


RESEARCH

Open Access



Retrospective analysis of hospitalization costs using two payment systems: the diagnosis related groups (DRG) and the Queralt system, a newly developed case-mix tool for hospitalized patients

Júlia Folguera^{1,2}, Elisabet Buj³, David Monterde^{2,3}, Gerard Carot-Sans^{1,2}, Isaac Cano⁴, Jordi Piera-Jiménez^{1,2,5*}  and Miquel Arrufat³

Abstract

Background Hospital services are typically reimbursed using case-mix tools that group patients according to diagnoses and procedures. We recently developed a case-mix tool (i.e., the Queralt system) aimed at supporting clinicians in patient management. In this study, we compared the performance of a broadly used tool (i.e., the APR-DRG) with the Queralt system.

Methods Retrospective analysis of all admissions occurred in any of the eight hospitals of the Catalan Institute of Health (i.e., approximately, 30% of all hospitalizations in Catalonia) during 2019. Costs were retrieved from a full cost accounting. Electronic health records were used to calculate the APR-DRG group and the Queralt index, and its different sub-indices for diagnoses (main diagnosis, comorbidities on admission, and complications occurred during hospital stay) and procedures (main and secondary procedures). The primary objective was the predictive capacity of the tools; we also investigated efficiency and within-group homogeneity.

Results The analysis included 166,837 hospitalization episodes, with a mean cost of € 4,935 (median 2,616; interquartile range 1,011–5,543). The components of the Queralt system had higher efficiency (i.e., the percentage of costs and hospitalizations covered by increasing percentages of groups from each case-mix tool) and lower heterogeneity. The logistic model for predicting costs at pre-established thresholds (i.e., 80th, 90th, and 95th percentiles) showed better performance for the Queralt system, particularly when combining diagnoses and procedures (DP): the area under the receiver operating characteristics curve for the 80th, 90th, 95th cost percentiles were 0.904, 0.882, and 0.863 for the APR-DRG, and 0.958, 0.945, and 0.928 for the Queralt DP; the corresponding values of area under the precision-recall curve were 0.522, 0.604, and 0.699 for the APR-DRG, and 0.748, 0.7966, and 0.834 for the Queralt DP. Likewise, the linear model for predicting the actual cost fitted better in the case of the Queralt system.

*Correspondence:
Jordi Piera-Jiménez
jpiera@catsalut.cat

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions The Queralt system, originally developed to predict hospital outcomes, has good performance and efficiency for predicting hospitalization costs.

Keywords Hospital costs, Case-mix tools, Diagnosis-related groups, Queralt system

Background

The increasing complexity of healthcare systems and the diversity of healthcare services experienced in the last decades have transformed how healthcare providers are paid for their services [1]. This interest has been particularly important in the hospital setting, which is the most costly of the healthcare system [2].

Traditional payment approaches, such as retrospective fee-for-service models —where providers are paid for the services and procedures delivered—, have discouraged efficiency and quality, often leading to unnecessary procedures and longer average values of the length of stay. Similarly, prospective per-diem payments — where providers are paid for a type of in-patient day, hampered healthcare planning and budgeting [3]. To overcome these drawbacks, many healthcare systems worldwide initiated a transition towards case-mix tools, which summarize a large number of items related to patient complexity and procedures into a limited number of groups with certain characteristics in common, allowing classification of all hospitalizations.

Case-mix tools for summarizing the clinical complexity of patients have been extensively used in healthcare. However, most assessments reported in the literature have investigated clinical or hard endpoints, such as mortality (in-hospital and at different time points after discharge), while the performance of these tools for cost estimate have been scarce [4]. More than 20 years ago, Ash et al. showed how diagnoses and diagnostic groups in general aligned with the actual costs for procedure [5]. Similarly, Huang et al. compared the ability of several case-mix tools developed for clinical purposes to inform healthcare costs. The authors found that, although all case-mix tools performed acceptably for predicting costs, proprietary tools were better, limiting their adaptability to specific environments [6]. Finally, it is worth mentioning that, recently, with the advent of machine learning use in healthcare, some authors have improved cost predictions by including information on social determinants [7].

The most popular case-mix tool is the diagnosis-related groups (DRG), which rapidly expanded across the globe. The DRG was created at Yale University in the 1960s to define the services provided by each hospital and manage costs and quality and has rapidly been implemented across the globe. In its more recent version, all-patient refined (APR-DRG), it has several advantages, including transparency, exhaustivity (e.g., accounts for severity and risk levels), and the possibility of measuring quality

indicators across hospitals and even countries [8]. However, it is costly and inflexible: it does not use all the administrative data available in each healthcare system and it is usually trained with the USA population and not the local one, thus precluding healthcare systems from tailoring this tool to their needs and capabilities in terms of data collection. Furthermore, the APR-DRG accounts for diagnoses and procedures but not complications that may occur during hospitalization and may impact the overall cost.

Recently, we developed a comprehensive risk adjustment tool for predicting health outcomes in hospitalized patients: the Queralt system [9]. The tool was designed as a family of indices that summarize the principal diagnosis (Queralt DxP), in-hospital complications (Queralt DxC), and diagnoses on admission (Queralt DxS) based on weighted sums of over 2,100 diagnostic codes. All indices can be summarized in the Queralt Dx. The Queralt System has shown a high capacity to predict critical events such as death or admission to the intensive care unit (ICU) in all-cause hospitalizations [9] and disease-specific hospitalizations [10]. Additionally, the Queralt System includes two indices for primary and secondary procedures, respectively. Like in the case of diagnoses, the procedure indices can be combined in a single one (i.e., Queralt Px). However, its performance as a paying system for hospital services has not been assessed. In this retrospective analysis of costs, we compared the performance of the APR-DRG and the Queralt Dx and Px indices for estimating the costs of hospital services.

