include combinations of age, race, gender, mental illness, pregnancy, drug or alcohol addiction, type of health plan, type of provider, number of therapeutic classes and number of medications prescribed. The CCI is fit for use with primary care data, but focuses primarily on the absence or presence of chronic conditions, apart from other demographics. Although there is no evidence in the included studies of use of the CCI with ICPC codes, the coding system used in Dutch primary care, there is evidence for use with Read codes, a British primary care coding system (24). Possibilities to use the model with coding systems other than International Classifications of Disease (ICD) codes, are therefore very likely.

The software algorithm for CCI is published and available (4).

Hierarchical Condition Categories: 7 studies

The third most frequently studied model (n=7) is the HCC. This model was first designed and implemented by the Centers for Medicare and Medicaid Services (CMS) to adjust capitation payments for enrolees with higher risk than others. The model uses demographic data of patients as well as ICD 10th revision (ICD-10) diagnosis codes. ICD codes are used in all American healthcare service providers

(25). The ICD classification is adapted in other countries, yet these are codes most prominently used in hospital administrative registries (26). Based on this information, the model categorizes a patient into one of 70 aggregated condition categories which contributes to an individualized risk score.

For this model, four diagnostic values are provided in two studies included in this literature review. For hospitalization an AUC value of 0.64 (25), and a C-statistic of 0.67 (4), are provided. The study by Haas et al. provides a C-statistic equal to 0.58 for prediction of emergency department visitation, but a much higher C-statistic of 0.70 for prediction of high total costs (4).

A major concern regarding this model, is that it makes use of ICD codes rather than ICPC codes, making it difficult to apply in the Dutch primary care settings.

Comparison studies

A total of ten papers compared more than one risk stratification tool applied within the same study populations. However, only five articles compared more than one of the three above mentioned risk stratification tools while providing statistical diagnostic values to compare the different tools with each other.

For ***hospitalization*** the ACG performs slightly better than the CCI with AUC values of 0.80 versus 0.78 (12), and C-statistics of 0.73 versus 0.68 (4). The ACG also outperforms the CCI regarding ***emergency department visitation*** with C-statistics of 0.67 versus 0.59 and ***high total costs*** with C-statistics of 0.76 versus 0.70 (4), and R^2 values of 0.41 versus 0.34 (14). Furthermore, the study by Shadmi and colleagues showed evidence of the ACG providing better results compared to the CCI regarding other ***healthcare utilization*** outcomes, such as numbers of hospitalizations (R^2 = 0.24 versus R^2 = 0.11), primary care visits (R^2 = 0.54 versus R^2 = 0.18), specialist visits (R^2 = 0.45 versus R^2 = 0.13) and diagnostic imaging tests (R^2 = 0.37 versus R^2 = 0.15) all within a study period of 12 months (16). In addition, Brilleman and colleagues find R^2 values of 0.37 for the ACG and 0.26 for the CCI with the number of general practitioner (GP) visits as the predicted outcome (15).

Remaining risk stratification tools

In addition to the three above mentioned risk stratification tools, 28 other tools were identified within this systematic literature review. One of the 28 identified risk stratification tools is called the Elixhauser Index or Method and was mentioned in five studies. The Elixhauser Index uses a set of 30 dichotomous variables as comorbidity measures (27). Outcomes concern high utilization and pharmaceutical expenditure. One out of the five studies, mentioning the Elixhauser Index, provided C-statistics between 0.62 and 0.74 for different health utilization outcomes (21). The study by Ou and colleagues compared those C-statistics to values between 0.61 and 0.64 for the CCI (21).

A number of the identified risk stratification tools include disease or medication counts as comorbidity measures, such as the Chronic Disease Score (CDS) (n=3), which is based on dispensed drugs history. The previously mentioned study by Ou and colleagues provided C-statistic values between 0.61 and 0.72 for the CDS (21). The remainder of the identified risk stratification tools were only mentioned a few times (n=1, 2 or 3) in the articles, typically including only one validation study per risk stratification tool. The infrequent use of these tools does not make a review possible. The Clinical Risk Groups (CRG), for example, emerged three times within our systematic review. However, all studies using the CRG as a risk stratification tool were application studies and thus lacking statistical diagnostic values. Most other studies, describe a new risk stratification tool developed for a specific situation. In supplementary table 2 all of the risk stratification tools are presented, organized by included studies.

Discussion

Summary of main findings

This literature review revealed a broad range of risk stratification tools that have been assessed on accuracy and validity. The most common predicted outcomes were future hospitalization, emergency department visitation, high healthcare utilization, and total cost. The three most frequently studied risk stratification tools were the ACG, CCI and HCC.

With most AUC and C-statistic values above 0.70, the ACG performs *good* on a wide variety of outcomes. The CCI scores *sufficient* for different outcomes, with the exception of high utilization of healthcare for which a low score yielded. With most AUC and C-statistic values between 0.60 and 0.70 the HCC can also be classified as *sufficient*. Comparing the results of the ACG, the CCI and HCC, more convincing evidence for accuracy and validity is found for the ACG. Previous research also indicated the high accuracy and validation of the ACG model (12,28-30). The model is considered one of the leading models regarding the accuracy of predicting hospitalizations (12), and is widely used to gain insight in future healthcare utilization of patients (31). The study by Ou and colleagues is making a compelling case for the validity of the Elixhauser Index and the CDS, compared to the CCI (21). However, this result is not robust as it is only based on a single study. Nevertheless, the Elixhauser Index may have future potential for use in a European primary care setting.

