EJoHP 1(6)/2026
https://doi.org/10.21697/ejhp.0112.02
e-ISSN 2957-2002

2026, 6, 1: 17-29

Clinical Risk Prediction Tools for Hospitalization-related Outcomes Based on Administrative Health Data: A Mini Narrative Review

Agnieszka Mazurek, MD
School of Public Health
Centre of Postgraduate Medical Education
ORCID: https://orcid.org/0009-0002-1176-4076
Email: agnieszka.mazurek@cmkp.edu.pl

Received: 3 May 2026; Revised: 29 May 2026; Accepted: 16 June 2026

Abstract

Risk prediction tools for hospitalization-related outcomes can be used clinically for care planning and operationally for resource allocation and quality-of-care evaluation. Administrative health data, collected routinely for billing, claims, statistics, or quality monitoring, can serve as an inexpensive and generalizable source of data for the development of risk prediction tools. The purpose of this narrative review was to evaluate tools developed from administrative health data and designed primarily to predict hospitalization-related outcomes. PubMed, Web of Science, and Scopus were searched for studies published between January 1, 2016 and April 17, 2026. Ultimately, 10 studies met the criteria for inclusion. Most tools were developed using logistic regression and converted into risk scores. One study implemented a random forest model as a web-based calculator. In general, the tools predicting in-hospital mortality showed the best discrimination, whereas readmission tools had the worst performance. Although tools based on administrative data may be valuable for risk stratification, further prospective clinical validation and an assessment of potential bias are needed before routine implementation.

Keywords: risk stratification; hospitalization outcomes; administrative data; logistic regression; risk scores

Introduction

Tools for predicting the risk of adverse hospitalization outcomes, such as hospital-acquired complications, in-hospital mortality, prolonged length of stay, and unplanned readmission, have many practical applications. These tools can be used by hospital administrators and clinicians to support resource allocation and care planning (Zaccardi et al., 2017; Ar Rochmah et al., 2026) and by researchers and policymakers to evaluate the quality of care between hospitals and healthcare systems (Wada et al., 2017; Uematsu et al., 2021).

Many of the tools currently used in clinical practice are in the form of scores. Most of these scores were developed to quantify illness severity, but some have also been shown to correlate with the risk of adverse hospitalization outcomes. Examples include APACHE II (Yelamanchi, 2023), APACHE IV (Zimmerman et al., 2006), CURB-65 (Lim et al., 2003), and National Early Warning Score 2 (NEWS2) (Welch et al., 2022). However, these scores are intended for acutely ill patients and require the targeted collection of vital signs and laboratory measurements, which may hinder their use in settings where such data are not routinely collected for all patients on admission (Ar Rochmah et al., 2026).

One source of routinely collected information that can be used for systematic risk prediction is administrative health databases. Administrative data are typically collected for billing and claims, statistics, or quality monitoring, making their use for secondary analysis relatively inexpensive compared with primary data collection (Emerson et al., 2024). These data often include variables that are consistently recorded across many hospitals and available regionally, or even nationally. Therefore, prediction tools developed from administrative data may provide greater generalizability than tools developed from data collected from only a small number of hospitals, where the study population may be narrower in terms of clinical severity, age, ethnicity, or socioeconomic status (Singh et al., 2022).

Some existing tools use variables that are typically found in administrative data sources. Perhaps the most prominent example is the Charlson Comorbidity Index (CCI), which was not originally developed using administrative data, but was later adapted for use with administrative diagnostic codes and has been shown to predict various hospitalization outcomes (Librero et al., 1999). Another example is the LACE index, which was developed using a combination of administrative and clinical data but was validated for use with administrative data alone (Walraven et al., 2010). However, of these two indices, only LACE was developed specifically as a direct prediction tool for hospitalization-related outcomes.

Therefore, the aim of this narrative review was to identify and describe readily usable clinical risk prediction tools developed using administrative health data and designed primarily to predict hospitalization-related outcomes.

