



Universiteit
Leiden
The Netherlands

A systematic review of risk stratification tools internationally used in primary care settings

Girwar, S.A.M.; Jabroer, R.; Fiocco, M.; Sutch, S.P.; Numans, M.E.; Bruijnzeels, M.A.

Citation

Girwar, S. A. M., Jabroer, R., Fiocco, M., Sutch, S. P., Numans, M. E., & Bruijnzeels, M. A. (2021). A systematic review of risk stratification tools internationally used in primary care settings. *Health Science Reports*, 4(3). doi:10.1002/hsr2.329

Version: Publisher's Version


License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3276369>

Note: To cite this publication please use the final published version (if applicable).

REVIEW

A systematic review of risk stratification tools internationally used in primary care settings

Shelley-Ann M. Girwar^{1,2}  | Robert Jabroer¹ | Marta Fiocco^{3,4,5} |
 Stephen P. Sutch^{1,6} | Mattijs E. Numans¹ | Marc A. Bruijnzeels^{1,2}

¹Department of Public Health and Primary Care, LUMC Campus the Hague, Leiden University Medical Centre, The Hague, The Netherlands

²Jan van Es Instituut, Ede, The Netherlands

³Mathematical Institute, Leiden University, Leiden, The Netherlands

⁴Medical Statistics Department of Biomedical Data Science, Leiden University Medical Center, Leiden, The Netherlands

⁵Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

⁶Department of Health Policy and Management, Bloomberg School of Public Health Johns Hopkins University, Baltimore, Maryland, USA

Correspondence

Shelley-Ann M. Girwar, Department of Public Health and Primary Care, LUMC Campus the Hague, Leiden University Medical Centre, Turfmarkt 99, The Hague, South Holland 2511 DP, The Netherlands.
 Email: s.a.girwar@lumc.nl

Funding information

Noaber Foundation

Abstract

Background and Aims: In our current healthcare situation, burden on healthcare services is increasing, with higher costs and increased utilization. Structured population health management has been 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 Organisation for Economic Co-operation and Development countries published in English-language journals were included. Search engines were utilized with keywords, for example, “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 R^2 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 were above 0.7, with ACG showing slightly improved scores compared with 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.

KEYWORDS

population health management, primary healthcare, risk assessment

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Health Science Reports* published by Wiley Periodicals LLC.

1 | INTRODUCTION

For several decades, healthcare costs have been rising. This has been attributed to aging 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 focuses on stratifying hospital populations, based on readily available hospital data. However, primary care data have 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.

2 | 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 the Supporting Information S1. 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).

2.1 | Inclusion criteria

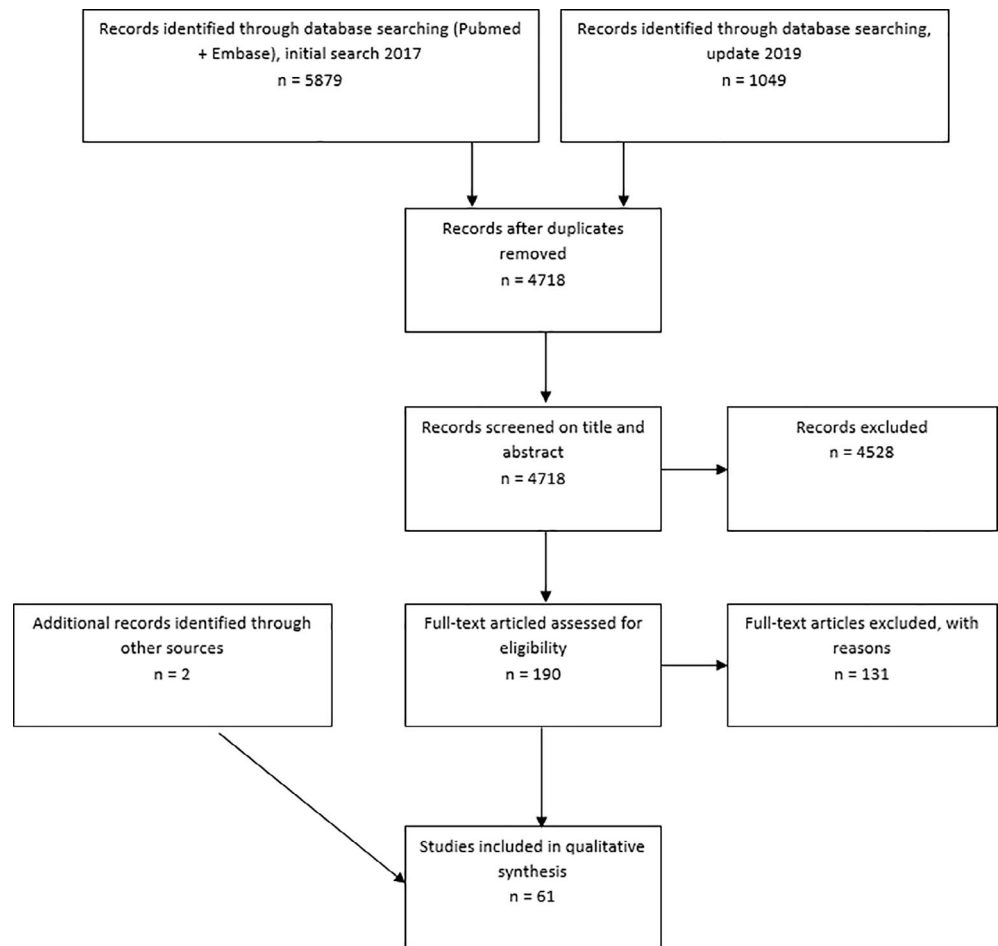
The search characteristics are specified by the *Population*, *Intervention*, *Control*, and *Outcome* method. In our research, the *population* is the primary care population. Therefore, we only included articles where models applied to primary care populations are discussed. The *interventions* investigated 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,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.