For the applicability in primary care, evidence shows that the ACG has the possibility to make use of ICPC codes, the coding system of the (Dutch) primary care registry. The CCI has not yet been proven usable with ICPC codes. Nevertheless, evidence has shown possibilities for the CCI to be used with Read codes (24), the UK's primary care coding system, making it highly likely that the CCI can be applied using other than ICD diagnosis codes. For the HCC model on the other hand, there is no evidence to use diagnosis codes other than the ICD coding system, making it difficult to use this model in Dutch primary care.

The results of this study support the idea that risk stratification tools are suitable for primary care data in a European context. However different models emphasize various aspects within the tools. As all applications are focusing on similar utilization outcomes, such as hospitalization, ED visits and costs, the ACG has an array of other indicators developed for risk stratification. Various applications in primary care show the potential of models for example in areas of improved resource allocation (32), and care management due to better insights into vulnerable populations (33). In addition, the ACG provides possibilities to efficiently prioritize sub-populations for tailored care interventions (34).

Limitations

Although our results support risk stratification using the ACG in primary care, there are some limitations.

We only selected studies that already performed risk stratification in primary care. As a consequence we could have missed stratification tools only applied in hospital or open source data, but with a strong potential for suitability in primary care.

Selection of studies was dependant on the inter observer reliability of the two researchers. Although inclusion and exclusion criteria were clearly formulated beforehand, the possibility remains that useful tools were missed given the relatively high number of disagreements.

We assessed the identified risk stratification tools in different studies, in an attempt to compare the statistical validity of the models with each other. However, the incomparable circumstances under which different studies are performed, such as study populations and data sources, make reasonable comparisons challenging.

We based our recommendation on diagnostic values of applied risk stratification tool reported by studies published in scientific literature. Due to publication bias promising risk stratification tools may not have emerged sufficiently from our findings.

Further research

From all the articles included in this study, a small percentage explicitly defines 'risk stratification'. With the growing need for tailored care and health management approaches, a precise definition will be useful. Risk stratification and other terms such as population segmentation are now used interchangeably. Studies contributing to a generalized definition of the term *risk stratification* will be of great scientific and practical value. By using the same definition, miscommunications regarding the meaning of risk stratification will be reduced, and information on highly performing methods and implementations thereof can be shared more effectively.

With this review, we studied which risk stratification tools are best suited for the European primary care setting. However, primary care settings differ between countries. To find the best suitable tool for a specific primary care system, the performance of different tools should be investigated within the same setting, centred on desired outcomes. Based on the results of this literature review, further studies assessing the performance of desired risk stratification models, will be beneficial for Dutch primary care.

Conclusion

In conclusion, based on application frequency, statistical validity and used diagnosis coding systems, we suggest the ACG as the best model for use in European primary care settings, such as Dutch Primary Care. However, further local assessment of the ACG system is needed to ensure proper implementation.

References

1. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. Journal of general internal medicine 2007;22 Suppl 3:391-5.
2. Luo G, Stone BL, Sakaguchi F, et al. Using Computational Approaches to Improve Risk-Stratified Patient Management: Rationale and Methods. JMIR research protocols 2015;4(4):e128.
3. Bernstein RH. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. J Ambul Care Manage 2007;30(1):39-51.
4. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. The American journal of managed care 2013;19(9):725-32.
5. Gentil ML, Cuggia M, Fiquet L, et al. Factors influencing the development of primary care data collection projects from electronic health records: a systematic review of the literature. BMC medical informatics and decision making 2017;17(1):139.
6. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine 2009;6(7):e1000097.
7. Organisation for Economic Co-operation and Development. Member Countries. http://www.oecd.org/about/members-and-partners/. Accessed March, 2020.
8. Ket H. Tips & trucs voor het uitvoeren van systematische reviews met EndNote.
9. Simundic AM. Measures of Diagnostic Accuracy: Basic Definitions. Ejifcc 2009;19(4):203-11.
10. Aguado A, Guino E, Mukherjee B, et al. Variability in prescription drug expenditures explained by adjusted clinical groups (ACG) case-mix: a cross-sectional study of patient electronic records in primary care. BMC health services research 2008;8:53.
11. Maltenfort MG, Chen Y, Forrest CB. Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins Adjusted Clinical Groups risk adjustment system. PloS one 2019;14(8):e0221233.
12. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. Medical care 2012;50(2):131-9
13. Sicras-Mainar A, Velasco-Velasco S, Navarro-Artieda R, et al. Obtaining the mean relative weights of the cost of care in Catalonia (Spain): retrospective application of the adjusted clinical groups case-mix system in primary health care. Journal of evaluation in clinical practice 2013;19(2):267-76.
14. Brilleman SL, Gravelle H, Hollinghurst S, et al. Keep it simple? Predicting primary health care costs with clinical morbidity measures. Journal of health economics 2014;35:109-22.
15. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. Family practice 2013;30(2):172-8.
16. Shadmi E, Balicer RD, Kinder K, et al. Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. BMC public health 2011;11:609.
17. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-83.
18. Charlson ME, Charlson RE, Peterson JC, et al. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. Journal of clinical epidemiology 2008;61(12):1234-40.