Methods

Study design and setting

This was a cost analysis of hospital services that included all hospitals in the Catalan Institute of Health (ICS for Catalan *Institut Català de la Salut*). The ICS is the leading healthcare provider in Catalonia (North-East Spain), and delivers public and universal healthcare to its reference population, accounting for approximately 30% of all hospitalizations in Catalonia. Three of the eight hospitals managed by the ICS are tertiary hospitals with specialized reference services that receive referrals from other centers of the public healthcare network. Hospitals of the ICS are currently paid by the public healthcare insurer (i.e., the Catalan Health Service) on an activity-based payment system. Service costs are established based on full cost accounting, which considers direct and indirect costs of healthcare service delivery, aimed at establishing the cost of a given provision (i.e., the smallest unit of care

received by the patient). The provision-based grouping allows estimating the cost per episode, patient, or diagnosis by grouping several provisions.

Two type of costs were considered: direct costs and indirect costs. Direct costs included all costs that the provider entity (i.e., the hospital) allocates to the units providing care activities (e.g., cardiology department, dermatology department...). These costs encompass personnel of the given unit, pharmacy, and supplies that are central for the healthcare activity. Considering that cost may vary according to services, a weight of relative unit of value is assigned to each unit. The cost for each service is added to each recorded patients within the hospitalization episode; the sum of all services represents the patient cost during the admission period.

Indirect costs include all attributed to units that do not perform healthcare activity per se. Examples of these units include admissions, clinical documentation, sterilization, and accounting. Indirect costs are distributed proportionally to healthcare units at the hospital level. Proportionality criteria are defined for the entire organization for each type of cost (e.g., cleaning expenses are assigned based on weighted square meters per unit type, management expenses based on staffing).

In this analysis, we retrospectively collected the costs of all regular hospitalizations (i.e., excluding hospitalizations because of major surgery and other hospital procedures) that occurred in any of the ICS hospitals between January 01 and December 31, 2019. Registries of DRG groups with less than 25 cases and those without grouping criteria in either the DRG or the Queralt system were removed. The cost estimate was conducted in euros for 2019.

Cost information was retrieved without personal or health information; therefore, the analysis did not meet the criteria for being considered a clinical study; therefore, approval by an independent ethics committee was deemed not necessary by the Bellvitge Independent Ethics Committee (Spain).

Case-mix tools

The development of the Queralt system for diagnoses is described elsewhere [9]; the software and user's manual of the Queralt system can be accessed from <https://ics.gencat.cat/ca/assistencia/coneixement-assistencial/sistema-de-Queralt/>. The Queralt system considers diagnoses and procedures separately, both summarized as indices (i.e., numerical values that measure the complexity of hospitalization). In the case of diagnoses, the system provides three sub-indices: the main diagnosis (QIDP), comorbidities present on admission (QIDA), and complications occurred during hospital stay (QIDC). The three indices regarding diagnoses can be summarized in the Queralt index for diagnoses (QID), which is

the sum of QIDP+QIDA+QIDC. Similarly, the Queralt system has two sub-indices for procedures: main procedure (QIPP) and secondary procedures (QIPS). Like the diagnosis indices, procedures indices can be summarized in a Queralt indices for procedures (QIP), which is the sum of QIPP+QIPS. The distribution of each of the summary indices (i.e., QIP and QID) is used to define complexity groups for diagnoses (D_strata) and procedures (P_strata). The supplementary methods (Supplementary appendix) provides details regarding the definition of the complexity groups or strata. In this analysis, we used both the numerical indices (QID and QIP) and the groups or strata (D_strata and P_strata). We also combined the two last (i.e., scalar product of D_strata and P_strata) to obtain an exhaustive stratification based on both, diagnoses and procedures (DP_strata).

The DRG system is a for-profit case-mix tool designed for assessing patients' complexity and service payment that allows assigning hospital services to comparable groups and determining the weight or price for each of the groups of products [11]. In Catalonia, the APR-DRG is used. The APR-DRG divides each base DRG into four risk levels (based on mortality risk) and four severity levels (based on the length-of-stay and other hospital outcomes). In this study, we used the DRG Severity (i.e., the Cartesian combination of the base DRG and severity levels related to resource utilization).

Outcomes and statistical approach

The primary objective was to compare the predictive capacity of costs between the different case-mix tools. This objective was addressed using two outcomes: the cost as a numerical (continuous) variable in order to study if the groupers could predict the distribution of costs, and the cost (as categorical, binary) stratified by the 80th, 90th, and 95th percentiles to investigate the capacity to predict high costs.

Before model building, we examined the classification systems using two different indicators: efficiency and within-group homogeneity. The efficiency was defined as the percentage of costs and hospitalizations covered by increasing percentages of groups from each case-mix tool. The within-group homogeneity was assessed by estimating the coefficient of variation (i.e., standard deviation over mean) for costs in each group of the case-mix tool.

