Note: Please note for the final poster the text will be shortened and all presented figures translated into English.

Title

Improving the information basis for decision makers: The critical validity of applying DRG-based patient severity indices for areas of application other than reimbursement[1].

Objective

In numerous cases patient severity indices are an important information basis for decision makers in the healthcare sector, however, in many instances they have been misapplied. For example, in face of increased pressure for outpatient treatment in Germany, the Patient Clinical Complexity Level has been employed to estimate the outpatient potential in German hospitals. This study aims to illustrate the risks of applying a risk-adjustment system for reimbursement in areas other than the intended one by comparing the performance of two sets of patient severity scales - reimbursement and clinical - applied to a clinical use case.

Research Design and Methods

In order to determine the outpatient potential of an anonymized German hospital (ca. 26,000 cases), we applied the following patient severity indices: Patient Clinical Complexity Level (PCCL), Disease Staging™ (DS)[2,3,4], Elixhauser Score (ES)[5], and the Charlson Index (CI)[6]. The former serves as a proxy for a patient severity index for reimbursement, the latter three are medical scales for clinical applications (ES and CI are the most adopted ones, while DS is the most comprehensive). Determining the outpatient potential of the sample hospital is our proxy for a clinical application. The employed methodology is descriptive statistics in the form of comparing frequency distributions of severity levels across those systems. Initially, we divide the PCCL according to the other scales to assess how comprehensive and detailed each system is in relative terms. When determining the outpatient potential, we adopt a prudent approach and set the cutoff values at the lowest level for each index.

It should be noted that a relative assessment between the clinical systems has some caveats, and a comparison of frequency distributions between them should be conducted with caution, as the results might be misleading (as opposed to the comparison between resource based and clinical based scales). An important distinction between DS and the other presented medical scales is that they are additive

---

[1] Based on the publication: Rothkopf, K.; Fischer, F. J.; Bareis, L.; Mraz, G.; (2023); Eignet sich der PCCL vor dem Hintergrund der Krankenhausreform und der zunehmenden Ambulantisierung als Parameter zur Abbildung der medizinischen, patientenindividuellen Fallschwere?; in: Fischer, FJ (Hrsg.), Risikoadjustierung und individualisierte Medizin, Kohlhammer-Verlag Stuttgart 2024.

[2] Gonella, J.S.; M.C. Hornbrook, D.Z. Louis; (1984), „Staging of Disease; A Case-Mix-Measurement"; Journal of the American Medical Association 251 (5), S. 637 -644

[3] www.jefferson.edu;

[4] https://hcup-us.ahrq.gov > nation> nis> disease Staging

[5] Eixhauser, A.; R.M. Andrews; Fox, S.; (1993); Clinical Classification for Health Policy Research: Discharge Statistics by Pricipal Diagnosis and Procedure; AHCPR Publication No. 93-004; Rockville MD: Agency for Healthcare Policy and Research, Public Helth Service

[6] Charlson, M.E.; Pompei, P; Ales, K.L.; MacKenzie, C.R.; (1987); A new Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation; in: Journal of Chronic Diseases 40(5); S. 373 – 383.

systems (a count of certain comorbidities) and treat every comorbidity equally (CI weights limited comorbidities to a small extent). This might result in somewhat arbitrary cutoff points, as it is unclear whether the presence of a more or less severe comorbidity is the determinant of a respective score, a differentiation is not possible. This means that for the clinical application at hand, only DS provides the most adequate assessment of the severity of individual comorbidities or the interaction thereof, which is to be expected as this is the most comprehensive system. For instance, DS classifies a certain type of mitral valve insufficiency as Stage 3.3 (one of the highest levels), where ES and CI only yield a score of 1 (the lowest score for covered comorbidities). This illustrates that many clinical scales cannot distinguish between individual comorbidities in terms of severity, an important aspect that needs to be taken into account when applying such scales to clinical problems. Further, there are significant differences in terms of the comprehensiveness between the systems. (17 diagnoses) compared to the ES (ca. 30 diagnoses) and the most comprehensive system DS (600 disease patterns and 5200 prognostic severity groups). For instance, 'chronic kidney disease, stage 5', or 'benign essential hypertension: with mention of a hypertensive crisis' are both classified as Stage 3.1 (high severity) for the DS scale, but are not considered (i.e. a score of 0) by the ES and CI.

Results

Fig. 2 shows the PCCL score of cases in an anonymized German hospital (approximately 26,000 cases), divided according to the respective medical scales. In more than half of the cases with a PCCL of 0 (lowest level), both DS and the ES classify cases into higher levels of severity. This suggests that these two systems provide a much more comprehensive and detailed assessment of clinical patient severity, also with regard to the extensive focus on the effect of comorbidities and interactions thereof (the latter only for DS). In comparison, the CI provides a much less detailed subdivision, reclassifying only about 20% of cases with a PCCL of 0. This is expected, as the CI considers by far the fewest number of diagnostic codes (see above).

In a further step, we can evaluate the collective declared as outpatient by the German AOP 2023 catalog concerning its actual outpatient potential according to each scale. In other words, we use different medical scales in addition to the official context factors (exclusion criteria) to assess, from a medical (diagnostic) perspective, how high the outpatient potential from the perspective of respective scales is. In contrast, we present the outpatient potential from the perspective of the PCCL and compare how homogeneously the respective scales indicate the outpatient potential.