19. Charlson ME, Wells MT, Kanna B, et al. Medicaid managed care: how to target efforts to reduce costs. BMC health services research 2014;14:461.
20. Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. Medical care 2008;46(7):726-31.
21. Ou HT, Mukherjee B, Erickson SR, et al. Comparative performance of comorbidity indices in discriminating health-related behaviors and outcomes. Health Outcomes Research in Medicine 2011;2(2):e91-e104.
22. Wallace E, McDowell R, Bennett K, et al. Comparison of count-based multimorbidity measures in predicting emergency admission and functional decline in older community-dwelling adults: a prospective cohort study. BMJ open 2016;6(9):e013089.
23. Ou HT, Mukherjee B, Erickson SR, et al. Comparative performance of comorbidity indices in predicting health care-related behaviors and outcomes among Medicaid enrollees with type 2 diabetes. Population health management 2012;15(4):220-9.
24. Khan NF, Perera R, Harper S, et al. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC family practice 2010;11:1.
25. Mosley DG, Peterson E, Martin DC. Do hierarchical condition category model scores predict hospitalization risk in newly enrolled Medicare advantage participants as well as probability of repeated admission scores? Journal of the American Geriatrics Society 2009;57(12):2306-10.
26. (RIVM) RvVeM. Internationale statistische classificatie van ziekten en met gezondheid verband houdende problemen. Tiende Revisie. 2015
27. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. Medical care 1998;36(1):8-27.
28. Dominick KL, Dudley TK, Coffman CJ, et al. Comparison of three comorbidity measures for predicting health service use in patients with osteoarthritis. Arthritis and rheumatism 2005;53(5):666-72.
29. Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. Journal of clinical epidemiology 2004;57(10):1040-8.
30. Huntley AL, Johnson R, Purdy S, et al. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. Annals of family medicine 2012;10(2):134-41.
31. Johns Hopkins University. https://www.hopkinsacg.org/. Accessed March, 2020.
32. Kristensen T, Rose Olsen K, Sortso C, et al. Resources allocation and health care needs in diabetes care in Danish GP clinics. Health policy (Amsterdam, Netherlands) 2013;113(1-2):206-15.
33. Burton LC, Skinner EA, Uscher-Pines L, et al. Health of Medicare Advantage plan enrollees at 1 year after Hurricane Katrina. The American journal of managed care 2009;15(1):13-22.
34. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health services research 2019;54(2):466-73.
35. Akazawa M, Imai H, Igarashi A, et al. Potentially inappropriate medication use in elderly Japanese patients. The American journal of geriatric pharmacotherapy 2010;8(2):146-60.

36. Beauchet O, Launay CP, Chabot J, et al. Prediction of unplanned hospital admissions in older community dwellers using the 6-item Brief Geriatric Assessment: Results from REPERAGE, an observational prospective population-based cohort study. Maturitas 2019;122:1-7.
37. Chang HY, Richards TM, Shermock KM, et al. Evaluating the Impact of Prescription Fill Rates on Risk Stratification Model Performance. Medical care 2017;55(12):1052-60.
38. Chung S, Romanelli RJ, Stults CD, et al. Preventive visit among older adults with Medicare's introduction of Annual Wellness Visit: Closing gaps in underutilization. Preventive medicine 2018;115:110-18.
39. Crane SJ, Tung EE, Hanson GJ, et al. Use of an electronic administrative database to identify older community dwelling adults at high-risk for hospitalization or emergency department visits: the elders risk assessment index. BMC health services research 2010;10:338.
40. Davis AC, Shen E, Shah NR, et al. Segmentation of High-Cost Adults in an Integrated Healthcare System Based on Empirical Clustering of Acute and Chronic Conditions. Journal of general internal medicine 2018;33(12):2171-79.
41. Dennis S, Taggart J, Yu H, et al. Linking observational data from general practice, hospital admissions and diabetes clinic databases: can it be used to predict hospital admission? BMC health services research 2019;19(1):526.
42. Duenas-Espin I, Vela E, Pauws S, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. BMJ open 2016;6(4):e010301.
43. Freund T, Kunz CU, Ose D, et al. Patterns of multimorbidity in primary care patients at high risk of future hospitalization. Population health management 2012;15(2):119-24.
44. Glazier RH, Agha MM, Moineddin R, et al. Universal health insurance and equity in primary care and specialist office visits: a population-based study. Annals of family medicine 2009;7(5):396-405.
45. Hamano J, Oishi A, Kizawa Y. Prevalence and Characteristics of Patients Being at Risk of Deteriorating and Dying in Primary Care. Journal of pain and symptom management 2019;57(2):266-72.e1.
46. Hewner S, Seo JY, Gothard SE, et al. Aligning population-based care management with chronic disease complexity. Nursing outlook 2014;62(4):250-8.
47. Hong CS, Atlas SJ, Ashburner JM, et al. Evaluating a Model to Predict Primary Care Physician-Defined Complexity in a Large Academic Primary Care Practice-Based Research Network. Journal of general internal medicine 2015;30(12):1741-7.
48. Hu T, Dattani ND, Cox KA, et al. Effect of comorbidities and medications on frequency of primary care visits among older patients. Canadian family physician Medecin de famille canadien 2017;63(1):45-50.
49. Hutchings HA, Evans BA, Fitzsimmons D, et al. Predictive risk stratification model: a progressive cluster-randomised trial in chronic conditions management (PRISMATIC) research protocol. Trials 2013;14:301.
50. Khanna S, Rolls DA, Boyle J, et al. A risk stratification tool for hospitalisation in Australia using primary care data. Scientific reports 2019;9(1):5011.
51. Martin S, Wagner J, Lupulescu-Mann N, et al. Comparison of EHR-based diagnosis documentation locations to a gold standard for risk stratification in patients with multiple chronic conditions. Applied clinical informatics 2017;8(3):794-809.