To test the predictive capacity of the case-mix tools, we built two model types for each grouper: D_strata, P_strata, and DRG severity. First, we performed a logistic regression for each grouper to predict high cost based on three thresholds of cost distribution: percentiles 80th, 90th, and 95th. Then, we conducted a lognormal regression for each grouper to predict the actual cost. Since risk and severity levels make sense only within a given

Table 1 Costs (in €) according to demographic characteristics

	Mean	Median (IQR)
Sex		
Male	5337.4	2671 (1003–6025)
Female	4509.9	2573 (1019–5122)
Age group		
0–9	3589.3	1652 (883–3208)
10–19	4537.5	2168 (967–4258)
20–29	4173.7	2503 (1213–4203)
30–39	4050.4	2548 (1414–4208.25)
40–49	5013.5	2749 (1081–5663)
50–59	5840.4	3139 (1178–6642)
60–69	6048.3	3241 (1220–7000)
70–79	5569.5	3032 (1074–6625)
80–89	4229.7	2374 (764–5169)
90–99	3006.2	1771 (617–4020)
100+	2782.9	1154 (526–4112)

IQR: interquartile range, defined by the 25th and 75th percentiles

DRG base group, we stratified all analyses for each DRG base group. DRG base groups with less than 25 registries were removed. The predictive validity of the models was assessed using the reduction in variance, as described elsewhere [12, 13]. Briefly, the reduction in variance indicates the amount of variation in the dependent variable explained by the classification of items. The performance of the logistic regression models for high costs was assessed using the receiving operating characteristics (ROC) and precision-recall (PR) curves. Additionally, we conducted linear models for descriptive purposes; the the performance of the linear models for actual costs was represented by plotting the expected and observed distribution of costs across groups.

Results

Overall costs

Our analysis included 166,837 hospitalization episodes, with a mean cost of € 4,935 (median 2,616; interquartile range 1,011–5,543). Table 1 summarizes the mean and median costs according to age groups and sex; the interactions between these demographic characteristics are provided in Table S1 (Supplementary Appendix). Overall, hospitalization costs within the investigated period amounted to 2,006 million euros; of them, 1,496 million were direct costs and 513 indirect costs. Figure 1 summarizes the distribution of costs per episode.

Figure S1 shows the distribution of average cost and number of hospitalizations for twenty relevant DRG groups. The average cost according to sex and age for eight relevant causes of hospitalization (either diagnoses or procedures, with overlap between DRG and Queralt groups) is shown in Figure S2.

Efficiency and heterogeneity

Figure 2 summarizes the values of costs (A) and frequency of hospitalization episodes (B) for each of the DP_strata (grouped first according to D_strata and, within them, the P_strata). The distribution of costs and hospitalizations revealed that the P_strata grouper explained better the differences in costs and episodes than the D_strata (Fig. 2). Also, we saw that less costly hospitalizations were remarkably more frequent than high-costly ones.

Regarding efficiency, we also investigated the percentage of costs and hospitalizations covered by each percentage of the Queralt and DRG groups (Fig. 3). The D_strata of the Queralt System showed the highest efficiency for hospitalizations (i.e., closest to the center diagonal, indicating proportionality). The highest efficiency

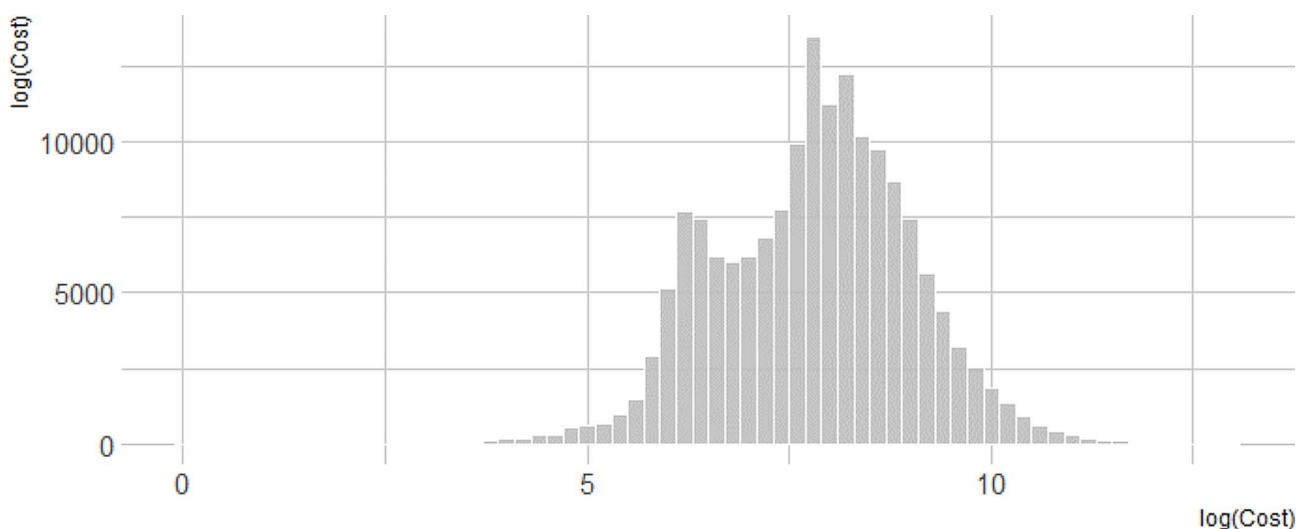


Fig. 1 Distribution of cost (logarithm)