Methods

PubMed, Web of Science, and Scopus were searched for original articles on tools for predicting the risk of adverse hospitalization outcomes based on administrative data. The outcomes of interest were hospital-acquired complications, prolonged length of stay, mortality, or unplanned readmission. Administrative data were defined as primarily non-clinical data collected routinely for the purposes of billing, claims, statistics, or other administrative activities and used for secondary analysis.

The search was restricted to studies published between January 1, 2016 and April 17, 2026. The following keywords were used: machine learning, artificial intelligence, regression, prediction model, prediction rule, risk score, point score, nomogram, calculator, scoring system, hospitalization, inpatient, in-hospital, administrative data, billing data, claims data, registry, register-based, mortality, length of stay, readmission, complication, and outcome.

A total of 717 studies were identified during the initial search; these were then assessed by title and abstract against the inclusion and exclusion criteria. Studies were included if they were written in English, provided an actual risk prediction tool, and had a stated use case of predicting a hospitalization-related outcome. Conversely, studies were excluded if no ready-to-use tool was developed. Studies that provided model artifacts without an accompanying tool, or logistic regression coefficients without an explicit description of how to combine them into a risk score were not included. Comorbidity indexes whose main stated utility was to adjust for case mix in prediction models were also excluded.

Following title and abstract screening, the remaining articles were subjected to a full-text review to decide on final eligibility. Given the narrative nature of this literature review, no formal screening framework was used. The author conducted the article screening independently. In cases of uncertainty, the article was moved to the end of the screening queue and reassessed after the remaining articles had been screened. In total, 10 studies were determined to meet all inclusion and exclusion criteria and were selected to form the sample for this narrative literature review.

Results

Populations, Clinical Contexts, and Outcomes

The 10 studies included in this review are presented in Table 1. Overall, the studies encompassed several different populations, clinical conditions, and hospitalization-related outcomes across three continents. Two studies included broad inpatient populations: one examined all acute-care admissions and the other all medical (i.e., non-surgical) admissions. One study included all older adult admissions (≥75 years old). Three of the studies that examined specific clinical conditions focused on acute conditions, two on chronic conditions, and two on acute presentations among patients with chronic disease. Five studies included adults only, while another five included all hospitalizations irrespective of patient age. The most commonly predicted outcome was in-hospital mortality (6 studies), followed by prolonged length of stay (3 studies), 30-day readmission (3 studies), and 30-day mortality (1 study).

Table 1. Summary of studies on the development of risk prediction tools using administrative data

Authors, Year	Country	Population	Outcome	Data Source	Algorithm	Tool
Zhang et al., 2026	USA	Adults with chronic lymphocytic leukemia/small lymphocytic lymphoma	In-hospital mortality	National Inpatient Sample	Logistic regression	Score
Ar Rochmah et al., 2026	Indo-nesia	Adults with type 2 diabetes and stroke	In-hospital mortality	Claims-based diabetes registry at Dr. Sardjito General Hospital, Yogyakarta	Random forest	Web--based calculator
Uematsu et al., 2021	Japan	Adult patients with community-acqui-red pneumonia	Prolonged length of stay	Quality Indicator/ Improvement Project/ Diagnosis Procedure Combination	Logistic regression	Score
Doctoroff and Herzig, 2020	USA	Adult medical hospitalizations	Prolonged length of stay	National Inpatient Sample	Logistic regression	Score
Pauly et al., 2019	France	All acute-care hospitalizations	30-day readmission[a]	Programme de Médicalisa-tion des Systèmes d'Information/Assistance Publique-Hôpitaux de Marseille	Logistic regression	Score
Gilbert et al.	England	All patients aged 75 years and older	30-day mortality/ prolonged length of stay/30-day readmission[a]	Hospital episode statistics	Logistic regression	Score
Zaccardi et al., 2017	England	All patients with type 1 and type 2 diabetes and hypoglycemia	In-hospital mortality/30--day readmission[b, c]	Hospital episode statistics	Logistic regression	Excel calculator
Wada et al., 2017	Japan	All patients with injury (trauma)	In-hospital mortality	Diagnosis Procedure Combination	Logistic regression	Score