52. Martin Lesende I, Mendibil Crespo LI, Castano Manzanares S, et al. Functional decline and associated factors in patients with multimorbidity at 8 months of follow-up in primary care: The functionality in pluripathological patients (FUNCIPLUR) longitudinal descriptive study. BMJ open 2018;8 (7) (no pagination)(e022377)
53. Metcalfe D, Masters J, Delmestri A, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. BMC medical research methodology 2019;19(1):115.
54. Milla-Perseguer M, Guadalajara-Olmeda N, Vivas-Consuelo D, et al. Measurement of health-related quality by multimorbidity groups in primary health care. Health and quality of life outcomes 2019;17(1):8.
55. Moran WP, Zhang J, Gebregziabher M, et al. Chaos to complexity: leveling the playing field for measuring value in primary care. Journal of evaluation in clinical practice 2017;23(2):430-38.
56. Muratov S, Lee J, Holbrook A, et al. Unplanned index hospital admissions among new older high-cost health care users in Ontario: a population-based matched cohort study. CMAJ open 2019;7(3):E537-e45.
57. Noyes K, Liu H, Temkin-Greener H. Medicare capitation model, functional status, and multiple comorbidities: model accuracy. The American journal of managed care 2008;14(10):679-90.
58. Ranstad K, Midlov P, Halling A. Active listing and more consultations in primary care are associated with shorter mean hospitalisation and interacting with psychiatric disorders when adjusting for multimorbidity, age and sex. Scandinavian journal of primary health care 2018;36(3):308-16.
59. Rohrer JE, Rasmussen N, Adamson SA. Illness severity and total visits in family medicine. Journal of evaluation in clinical practice 2008;14(1):65-69.
60. Salisbury C, Johnson L, Purdy S, et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. The British journal of general practice : the journal of the Royal College of General Practitioners 2011;61(582):e12-21.
61. Sibley LM, Moineddin R, Agha MM, et al. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. Medical care 2010;48(2):175-82.
62. Sibley LM, Glazier RH. Evaluation of the equity of age-sex adjusted primary care capitation payments in Ontario, Canada. Health policy (Amsterdam, Netherlands) 2012;104(2):186-92.
63. Sicras-Mainar A, Serrat-Tarres J, Navarro-Artieda R, et al. Adjusted Clinical Groups use as a measure of the referrals efficiency from primary care to specialized in Spain. European journal of public health 2007;17(6):657-63.
64. Sicras-Mainar A, Velasco-Velasco S, Navarro-Artieda R, et al. Adaptive capacity of the Adjusted Clinical Groups Case-Mix System to the cost of primary healthcare in Catalonia (Spain): a observational study. BMJ open 2012;2(3).
65. Sino CG, Stuffken R, Heerdink ER, et al. The association between prescription change frequency, chronic disease score and hospital admissions: a case control study. BMC pharmacology & toxicology 2013;14:39.
66. Snooks H, Bailey-Jones K, Burge-Jones D, et al. Effects and costs of implementing predictive risk stratification in primary care: a randomised stepped wedge trial. BMJ quality & safety 2019;28(9):697-705.