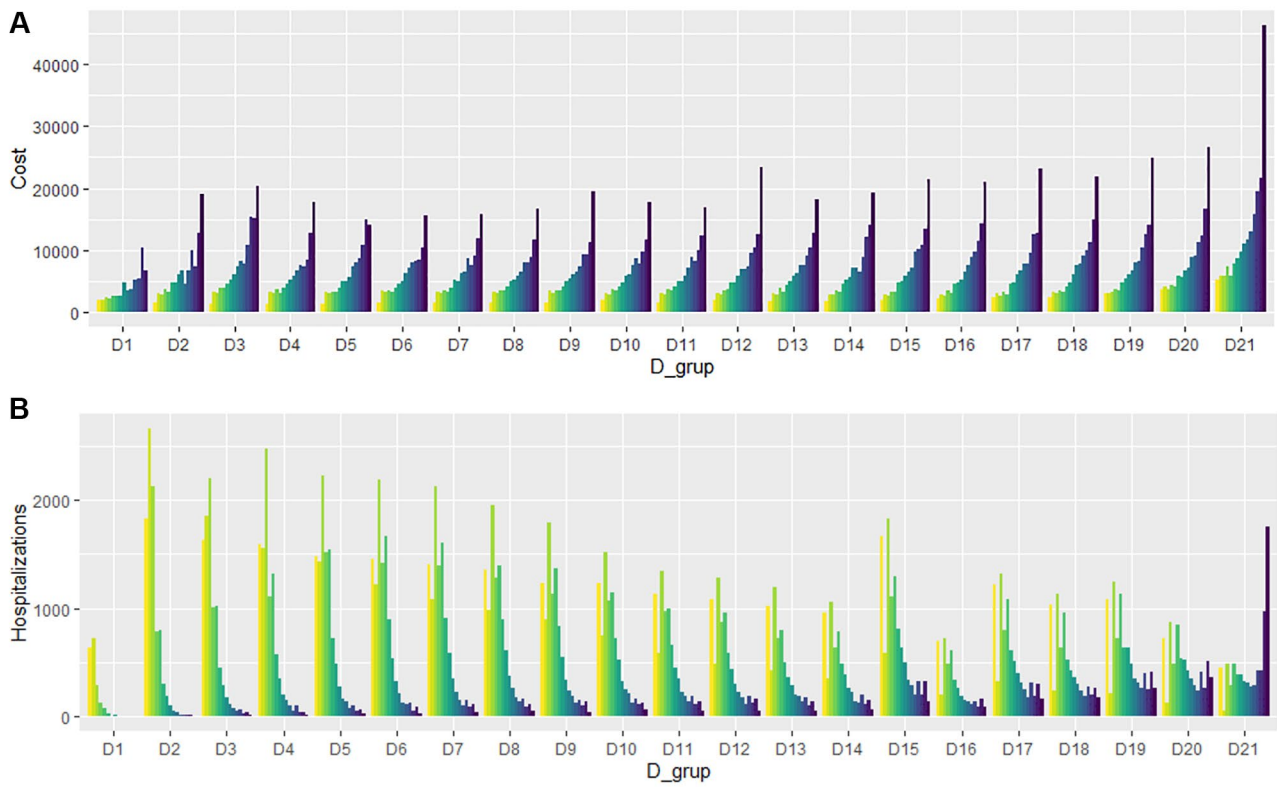


Fig. 2 Distribution of costs (A) and hospitalizations (B) within each of the D_strata and, within them, P_strata. Colors show the relative values for the P_strata, from yellow (lowest values) to dark blue (highest values)

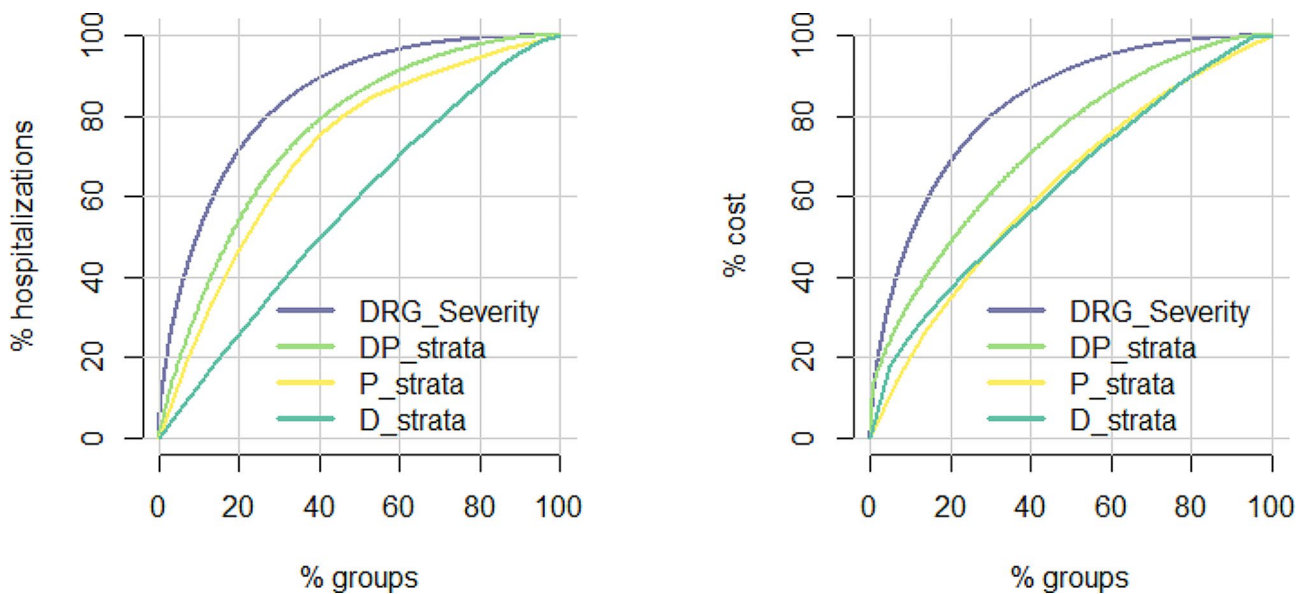


Fig. 3 Efficiency analysis of the DRG and Queralt System components

for costs was also observed for the P_strata and D_strata components. In the case of the DRG system, a moderate percentage of groups covered high percentages of costs and hospitalizations, and the coverage improved very

little when increasing the number of groups, indicating the low efficiency of the system.