2.2 | Exclusion criteria

Articles that used risk stratification tools on populations consisting of hospitalized patients or patients seeking consultation with a specialist

FIGURE 1 PRISMA flowchart displaying numbers of included and excluded articles



(eg, 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 the title and abstract, based on the criteria mentioned earlier. Seventy-eight percentage of the screening, based on the title and abstract, was performed by two researchers independently (R.J. and 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 were only screened by one researcher. After screening the 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.

2.3 | 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 statistical 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 continuous 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 are 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.

3 | 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 the Supporting Information S2.

3.1 | 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.

3.2 | 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 1-5 or more conditions as category 6).^{18,19} The model investigates the effect of multimorbidity and predicts several outcomes. Variations of the CCI exist, and the validity on predictions has 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 corresponds to an accuracy of *sufficient* and *good*. For *emergency department visitation*, C-statistics between 0.58 and 0.63 is 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

3.3 | Hierarchical condition categories: Seven 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

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:				
<i>Hospitalization</i> (Number of hospitalizations) (Unplanned hospitalizations)	Haas ⁴	C = 0.73	C = 0.68	C = 0.67
	Lemke ¹²	AUC = 0.80	AUC = 0.78	
	Shadmi ¹⁶	R ² = .24	R ² = .11	
	Maltenfort ¹¹	AUC = 0.82		
	Inouye ²⁰		C = 0.72	
	Ou ²¹		C = 0.61	
	Mosley ²⁵			AUC = 0.64
<i>Emergency department visits</i>	Haas ⁴	C = 0.67	C = 0.59	C = 0.58
	Ou ²¹		C = 0.63	
	Wallace ²²		C = 0.58	
<i>Costs</i>				
(Top 10% total costs)	Haas ⁴	C = 0.76	C = 0.70	C = 0.70
(Total costs)	Brilleman ¹⁴	R ² = .41	R ² = .34	
(Pharmaceutical costs)	Aguado ¹⁰	R ² = .39		
(Total costs)	Sicras-Mainar ¹³	R ² = .37		
(Total costs)	Charlson ¹⁸		R ² = .22	
(Total costs)	Charlson ¹⁹		R ² = .20	
(High total costs)	Ou ²¹		C = 0.64	
<i>Utilization of different healthcare services</i>				
(GP visits)	Brilleman ¹⁵	R ² = .37	R ² = .26	
(Primary care visits)	Shadmi ¹⁶	R ² = .54	R ² = .18	
(Specialist visits)	Shadmi ¹⁶	R ² = .45	R ² = .13	
(Number of diagnostic imaging tests)	Shadmi ¹⁶	R ² = .37	R ² = .15	
(Visits)	Sicras-Mainar ¹³	R ² = .42		
(Number of diagnoses/reasons for visit)	Sicras-Mainar ¹³	R ² = .77		
(High outpatient visits)	Ou 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

Abbreviations: AUC, area under the ROC curve; C, C-statistic; R², R-square.

enrollees 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 0.67⁴ 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.