Authors, Year	Country	Population	Outcome	Data Source	Algorithm	Tool
Sato et al., 2017	Japan	All patients with liver cirrhosis hospitalized for elective surgery	In-hospital mortality	Diagnosis Procedure Combination	Logistic regression	Score / Mobile app
Qu et al., 2016	China	All patients with acute myocardial infarction	In-hospital mortality	Discharge summary reports	Logistic regression	Score

Footnotes: [a] Unplanned 30-day readmission; [b] 30-day readmission for hypoglycemia only; [c] 24-hour discharge was also evaluated as an outcome, but this outcome was excluded from this review because it was not adverse.

Data Sources and Available Variables

Across the ten studies, seven different sources of administrative data were used. Most datasets contained basic demographic characteristics, such as age and sex, as well as hospitalization characteristics, including admission mode, admission and discharge dates, discharge disposition, ICD diagnosis codes for primary and secondary diagnoses, and care-related information, such as procedures. Although all studies used data collected primarily for administrative purposes, some data sources included limited clinical information. The Japanese Diagnosis Procedure Combination database included the Child–Pugh score for patients with liver cirrhosis (Sato et al., 2017). The Japanese Quality Indicator/Improvement Project database included body mass index, Japan Coma Scale, Barthel index, and A-DROP pneumonia severity scores (Uematsu et al., 2021). The French database included a disease severity measure (Pauly et al., 2019). Socioeconomic information was also available in several databases. For instance, the English Hospital Episode Statistics included the Index of Multiple Deprivation (Zaccardi et al., 2017). Therefore, the contents of administrative data were not uniform across studies.

Modeling Approaches and Algorithm Selection

Logistic regression was the most commonly used modeling method in the studies selected for the review. All 10 studies tested logistic regression and eight used it exclusively. The two most recently published studies also evaluated more complex algorithms, including random forest, XGBoost, k-nearest neighbors, elastic net, LightGBM, support vector machines, and multilayer perceptrons (Ar Rochmah et al., 2026; Zhang et al., 2026). However, only the study predicting in-hospital mortality in stroke patients with diabetes deemed a machine-learning algorithm (random forest) to be clearly superior

(Ar Rochmah et al., 2026). In contrast, the study developing a model for in-hospital mortality in chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL) found that LightGBM provided only minimal improvements in area under the curve (AUC) and sensitivity, while reducing specificity and precision (Zhang et al., 2026). For this reason, the authors considered logistic regression to be the preferred algorithm and noted that one of the major advantages advanced algorithms have over logistic regression is that they can effectively capture non-linear relationships between predictors, but that this advantage may be reduced when the available data consist mostly of sparse, binary variables (Zhang et al., 2026).

Types of Tools and Development Methods

The most straightforward way to translate a prediction model into a usable clinical prediction tool is to create a calculator that applies the underlying logistic regression equation or model artifact directly. This approach was used by Zaccardi et al. (2017), who developed a simple, logistic-regression-based calculator in Microsoft Excel for in-hospital mortality in patients with diabetes admitted for hypoglycemia. Similarly, Ar Rochmah et al. (2026) chose a web-based calculator for their random forest model for predicting in-hospital mortality in patients with type 2 diabetes admitted for stroke.

The remaining eight studies developed risk scores from logistic regression models. This approach is understandable, given that risk scores are familiar in medicine and may therefore be more readily adopted in clinical practice. The methods used to convert models into scores differed across studies. The plainest approach was taken by Wada et al. (2017), who simply summed the regression (β) coefficients derived from the model to produce the final risk score. Gilbert et al. (2018) multiplied the coefficients by 5 before summing them. Sato et al. (2017) rounded the regression coefficients to whole numbers and summed them to obtain a risk score. Uematsu et al. (2021) used a similar approach, but multiplied the coefficients by 10 before rounding them. Qu et al. (2016) used the same approach as Sato et al. (2017), but with odds ratios instead of regression coefficients. Zhang et al. (2026) used a more elaborate method that involved scaling the non-zero regression coefficients by a constant and rounding them to whole numbers, then further setting the negative points to zero in order to produce a non-negative final score. Likewise, Doctoroff and Herzig (2020) transformed and rescaled the odds ratios into non-negative integer points for summation into a final score. Pauly et al. (2019) similarly rescaled the regression coefficients so that the final summed score ranged from 0 to 100.