The components of the Queralt system showed a relatively stable coefficient of variation around 1 across most groups. On the other hand, the DRG groups showed a

more heterogeneous profile, with some groups associated with the lowest coefficient of variation and some others with the highest (Fig. 4).

Predictive model

The reduction in variance of the predictive based on DRG_Severity grouper was 0.487. The corresponding values for the D_strata, P_strata, and DP_strata of the Queralt system were 0.149, 0.326, and 0.379, respectively, indicating that the DRG system (DRG_Severity) has slightly higher predictive validity.

According to the ROC analysis, the DP_strata of the Queralt system had the highest performance in predicting high hospital costs at all thresholds established (i.e., 80th, 90th, and 95th percentiles) (Fig. 5). The other components analyzed showed little differences. In the corresponding PR analysis, the superiority of the DP_Strata of the Queralt system was even more notorious, and the other two components of this system (i.e., the P_strata and D_strata) were slightly better than the DRG components.

Regarding the predictive capacity of the actual cost, the linear regression model based on the DP_strata of the Queralt system showed the closest distribution to the actual cost (Fig. 6).

Discussion

Summary of main findings

In this retrospective analysis of analytical costs in public hospitals in Catalonia, we found that a recently-developed case-mix tool for risk stratification in the hospital setting shows better performance (in terms of both efficiency and ability to predict costs) than the APR-DRG, a tool broadly used for reimbursement of hospital services.

Diagnoses and procedures separately

Classical case-mix tools developed as an approach to reimbursement procedures, including but not limited to DRG, use pre-defined combinations of diagnoses and procedures to build risk groups. Alternatively, the Queralt tool was conceived as a system of independent indices (including diagnoses and procedures) that can be used separately or combined, depending on the needs [9]. In our analysis, the distribution of costs and hospitalization events across diagnostic and procedure groups (D_strata and P_strata, respectively) revealed that procedures account for higher differences in the overall costs than diagnoses. This finding indicates that within a given diagnostic group, procedures may make an important difference in costs, which is consistent with cost analyses in specific settings that show how heterogeneous and difficult to predict may be the costs of hospital