3.4 | Comparison studies

A total of 10 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 vs 0.78¹² and C-statistics of 0.73 vs 0.68.⁴ The ACG also outperforms the CCI regarding emergency department visitation with C-statistics of 0.67 vs 0.59 and high total costs with C-statistics of 0.76 vs 0.70,⁴ and R^2 values of .41 vs .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 ($R^2 = .24$ vs $R^2 = .11$), primary care visits ($R^2 = .54$ vs $R^2 = .18$), specialist visits ($R^2 = .45$ vs $R^2 = .13$), and diagnostic imaging tests ($R^2 = .37$ vs $R^2 = .15$) all within a study period of 12 months.¹⁶ In addition, Brilleman and colleagues find R^2 values of .37 for the ACG and 0.26 for the CCI with the number of general practitioner visits as the predicted outcome.¹⁵

3.5 | 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 ($n = 1, 2, \text{ or } 3$) 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 Supporting Information S2, all of the risk stratification tools are presented and organized by included studies.

4 | DISCUSSION

4.1 | 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 hospitalizations¹² 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

4.2 | 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 interobserver 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.

4.3 | 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, centered 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.

5 | 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.

ACKNOWLEDGMENTS

The authors acknowledge the Walaeus library of the Leiden University Medical Center for their collaboration in generating the search strategy for our PubMed and Embase search. The authors also acknowledge the contribution of Josanne Mansveld in the study screening process.

CONFLICT OF INTEREST

All authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

AUTHOR CONTRIBUTIONS

Conceptualizing: Shelley-Ann Girwar, Robert Jabroer, Marc Bruijnzeels

Data Curation: Shelley-Ann Girwar, Robert Jabroer

Funding Acquisition: Marc Bruijnzeels

Investigation: Shelley-Ann Girwar, Robert Jabroer.

Methodology: Shelley-Ann Girwar, Robert Jabroer, Marc Bruijnzeels

Supervision: Marc Bruijnzeels, Mattijs Numans

Visualization: Shelley-Ann Girwar, Robert Jabroer, Stephen Sutch, Marta Fiocco, Marc Bruijnzeels

Writing—Original Draft Preparation: Robert Jabroer, Shelley-Ann Girwar

Writing—Review & Editing: Shelley-Ann Girwar, Robert Jabroer, Stephen Sutch, Marta Fiocco, Marc Bruijnzeels, Mattijs Numans.

All authors have read and approved the final version of the manuscript.

TRANSPARENCY STATEMENT

Authors affirm that the manuscript is an honest, accurate and transparent account of the study being reported, that no important aspects of this study have been omitted and that there were no discrepancies from the study as planned.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no new data were created or analyzed during this study.

ETHICS STATEMENT

Ethical approval was not required for this study.

ORCID

Shelley-Ann M. Girwar  <https://orcid.org/0000-0001-9104-9624>

REFERENCES

1. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. *J Gen Internal Med.* 2007;22(3):391-395. <https://doi.org/10.1007/s11606-007-0322-1>
2. Luo G, Stone BL, Sakaguchi F, et al. Using computational approaches to improve risk-stratified patient management: rationale and methods. *JMIR Res Prot.* 2015;4(4):e128. <https://doi.org/10.2196/resprot.5039>
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. <https://doi.org/10.1097/00004479-200701000-00007>

4. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. *Am J Manag Care*. 2013;19(9):725-732.
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 Med Inform Decis Mak*. 2017;17(1):139. <https://doi.org/10.1186/s12911-017-0538-x>
6. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
7. Development Of EC-OA. Member Countries.
8. Ket H. Tips & trucs voor het uitvoeren van systematische reviews met EndNote.
9. Simundic AM. Measures of diagnostic accuracy: basic definitions. *eJFCC*. 2009;19(4):203-211.
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 Serv Res*. 2008;8:53. <https://doi.org/10.1186/1472-6963-8-53>
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. <https://doi.org/10.1371/journal.pone.0221233>
12. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. *Med Care*. 2012;50(2):131-139. <https://doi.org/10.1097/MLR.0b013e3182353ceb>
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. *J Eval Clin Pract*. 2013;19(2):267-276. <https://doi.org/10.1111/j.1365-2753.2012.01818.x>
14. Brilleman SL, Gravelle H, Hollinghurst S, et al. Keep it simple? Predicting primary health care costs with clinical morbidity measures. *J Health Econ*. 2014;35:109-122. <https://doi.org/10.1016/j.jhealeco.2014.02.005>
15. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. *Family Pract*. 2013;30(2):172-178. <https://doi.org/10.1093/fampra/cms060>
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. <https://doi.org/10.1186/1471-2458-11-609>
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):373-383. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)
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. *J Clin Epidemiol*. 2008;61(12):1234-1240. <https://doi.org/10.1016/j.jclinepi.2008.01.006>
19. Charlson ME, Wells MT, Kanna B, et al. Medicaid managed care: how to target efforts to reduce costs. *BMC Health Serv Res*. 2014;14:461. <https://doi.org/10.1186/1472-6963-14-461>
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. *Med Care*. 2008;46(7):726-731. <https://doi.org/10.1097/MLR.0b013e3181649426>
21. Ou HT, Mukherjee B, Erickson SR, et al. Comparative performance of comorbidity indices in discriminating health-related behaviors and outcomes. *Health Outcomes Res Med*. 2011;2(2):e91-e104. <https://doi.org/10.1016/j.ehrm.2011.06.002>
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. <https://doi.org/10.1136/bmjopen-2016-013089>
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. *Popul Health Manag*. 2012;15(4):220-229. <https://doi.org/10.1089/pop.2011.0037>
24. Khan NF, Perera R, Harper S, et al. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Family Pract*. 2010;11:1. <https://doi.org/10.1186/1471-2296-11-1>
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? *J Am Geriatr Soc*. 2009;57(12):2306-2310. <https://doi.org/10.1111/j.1532-5415.2009.02558.x>
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. *Med Care*. 1998;36(1):8-27. <https://doi.org/10.1097/00005650-199801000-00004>
28. Dominick KL, Dudley TK, Coffman CJ, et al. Comparison of three comorbidity measures for predicting health service use in patients with osteoarthritis. *Arthritis Rheum*. 2005;53(5):666-672. <https://doi.org/10.1002/art.21440>
29. Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol*. 2004;57(10):1040-1048. <https://doi.org/10.1016/j.jclinepi.2004.03.002>
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. *Ann Family Med*. 2012;10(2):134-141. <https://doi.org/10.1370/afm.1363>
31. Johns Hopkins University. <https://www.hopkinsagc.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*. 2013;113(1-2):206-215. <https://doi.org/10.1016/j.healthpol.2013.09.006>
33. Burton LC, Skinner EA, Uscher-Pines L, et al. Health of Medicare Advantage plan enrollees at 1 year after Hurricane Katrina. *Am J Manag 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 Serv Res*. 2019;54(2):466-473. <https://doi.org/10.1111/1475-6773.13094>
35. Akazawa M, Imai H, Igarashi A, et al. Potentially inappropriate medication use in elderly Japanese patients. *Am J Geriatr Pharmacother*. 2010;8(2):146-160. <https://doi.org/10.1016/j.amjopharm.2010.03.005>
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. <https://doi.org/10.1016/j.maturitas.2019.01.002>
37. Chang HY, Richards TM, Shermock KM, et al. Evaluating the impact of prescription fill rates on risk stratification model performance. *Med Care*. 2017;55(12):1052-1060. <https://doi.org/10.1097/mlr.0000000000000825>
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. *Prevent Med*. 2018;115:110-118. <https://doi.org/10.1016/j.ypmed.2018.08.018>
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 Serv Res.* 2010;10:338. <https://doi.org/10.1186/1472-6963-10-338>
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. *J Gen Internal Med.* 2018;33(12):2171-2179. <https://doi.org/10.1007/s11606-018-4626-0>
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 Serv Res.* 2019;19(1):526. <https://doi.org/10.1186/s12913-019-4337-1>
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. <https://doi.org/10.1136/bmjopen-2015-010301>
43. Freund T, Kunz CU, Ose D, et al. Patterns of multimorbidity in primary care patients at high risk of future hospitalization. *Popul Health Manag.* 2012;15(2):119-124. <https://doi.org/10.1089/pop.2011.0026>