67. Sternberg SA, Bentur N, Abrams C, et al. Identifying frail older people using predictive modeling. The American journal of managed care 2012;18(10):e392-7.
68. Takahashi PY, Ryu E, Olson JE, et al. Hospitalizations and emergency department use in Mayo Clinic Biobank participants within the employee and community health medical home. Mayo Clinic proceedings 2013;88(9):963-9.
69. Vest JR, Menachemi N, Grannis SJ, et al. Impact of Risk Stratification on Referrals and Uptake of Wraparound Services That Address Social Determinants: A Stepped Wedged Trial. American journal of preventive medicine 2019;56(4):e125-e33.
70. Violan C, Plana-Ripoll O, Foguet-Boreu Q, et al. Relationship between efficiency and clinical effectiveness indicators in an adjusted model of resource consumption: a cross-sectional study. BMC health services research 2013;13:421.
71. Vivas-Consuelo D, Uso-Talamantes R, Trillo-Mata JL, et al. Predictability of pharmaceutical spending in primary health services using Clinical Risk Groups. Health policy (Amsterdam, Netherlands) 2014;116(2-3):188-95.
72. Vuik SI, Mayer E, Darzi A. Enhancing risk stratification for use in integrated care: a cluster analysis of high-risk patients in a retrospective cohort study. BMJ open 2016;6(12):e012903.
73. Wennberg JE, Staiger DO, Sharp SM, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. BMJ (Clinical research ed) 2013;346:f549.
74. Xu J, Williams-Livingston A, Gaglioti A, et al. A Practical Risk Stratification Approach for Implementing a Primary Care Chronic Disease Management Program in an Underserved Community. Journal of health care for the poor and underserved 2018;29(1):202-13.
75. Zhou YY, Wong W, Li H. Improving care for older adults: a model to segment the senior population. The Permanente journal 2014;18(3):18-21.

Supplementary Material

Supplementary table 1: Search strategy

((((”risk”[tw] OR ”risks”[tw] OR ”Risk”[Mesh]) **AND** (”stratification”[tw] OR ”stratifications”[tw] OR ”stratified”[tw] OR ”stratify”[tw] OR ”stratifies”[tw] OR ”stratifying”[tw])) **AND** (”model”[tw] OR ”models”[tw] OR ”tool”[tw] OR ”tools”[tw] OR ”method”[tw] OR ”methods”[tw] OR ”measure”[tw] OR ”measures”[tw] OR ”measured”[tw] OR ”measuring”[tw] OR ”algorithm”[tw] OR ”algorithms”[tw] OR ”metric”[tw] OR ”metrics”[tw] OR ”score”[tw] OR ”scores”[tw] OR ”scoring”[tw] OR ”index”[tw] OR ”indexes”[tw] OR ”indices”[tw] OR ”indexed”[tw] OR ”count”[tw] OR ”counts”[tw])) OR ”adjusted clinical groups”[tw] OR ”Minnesota tiering”[tw] OR ”Hierarchical Condition Categories”[tw] OR ”elder risk assessment index”[tw] OR ”chronic condition count”[tw] OR ”Charlson comorbidity index”[tw] OR ”chronic disease score”[tw]) **AND** (”General Practitioners”[Mesh] OR ”General Practitioner”[tw] OR ”General Practitioners”[tw] OR ”physician”[tw] OR ”physicians”[tw] OR ”general practice”[tw] OR ”General Practice”[Mesh] OR ”GP”[ti] OR ”GPs”[ti] OR ”GP’s”[ti] OR ”Primary Health Care”[Mesh] OR ”Primary Health Care”[tw] OR ”primary care”[tw] OR ”Family doctor”[tw] OR ”family doctors”[tw] OR ”Family practice”[tw] OR ”Family practices”[tw] OR ”Family medicine”[tw] OR ”general medicine”[tw] OR ”Accountable Care Organizations”[Mesh] OR ”ACO”[ti] OR ”ACOs”[tw] OR ”ACO’s”[tw] OR ”care organization”[tw] OR ”care organisation”[tw] OR ”care organizations”[tw] OR ”care organisations”[tw] OR ”Ambulatory Care Facilities”[Mesh:NoExp] OR ”Community Health Centers”[Mesh:NoExp] OR ”health center”[tw] OR ”health centers”[tw] OR ”health centre”[tw] OR ”health centres”[tw] OR ”Health Maintenance Organizations”[Mesh] OR ”maintenance organization”[tw] OR ”maintenance organisation”[tw] OR ”maintenance organizations”[tw] OR ”maintenance organisations”[tw] OR ”HMO”[ti] OR ”HMOs”[ti] OR ”HMO’s”[ti] OR ”MCO”[ti] OR ”MCOs”[ti] OR ”MCO’s”[ti] OR ”managed care”[tw] OR ”integrated care”[tw])

Filters: Date 2007-2019, Language English

Supplementary table 2: Overview of included articles. Study-population, Outcome measure, Risk stratification model, Study type, Country and Journal of publication are shown.

Reference	Population	Outcome	Model	Study type	Country	Journal
(Aguado et al., 2008) [10]	Patients from five primary care centres n=65,630	Explained variability in drug expenditures	Adjusted Clinical Groups	Validation study	Spain	BMC Health Services Research
(Akazawa, Imai, Igarashi, & Tsutani, 2010) [35]	Adults aged 65 years old or older n=6,628	Incidence, healthcare utilization and costs associated with Potentially Inappropriate Medication Use	Elixhauser Classification System, Modified Beers criteria	Application study	Japan	The American Journal of Geriatric Pharmacotherapy
(Beauchet et al., 2019) [36]	Community dwelling adults aged 80 years or older, who visited a participating general practice within the study period n = 668	Hospital admissions (6 month follow-up)	6-item Brief Geriatric Assessment	Validation study	France	Maturitas
(Brilleman & Salisbury, 2013) [15]	Primary care population of general practice n=95,372	Three year mortality and consultation rate	Charlson Comorbidity Index, Adjusted Clinical Groups, Quality and Outcomes Framework	Validation and comparison study	UK	Oxford University Press
(Brilleman et al., 2014) [14]	Primary care population of general practice n=86,100	Primary health costs	Charlson Comorbidity Index, Adjusted Clinical Groups, Quality and Outcomes Framework	Validation and comparison study	UK	Journal of Health Economics
(Burton et al., 2009) [33]	Non-institutionalized Peoples Health (Managed Care Organization) managed care beneficiaries n=20,612	Morbidity level	Adjusted Clinical Groups	Application study	USA	The American Journal of Managed Care