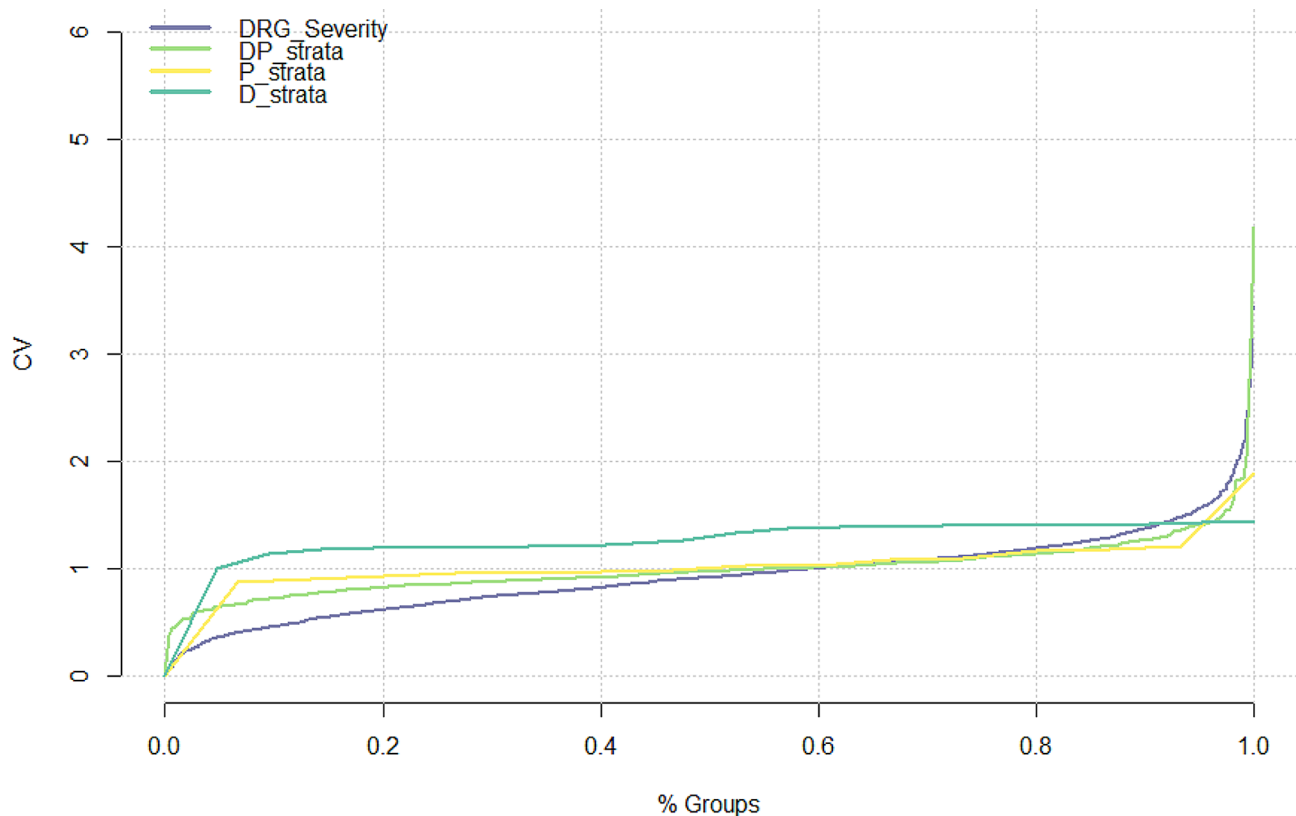


Fig. 4 Coefficient of variation of groups from the different case-mix tools

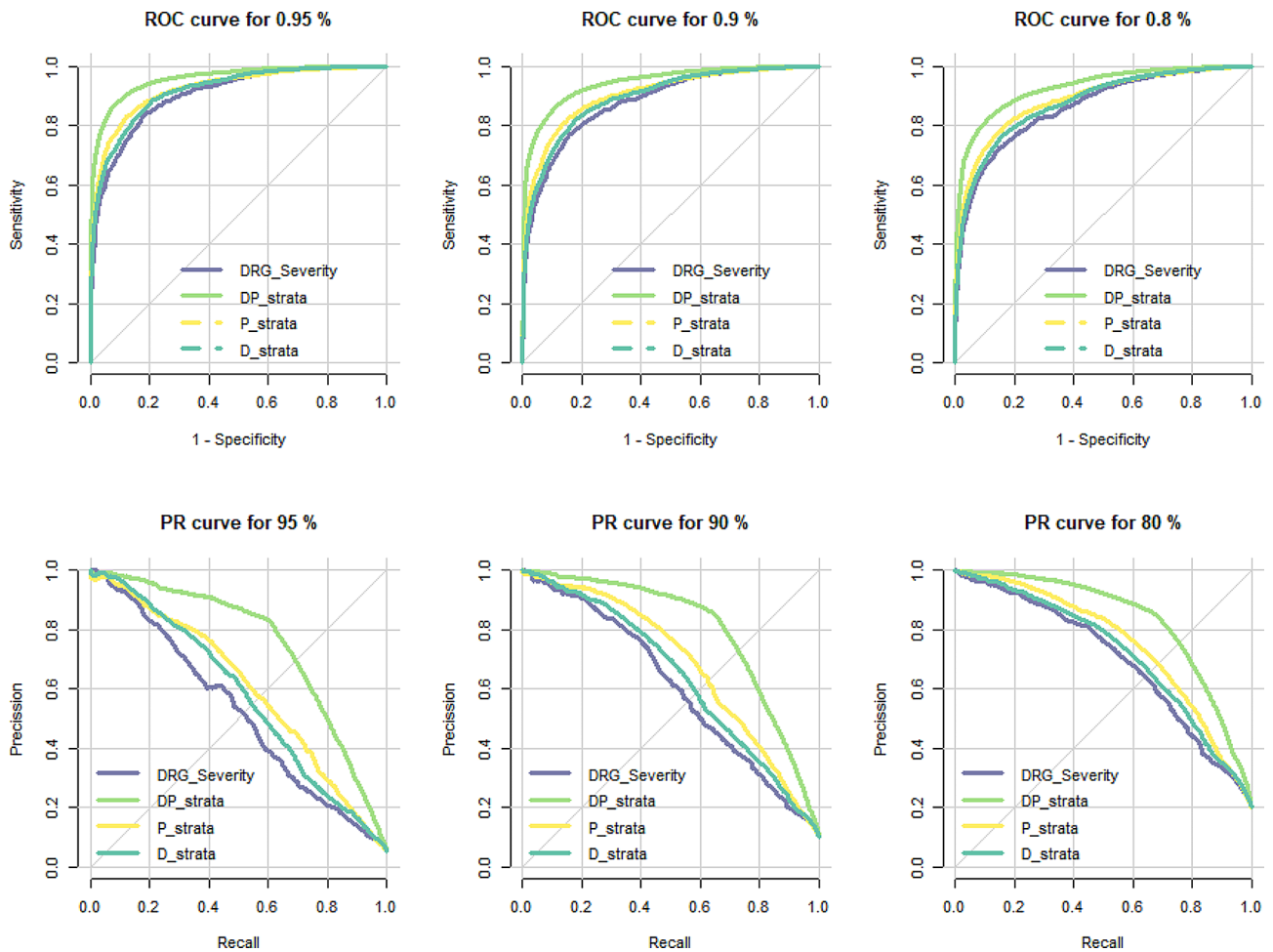


Fig. 5 Performance of the logistic regression models with components from the DRG and Queral systems for predicting high hospital costs when considering three percentile thresholds of the actual cost: 80th, 90th, and 95th. ROC: receiving operating characteristics. PR: precision-recall. The values for area under the curve are provided in Table S2 (Supplementary appendix)