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. *Ann Family Med.* 2009;7(5):396-405. <https://doi.org/10.1370/afm.994>
45. Hamano J, Oishi A, Kizawa Y. Prevalence and characteristics of patients being at risk of deteriorating and dying in primary care. *J Pain Symp Manag.* 2019;57(2):266-72.e1. <https://doi.org/10.1016/j.jpainsymman.2018.11.006>
46. Hewner S, Seo JY, Gothard SE, et al. Aligning population-based care management with chronic disease complexity. *Nurs Outlook.* 2014;62(4):250-258. <https://doi.org/10.1016/j.outlook.2014.03.003>
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. *J Gen Internal Med.* 2015;30(12):1741-1747. <https://doi.org/10.1007/s11606-015-3357-8>
48. Hu T, Dattani ND, Cox KA, et al. Effect of comorbidities and medications on frequency of primary care visits among older patients. *Can Fam Phys.* 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. <https://doi.org/10.1186/1745-6215-14-301>
50. Khanna S, Rolls DA, Boyle J, et al. A risk stratification tool for hospitalisation in Australia using primary care data. *Sci Rep.* 2019;9(1):5011. <https://doi.org/10.1038/s41598-019-41383-y>
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. *Appl Clin Inform.* 2017;8(3):794-809. <https://doi.org/10.4338/aci-2016-12-ra-0210>
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):e022377.
53. Metcalfe D, Masters J, Delmestri A, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. *BMC Med Res Methodol.* 2019;19(1):115. <https://doi.org/10.1186/s12874-019-0753-5>
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 Qual Life Outcome.* 2019;17(1):8. <https://doi.org/10.1186/s12955-018-1063-z>
55. Moran WP, Zhang J, Gebregziabher M, et al. Chaos to complexity: leveling the playing field for measuring value in primary care. *J Eval Clin Pract.* 2017;23(2):430-438. <https://doi.org/10.1111/jep.12298>
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. <https://doi.org/10.9778/cmajo.20180185>
57. Noyes K, Liu H, Temkin-Greener H. Medicare capitation model, functional status, and multiple comorbidities: model accuracy. *Am J Manag Care.* 2008;14(10):679-690.
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. *Scand J Primary Health Care.* 2018;36(3):308-316. <https://doi.org/10.1080/02813432.2018.1499514>
59. Rohrer JE, Rasmussen N, Adamson SA. Illness severity and total visits in family medicine. *J Eval Clin Pract.* 2008;14(1):65-69. <https://doi.org/10.1111/j.1365-2753.2007.00797.x>
60. Salisbury C, Johnson L, Purdy S, et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract.* 2011;61(582):e12-e21. <https://doi.org/10.3399/bjgp11X548929>
61. Sibley LM, Moineddin R, Agha MM, et al. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. *Med Care.* 2010;48(2):175-182. <https://doi.org/10.1097/MLR.0b013e3181c16102>
62. Sibley LM, Glazier RH. Evaluation of the equity of age-sex adjusted primary care capitation payments in Ontario, Canada. *Health Policy.* 2012;104(2):186-192. <https://doi.org/10.1016/j.healthpol.2011.10.008>
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. *Europ J Public Health.* 2007;17(6):657-663. <https://doi.org/10.1093/eurpub/ckm044>
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): an observational study. *BMJ open.* 2012;2(3). <https://doi.org/10.1136/bmjopen-2012-000941>
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 Pharm Toxicol.* 2013;14:39. <https://doi.org/10.1186/2050-6511-14-39>
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 Qual Safe.* 2019;28(9):697-705. <https://doi.org/10.1136/bmjqs-2018-007976>
67. Sternberg SA, Bentur N, Abrams C, et al. Identifying frail older people using predictive modeling. *Am J Manag Care.* 2012;18(10):e392-e397.
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 Clin Proc.* 2013;88(9):963-969. <https://doi.org/10.1016/j.mayocp.2013.06.015>
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. *Am J Prevent Med.* 2019;56(4):e125-e133. <https://doi.org/10.1016/j.amepre.2018.11.009>
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 Serv Res.* 2013;13:421. <https://doi.org/10.1186/1472-6963-13-421>
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.* 2014;116(2-3):188-195. <https://doi.org/10.1016/j.healthpol.2014.01.012>
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. <https://doi.org/10.1136/bmjopen-2016-012903>
73. Wennberg JE, Staiger DO, Sharp SM, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. *BMJ Clin Res*. 2013;346:f549. <https://doi.org/10.1136/bmj.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. *J Health Care Poor Underserv*. 2018;29(1):202-213. <https://doi.org/10.1353/hpu.2018.0014>
75. Zhou YY, Wong W, Li H. Improving care for older adults: a model to segment the senior population. *Perm J*. 2014;18(3):18-21. <https://doi.org/10.7812/tpp/14-005>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Girwar S-AM, Jabroer R, Fiocco M, Sutch SP, Numans ME, Bruijnzeels MA. A systematic review of risk stratification tools internationally used in primary care settings. *Health Sci Rep*. 2021;4:e329. <https://doi.org/10.1002/hsr2.329>