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Chang et al., 2017) [37]	Primary care population of the Health Partners Network n=43,097	Effect of prescription fill rates on risk stratification model performance	Adjusted Clinical Groups	Comparison study	USA	Medical Care
(Charlson et al., 2008) [18]	Primary care population of a general practice in an academic hospital n=5,861	Predict costs of disease in primary care patients	Adapted Charlson Comorbidity Index	Validation study	USA	Journal of Clinical Epidemiology
(Charlson et al., 2014) [19]	Medicaid managed care beneficiaries, adults and children, who received primary care in a specific medical center, not including Medicare/ Medicaid patients (In this review presented results concern adult population) n = 4,614 (2,218 adults and 2,396 children)	Identification of potential high cost beneficiaries	Charlson Comorbidity Index	Application study	USA	BMC Health Services Research
(Chung, Romanelli, Stults, & Luft, 2018) [38]	Medicare beneficiaries, aged 65 to 85 years, who were primary care patients in a large, mixed payer outpatient healthcare organization n=108,734	Preventive primary care visits	Own model including Charlson Comorbidity Index as comorbidity measure, age category, visit frequency and insurance	Application study	USA	Preventive Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Crane et al., 2010) [39]	Older community dwelling adults of a general practice n=12,650	High-Risk Hospitalization or ED admission	The Elders Risk Assessment	Validation study	USA	BMC Health Services Research
(Davis et al., 2018) [40]	Patients from an integrated healthcare delivery system and health plan, with continuous coverage; n = 2,118,343	Health care costs	Hierarchical Condition Categories	Application study	USA	Journal of General Internal Medicine
(Dennis et al., 2019) [41]	Patients with type 2 diabetes who had been treated at a general practice n = 161,575	Hospital admission	Own prediction model	Validation study	Australia	BMC Health Services Research
(Duenas-Espin et al., 2016) [42]	Primary care populations n=2,100,000 (B) n=7,500,000 (C) n=100,000) (L) n=3,400,000 (Sc)	Health risk assessment	Adjusted Clinical Groups, Adjusted Morbidity Groups known as GMA, Clinical Risk Group	Application and comparison study	Spain, Italy, Scotland	BMJ Open
(Freund, Kunz, Ose, Szecsenyi, & Peters-Klimm, 2012) [43]	Primary care population n=6,026	Prediction risk future hospitalization	Hierarchical Condition Categories	Application study	Germany	Population Health Management
(Glazier, Agha, Moineddin, & Sibley, 2009) [44]	Primary care population n=25,558	Diagnosed health status	Adjusted Clinical Groups	Application study	Canada	Annals of Family Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Haas et al., 2013) [4]	Adult patients empaneled in 2009 and 2010 in a primary care practice n=83,187	Hospitalization, emergency department visits, 30-day readmission, high expenditures	Adjusted Clinical Group, Hierarchical Condition Categories, Elder Risk Assessment, Chronic Comorbidity Count, Charlson Comorbidity Index, Minnesota Tiering	Comparison study	USA	The American Journal of Managed Care
(Hamano, Oishi, & Kizawa, 2019) [45]	Primary care patients aged 65 years or older, who visited a participating general practice within the study period; n = 382	Detoriation and dead	Supportive and Palliative Care Indicators Tool	Validation study	Japan	Journal of Pain and Symptom Management
(Hewner, Seo, Gothard, & Johnson, 2014) [46]	Primary care population of Medicare, Medicaid and privately insured n=477,407	Risk-stratified cohorts based on chronic disease and complexity	COMPLEXedex clinical algorithm	Application study	USA	Nursing Outlook
(Hong et al., 2015) [47]	Primary care adult patients in a practice-based research network n=143,372	Prediction of complexity	Outpatient Charlson Score & Commercial Risk Prediction	Validation and comparison study	USA	Journal of General Internal Medicine
(Hu et al., 2017) [48]	Primary care population n=265	Predictors of frequent visits to family physicians	Charlson Comorbidity Index, Beers Criteria	Application study	Canada	Canadian Familiy Physician
(Huntley, Johnson, Purdy, Valderas, & Salisbury, 2012) [30]	Primary care population	Care utilization, costs, mortality, quality of life	Adjusted Clinical Groups, Charlson Comorbidity Index, Chronic Disease Score, Cumulative Illness Rating Scale, Duke Severity Index	Review	-	Annals of Family Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Hutchings et al., 2013) [49]	Primary care population n=2,400	Estimate effects on the delivery of care, patient satisfaction, quality of life and resources used.	An emergency admission risk prediction tool called PRISM	Validation study	Wales	Trial Journal
(Inouye et al., 2008) [20]	Community dwelling elderly aged 70 years and older in primary care clinics of an academic medical center n=3,919	Unplanned hospitalization in one year	Deyo-Charlson,	Validation study	USA	Medical Care
(Khan et al., 2010) [24]	Primary care cancer patients and healthy controls n=146,441	Mortality	Adapted Charlson Score for use with Read/ OXMIS instead of ICD10 diagnosis codes	Validation study	UK	BMC Family Practice
(Khanna et al., 2019) [50]	Patients who attended their primary care clinics at least once n = 393,229	1 year hospitalization, ED visit	Own model	Validation study	Australia	Scientific Reports
(Kristensen et al., 2013) [32]	Primary care patients with type 2 diabetes; n = 6,706	Fee for service costs	Adjusted Clinical Groups	Application study	Denmark	Health Policy
(Lemke et al., 2012) [12]	Primary care population n=4,700,000	Predicting hospitalization	Adjusted Clinical Groups, Charlson Comorbidity Index	Validation study	USA	Medical Care
(Maltenfort et al., 2019) [11]	Children seen in a large primary and specialty care outpatient network n= 920,051 (70% for training and 30% for testing the model)	Unplanned 30-day hospitalization	Own model with predictors derived from the Adjusted Clinical Groups	Validation study	USA	PLoS One