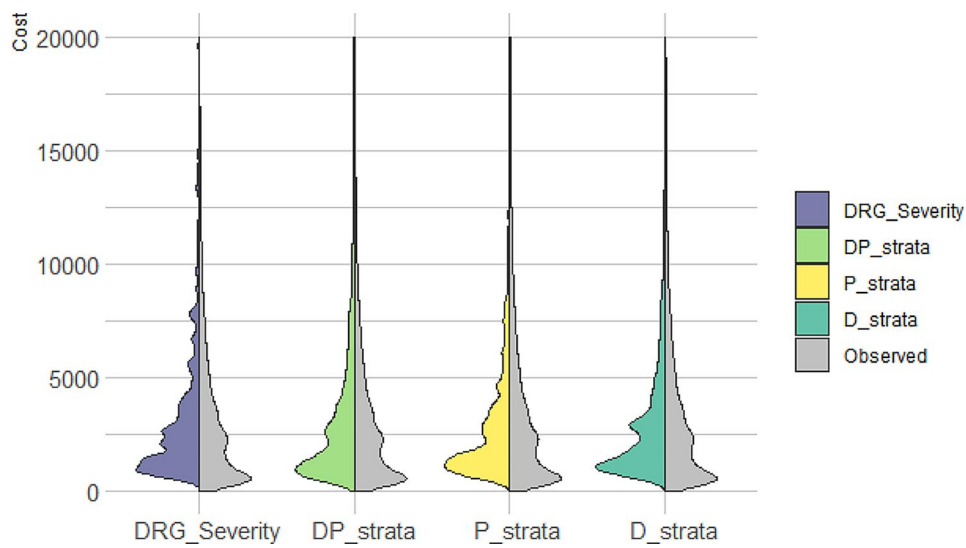


Fig. 6 Distribution of the observed and expected cost according to linear models built based on each of the grouper system

procedures despite sharing the same main diagnosis [14, 15]. Therefore, although the Queralt system ends up with a lower number of groups, the flexibility of combining them is likely to better describe the actual scenario of hospitalizations.

The number of groups needed

Case mix tools for reimbursement face the challenge of finding the optimal balance in the trade-off between the number of groups and the efficiency of the tool (briefly, what percentage of these groups really contribute to explaining costs). In general, it is assumed that larger number of cost groups produce more accurate cost information and, statistically, increasingly larger number of groups are expected to yield better values of the reduction in variance. However, Krabbe-alkemade et al. noticed that, in some cases, a more fine-grained cost accounting system may not produce more accurate cost information due to measurement or classification errors [8]. Furthermore, they observed that higher number of groups with high reduction in variance may reduce efficiency in the sense of including many groups that do not match actual cases and could be, therefore, considered unnecessary. This was the case of our analysis, in which a relatively small APR-DRG groups covered a very high percentage of costs, suggesting that the inclusion of that many groups does not add much benefit to the overall estimate. On the other hand, the distribution of D_groups in the Queralt system showed high proportionality respect the percentage of costs and, more notably, hospitalizations, suggesting higher efficiency.

Numeric measures vs. groups

Regardless of the different efficiency and performance of risk groups of the compared tools, it is worth mentioning that the Queralt system can be used either as group or numerical index. This feature overcomes several drawbacks of forcing cases into groups, such as information loss, underestimation of the variation in outcome, and concealment of non-linearity between variables and outcomes [16]. The use of the numerical index resulted in a relatively good symmetry between the observed and estimated costs using linear models; in this analysis, the APR-DRG showed a discontinuous shape, probably due to the group-based approach.

Predicting costs

The broad range and heterogeneity of hospitalizations precludes an exact prediction of costs and, therefore, certain degree of disparity between predicted and actual costs must be assumed. In this regard, it is particularly important that these tools have high capacity to predict the highest costs. In our analysis, all tools performed well in predicting hospital costs at the 80th, 90th, and 95th

percentiles, with slightly better performance of the Queralt DP_Strata in the three scenarios. In the PR analysis, which reflects the level of false positive and false negative rates, the DP_Strata clearly outstood compared with all other models.

Limitations

Our analysis has some limitations that need to be mentioned. First, as all retrospective analyses based on information stored in electronic health records, the accuracy of our prediction depends on the quality of the data. However, since we are comparing two models using the same dataset, inaccuracies regarding resource use and diagnoses is expected to affect the two tools similarly. Regarding the origin of data, it is also worth mentioning that the Queralt system was developed using data from the same population—albeit collected in 2018. Thus, although costs were not considered in its development, an external validation, replicating the analysis in another population would increase the strength of our conclusions and inform on the generalizability of the model.

Concluding remarks

In summary, the Queralt system offers valuable tools not only for predicting hospital outcomes (as originally intended) but also for predicting hospitalization costs. The combination of diagnosis and procedures strata yields approximately 600 groups that accurately reflect resource utilization during hospital stay by considering the main diagnosis, patient complexity on admission, and procedures needed. The Queralt system also had higher efficiency, with better match with the case-mix of hospitalizations found in routine care. Although future studies assessing the performance of this tool in other countries are warranted, the Queralt system offer healthcare managers of systems with comprehensive collection of information regarding diagnoses and procedures with a free tool that accurately reflect hospital costs and can be tailored, if necessary, to the specific needs of each environment.

Abbreviations

APR-DRG	All-patient refined DRG
DRG	Diagnosis-related groups
ICS	Catalan Institute of Health
ICU	Intensive care unit
Queralt DxC	Queralt index for in-hospital complications
Queralt DxP	Queralt index for principal diagnosis
Queralt DxS	Queralt index for diagnoses on admission
Queralt Px	Queralt index for procedures

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13561-024-00522-6>.