In the examined hospital, ca. 30% of cases (~7000) fall under the inclusion criteria for outpatient treatment. When evaluating this collective according to the AOP 2023 context factors (exclusion criteria) and the PCCL on one side, and the medical scales on the other side, based on severity, the outpatient potential results in different collectives (Fig. 3). In relation to the collective determined by the AOP 2023 catalogue affected by the inclusion criteria, the AOP 2023 catalogue indicates an outpatient potential of 19% (~5000 cases). Based on this population, the PCCL (< 3[7]) indicates a potential of 18% (relative to the entire population, i.e., an improvement by 1%), and the medical scales DS, ES, and CI show potentials of ca. 10% (Stage 1), ca. 8% (= 0), and ca. 14% (= 0), respectively (Fig. 3). Comparing these percentages (DS and ES exclude more than twice as many cases as the PCCL, and the less comprehensive CI (ca. 25%) it becomes apparent that, regarding the studied scales, the PCCL is the least

---

[7] Cut-off value based on the application within the aG-DRG system ("ohne äußerst schwere CC")

comprehensive severity scale for the application at hand, as it does not consider relevant factors in assessing outpatient potential. For instance, cases with a Ventricular Tachycardia or Diabetes mellitus type 2 (decompensated) are classified as outpatient, which fall into higher severity levels according to the broader medical scales (e.g., DS Stage 2, not recognized by ES and CI).

Conclusion

The practice of applying a resource-based patient severity index to a clinical problem is associated with some challenges. For instance, when applied to determine the outpatient potential of a sample hospital, clinical scales exclude ca. twice as many cases as the considered resource-based index. This could be associated with tangible risks when decisions regarding clinical problems are based on misleading information. Hence, awareness regarding the correct application of risk adjustment tools should be considered an important constituent of good decision making in the healthcare sector.

Table 1: Comparison between the scales Disease Staging™ (clinical) and the Patient Clinical Complexity Level (resource)

	Disease Staging™ (DS)	PCCL
Präzise Diagnostik inkl. Einbezug der Ätiologie (DS kann bis zu 99 Diagnosen für die EW nutzen)	JA	NEIN
Die identifiziertem Fallschweregruppen sind klinisch homogen und prädiktiv	JA	NEIN
Die Behandlung selbst beeinflusst den Fallschweregrad	NEIN	JA
Das System kann ambulant wie stationär eingesetzt werden (z.B. zur Bestimmung einer rechtzeitigen Klinikeinweisung)	JA	NEIN
Der medizinische Ressourceneinsatz kann überprüft werden (Arzneimittel, Diagnostik, Blutprodukte etc.)	JA	JA
Überprüfung der Angemessenheit einer Behandlung	JA	NEIN
Einsatz über die Systemgrenzen (Data-Pooling) hinweg	JA	NEIN (nur Basis Algorithmus)

Figure 1: Case volume of the aG-DRG G67 according to the DRG severity levels (A/B/C) and Disease Staging™ - low concordance between resource-based and clinical-based case severity.
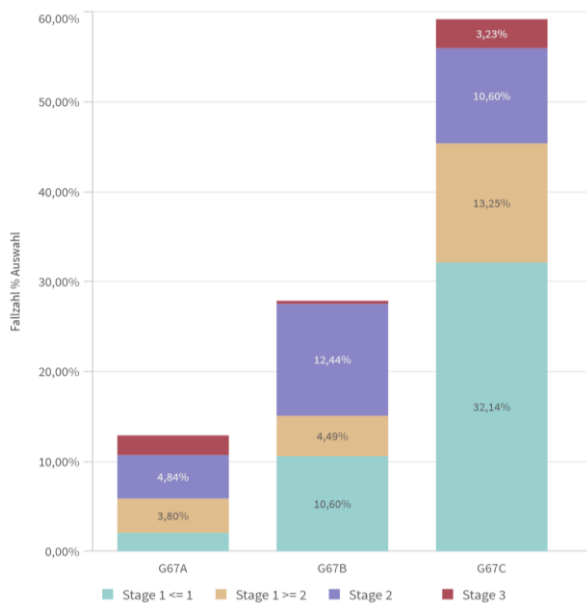
Figure 2: Frequency distribution of all cases from the sample hospital according to the PCCL and respective clinical scales
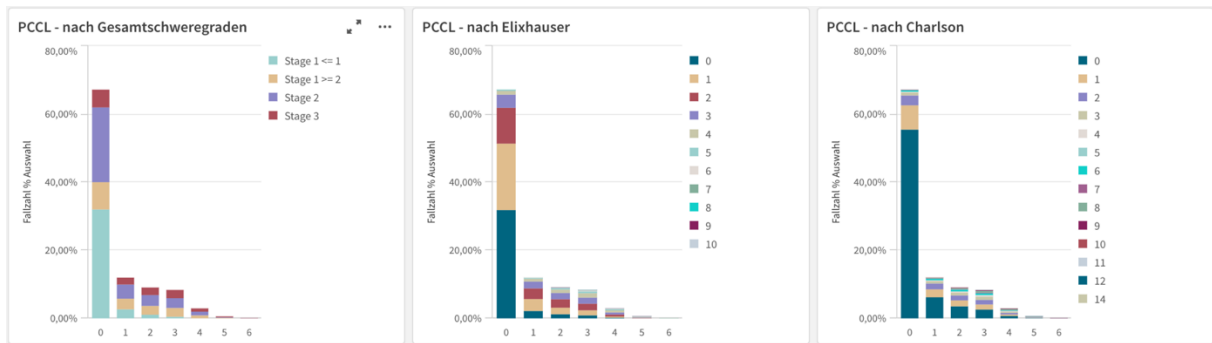

Figure 3: Frequency distribution of the officially recognized outpatient potential in Germany for the sample hospital based on the AOP 2023 catalogue according to studied scales