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Martin et al., 2017) [51]	Patients from three primary care clinics participating in the Integrated Care Coordination Information System study data set, who had at least one of the selected conditions and were seen from 2008 to 2012 n = 750	ED visits, hospitalization and heathcare costs	A modified Charlson Comorbidity Index, Hierarchical Condition Categories, count of chronic conditions defined by Affordable Care Act	Application study	USA	Applied Clinical Informatics
(Martin Lesende et al., 2018) [52]	Patients aged 65 years and older n = 241	Top 5% of the 'Kaiser Permanente pyramid'	Adjusted Clinical Groups	Application study	Spain	BMJ Open
(Metcalfe et al., 2019) [53]	Patients from primary care practices: age- and sex-matched controls (n=26,860) for hip fractured patients (n = 13,974)	Mortality: 30-day and 1-year	Charlson Comorbidity Index, Elixhauser method	Validation study	UK	BMC Medical Research Methodology
(Milla-Perseguer, Guadalajara Olmeda, Vivas-Consuelo, & Uso-Talamantes, 2019) [54]	All citizens registered in a specific health district n = 32,667	Morbidity measure	Clinical Risk Groups	Application study	Spain	Health and Quality of Life Outcomes
(Moran et al., 2017) [55]	Primary care population (Medical University) n=10,408	Clustering patients with stratification on risk for hospital and ED utilization	Own model	Validation study	USA	Journal of Evaluation in Clinical Practice
(Mosley et al., 2009) [25]	Primary care population, Medicare beneficiaries; n = 4,506	Hospitalization	Hierarchical Condition Categories	Application study	USA	Journal of the American Geriatrics Society

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Muratov et al., 2019) [56]	Adults aged 66 years old or more High cost users (n=175,847) and age and sex matched non-high cost users (n=527,541)	Hospital admission	Adjusted Clinical Groups	Validation study	Canada	CMAJ Open
(Noyes, Liu, & Temkin Greener, 2008) [57]	Community dwelling (not institutionalized for 90 days at a time) Medicare beneficiaries with continuous part A and B enrollment for at least two calender years n = 46,790	Healthcare Costs	Hierarchical Condition Categories	Validation study	USA	American Journal of Managed Care
(Ou et al., 2011) [21]	Medicaid enrolees with type 2 diabetes n=9,832	Health care behaviors (physician's diabetes .adherence standard adherence, patient's medication adherence), health care utilization and expenditures	A modified version of Romano-adapted Charlson index, Elixhauser index, Chronic Disease Score, Health-related Quality of Life Comorbidity Index	Validation and Comparison study	USA	Health Outcomes Research in Medicine
(Ou et al., 2012) [23]	Medicaid enrolees with type 2 diabetes n=9,832	Health care utilization and expenditures	A modified version of Romano-adapted Charlson index, Elixhauser index, Chronic Disease Score, Health-related Quality of Life Comorbidity Index	Validation and Comparison study	USA	Population Health Management