Supplementary Material 1

Acknowledgements

The authors would like to thank Spyros Kolovos for providing insightful comments to the manuscript.

Author contributions

The study was originally conceived by DM. The study was designed by JF, DM, GC-S, IC, JP-J, and M-A. JF, EB, and DM performed the analysis. Data were provided by JP-J and MA. The manuscript was first drafted by JF, DM, and GC-S, and subsequently revised by JF, EB, DM, GC-S, IC, JP-J, and MA. All authors approved the final version of the manuscript and consented their submission.

Funding

The study did not receive specific funding.
Open Access funding provided by Servei Català de la Salut

Data availability

The dataset contains financial data from the Catalan Institute of Health and, therefore, is not publicly available. The software and users' manual for the Queralt system can be freely accessed from <https://ics.gencat.cat/ca/assistencia/coneixement-assistencial/sistema-de-Queralt/>.

Declarations

Ethics approval and consent to participate

Based on the only use of aggregated data, the Ethics Committee of the Bellvitge Hospital (Spain) deemed not necessary an approval of the study protocol.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests regarding this work.

Author details

¹Catalan Health Service, Gran Via de les Corts Catalanes 587, Barcelona 08007, Spain

²Digitalization for the Sustainability of the Healthcare System (DS3) – Institut d'Investigacions Biomèdiques de Bellvitge (IDIBELL), Barcelona, Spain

³Catalan Institute of Health, Barcelona, Spain

⁴Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer (FRCB- IDIBAPS), Universitat de Barcelona, Barcelona, Spain

⁵Faculty of Informatics, Telecommunications and Multimedia, Universitat Oberta de Catalunya, Barcelona, Spain

Received: 18 July 2023 / Accepted: 14 June 2024

Published online: 26 June 2024

References

1. Cylus J, Papanicolas I, Smith PC. Health system efficiency. How to make measurement matter for policy and management. European Observatory on

- Health Systems and Policies Series (WHO Partnership). 2016 [cited 2022 Oct 10]. https://www.euro.who.int/__data/assets/pdf_file/0004/324283/Health-System-Efficiency-How-make-measurement-matter-policy-management.pdf
2. Vela E, Clèries M, Vella VA, Adroher C, García-Altés A. Population-based analysis of the healthcare expenditure in Catalonia (Spain): what and who consumes more resources? *Gac Sanit*. 2019;33:24–31.
3. Bredekamp C, Bales S, Kahur K. Transition to diagnosis-related group payments for health: lessons from case studies. World Bank. 2020 [cited 2022 Sep 29]. <https://openknowledge.worldbank.org/handle/10986/33034>
4. Sharabiani MTA, Aylin P, Bottle A. Systematic review of comorbidity indices for administrative data. *Med Care*. 2012;50(12):1109–18.
5. Ash AS, Ellis RP, Pope GC, Ayanian JZ, Bates DW, Burstin H, et al. Using diagnoses to describe populations and predict costs. *Health Care Financ Rev*. 2000;21(3):7.
6. Huang X, Peterson S, Lavergne R, Ahuja M, Mcgrail K. Predicting the cost of health care services. *Med Care*. 2019;00(00):1–6.
7. Hammill BG, Hoffman MN, Clark AG, Bae JG, Shannon RP, Curtis LH. Incorporating acute conditions into risk-adjustment for provider profiling: the case of the US news and world report best hospitals rankings methodology. *Am Coll Med Qual*. 2024;39(2):69–77.
8. Krabbe-alkemade Y, Groot T. Performance differences between the episode-based DBC and diagnosis-related DRG case mix systems. *Manag Acc*. 2017;91:224–35.
9. Monrde D, Cainzos-Achirica M, Cossio-Gil Y, García-Eroles L, Pérez-Sust P, Arrufat M, et al. Performance of comprehensive risk adjustment for the prediction of in-hospital events using administrative healthcare data: the queralt indices. *Risk Manag Healthc Policy*. 2020;13:271–83.
10. Monrde D, Carot-Sans G, Cainzos-Achirica M, Abilleira S, Coca M, Vela E, et al. Performance of three measures of comorbidity in predicting critical covid-19: a retrospective analysis of 4607 hospitalized patients. *Risk Manag Healthc Policy*. 2021;14:4729–37.
11. Scheller-Kreinsen D, Geissler A, Busse R. The ABC of DRGs. *Euro Observer*. 2009.
12. Palmer G, Reid B. Evaluation of the performance of diagnosis-related groups and similar casemix systems: methodological issues. *Heal Serv Manag Res*. 2001;14(2):71–81.
13. Reid B, Palmer G, Aisbett C. The performance of Australian DRGs. *Aust Heal Rev*. 2000;23(2):20–31.
14. Lyons KW, Klare CM, Kunkel ST, Lemire JR, Bao M, McGuire KJ, et al. A 5-year review of hospital costs and reimbursement in the surgical management of degenerative spondylolisthesis. *Int J Spine Surg*. 2019;13(4):378–85.
15. Virk SS, Phillips FM, Moffatt-Bruce SD, Khan SN. What factors influence reimbursement for 1 to 2 level anterior cervical discectomy and fusion procedures? *Spine (Phila Pa 1976)*. 2019;44(1):E33–8.
16. Altman DG, Royston P. The cost of dichotomising continuous variables. *Br Med J*. 2006;332(7549):1080.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.