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Ranstad, Midlov, & Halling, 2018) [58]	Primary care listed patients n = 151,731	Hospitalization (binary) and number of days hospitalized	Own model with doctor-patient relationship, contribution of complex diagnosis patterns (psychiatric disorders) and morbidity burden (Adjusted Clinical Groups based)	Application study	Sweden	Scandinavian Journal of Primary Health Care
(Rohrer, Rasmussen, & Adamson, 2008) [59]	Primary care population n=698	High utilization, Illness severity	Charlson Comorbidity Index	Application study	USA	Journal of Evaluation in Clinical Practice
(Salisbury, Johnson, Purdy, Valderas, & Montgomery, 2011) [60]	Primary care population n=99,997	Multimorbidity	Adjusted Clinical Groups, Quality of Outcome Framework	Application study	UK	British Journal of General Practice
(Shadmi et al., 2011) [16]	Adult enrolees of Clalit Health Services, Israel's largest health care organization n = 279,241	Numbers of (1) primary care encounters, (2) specialist visits, (3) diagnostic imaging tests and (4) hospitalizations	Adjusted Clinical Groups, Charlson Comorbidity Index	Application, validation and comparison study	Israel	BMC Public Health
(Sibley, Moineddin, Agha, & Glazier, 2010) [61]	Primary care patients n=25,558	Predicting physician utilization	Adjusted Clinical Groups	Application study	Canada	Medical Care
(Sibley & Glazier, 2012) [62]	Primary care data (family health networks) n=487,131	Expected healthcare utilization	Adjusted Clinical Groups	Application study	Canada	Health Policy
(Sicras-Mainar et al., 2007) [63]	Patients attending five primary care teams n = 81,335	Referral Rate	Adjusted Clinical Groups	Application study	Spain	European Journal of Public Health

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Sicras-Mainar et al., 2012) [64]	Patients from 13 primary care teams (86,5% primary care, 13,5% peadiatrics) n = 227,235	Heatlh costs	Adjusted Clinical Groups	Application study	Spain	BMJ Open
(Sicras-Mainar et al., 2013) [13]	Patients from 13 primary care teams n = 227,235	Explaining variance of: Visits, number of diagnoses, total costs	Adjusted Clinical Groups	Validation Study	Spain	Journal of Evaluation in Clinical Practice
(Sino et al., 2013) [65]	Hospitalized (cases) and matched non-hospitalized patients from pharmacy registries n = 2 x 8,681 (cases and controls)	Prediction of hospital admission	Combining Prescription Changes Frequency, Chronic Disease Score	Application study	the Netherlands	BMC Pharmacology and Toxicology
(Snooks et al., 2019) [66]	Patients from primary care practices n = 230,099	Unscheduled hospital admission	An emergency admission risk prediction tool called PRISM	Application study	UK	BMJ Quality & Safety

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Soto-Gordoa et al., 2019) [34]	Patients with multimorbidity (at least two out of three conditions: diabetes, heart failure and chronic obstructive pulmonary disease) aged 65 years or more n = 4,225	Hospital admission	Adjusted Clinical Groups	Application study	Spain	Health Services Research
(Sternberg et al., 2012) [67]	Community dwelling elderly n=221	Predict hospitalizations and emergency department visitation	Adjusted Clinical Groups, Vulnerable Elderly Survey	Application study	Israel	American Journal of Managed Care
(Takahashi et al., 2013) [68]	Primary care population ECH Mayo clinic biobank n=22,916	Predict hospitalizations and emergency department visitation	Minnesota Tiering	Application study	USA	Mayo Clinic Proceedings
(Vest et al., 2019) [69]	Patients from primary care clinics linked to a public hospital: intervention group using risk stratification modelling (n=62,254) versus control group n = 175,833	Risk scores for needing a referral to different wraparound services such as behavioural health services, dietitian councelling and social services.	A machine learning algorithm, Elixhauser score	Application study	USA	The American Journal of Preventive Medicine
(Violan et al., 2013)[70]	Primary care population of 13 Primary Healthcare Centers n=196,593	Efficiency and effectiveness indicators on resource consumption	Adjusted Clinical Groups	Application study	Spain	BMC Health Service Research

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Vivas-Consuelo et al., 2014) [71]	Primary care population of general practice n=261,054	Predict pharmaceutical spending	ATC model*, Clinical Risk Groups	Application study	Spain	Health Policy
(Vuik, Mayer, & Darzi, 2016) [72]	Primary and Secondary care population; n=300,000	Identification high risk patients	Create risk scores, no tools mentioned	Validation study	UK	BMJ Open
(Wallace et al., 2016) [22]	Older community dwelling adults aged 70 years or older n = 862	Predicting emergency hospital admission	Total disease count, Selected conditions disease count, Charlson Comorbidity Index, number of dispensed medication classes, RxRisk-V	Comparison study	Ireland	BMJ Open
(Wennberg et al., 2013) [73]	Medicare beneficiaries n=5,153,877	Predicted one-year mortality	Charlson Comorbidity Index, Lezzoni chronic condition count, Hierarchical Condition Categories	Application study	USA	BMJ
(Xu, Williams-Livingston, Gaglioti, McAllister, & Rust, 2018) [74]	Patients aged 18 year and older, seen in an urban academic family medicine clinic over a two-year period; n=5,364	Utilization (high costs)	Elixhauser	Application study	USA	Journal of Health Care for the Poor and Underserved
(Zhou, Wong, & Li, 2014) [75]	Elderly aged 65 years and older n = 91,189	Mortality, hospitalization and costs of care	Senior Segmentation Algorithm	Validation study	USA	Permanente Journal

**ATC model categorizes patients into nine categories based on Anatomic Therapeutical Chemical (ATC) codes for medications.*