Predictive Performance

Any direct comparison of the ten risk prediction tools should be approached with caution, since these tools were designed to predict different outcomes in different patient populations. However, it is noteworthy that in terms of discrimination (measured by AUC or the C-statistic), the tools predicting in-hospital mortality appeared to perform best, with most reporting values above 0.80. Tools predicting prolonged length of stay had intermediate performance (AUCs of 0.80 or slightly below) and tools predicting 30-day readmission performed worst (AUCs below 0.75). The performance for predicting 30-day readmission in patients with diabetes hospitalized for hypoglycemia was so poor (AUC = 0.55) that the authors refrained from developing a calculator for that outcome (Zaccardi et al., 2017).

A more informative comparison may be performance relative to existing, clinically validated risk scores and indices. For in-hospital mortality, the risk score for patients hospitalized for acute myocardial infarction outperformed the CCI (AUC = 0.824 vs. 0.710) (Qu et al., 2016); the ADOPT-LC score for patients with cirrhosis undergoing major surgery outperformed the Child–Pugh score (AUC = 0.881 vs. 0.803) (Sato et al., 2017); and the risk score for patients hospitalized for trauma moderately outperformed the multiplicative ICD-10-based Injury Severity Score and the single worst injury score (AUC = 0.833 vs. 0.825 and 0.822, respectively) (Wada et al., 2017). The risk score predicting 30-day readmission after acute hospitalization, despite generally modest performance, still outperformed the LACE index (AUC = 0.740 vs. 0.660) (Pauly et al., 2019).

Even though AUC is a useful metric for assessing model performance in general, it does not provide much information about clinical performance, such as how many patients with the outcome would be missed and how many patients without the outcome would be mistakenly identified as high-risk (Zhang et al., 2026). Therefore, other metrics such as sensitivity, specificity, positive predictive value (PPV), negative predictive value, and calibration should also be taken into consideration for a thorough assessment of predictive performance. For example, both the calculator for risk of in-hospital death in patients with type 2 diabetes and stroke and the ADOPT-LC score had high AUCs and high sensitivity (above 90%) but lower specificity, where approximately 30%–36% of patients without the outcome would be classified as false positives (Sato et al., 2017; Ar Rochmah et al., 2026). Similarly, the risk score for prolonged hospitalization among general medical patients had an AUC of 0.80 but a PPV of 15%, which means that most patients classified as high-risk would not actually experience prolonged hospitalization (Doctoroff and Herzig, 2020). The risk score for in-hospital mortality in patients with CLL/SLL, despite an AUC of 0.842, was found to underestimate the risk for very high-risk patients and overestimate it for very low-risk patients due to less-than-ideal calibration (Zhang et al., 2026).

Discussion

Clinical and Operational Utility

The primary application of risk prediction tools in clinical practice is risk stratification. Patients can be assessed on admission for risk of an adverse hospitalization outcome, and appropriate measures can be taken by clinicians to reduce the risk of those outcomes. Tools predicting the risk of in-hospital mortality can be used to identify the need for more frequent monitoring of patients and plan for escalation of care, including the preparation of adequate staff and physical resources, such as beds and equipment (Zhang et al., 2026). Tools predicting prolonged length of stay can be used by hospital administrators to forecast bed use and by clinicians to employ strategies to reduce length of stay, such as the early planning of appropriate post-discharge placement (Doctoroff and Herzig, 2020; Uematsu et al., 2021). Finally, tools predicting readmission can help identify patients who might benefit from additional medical and social support to facilitate the transition from the hospital back to the community (Pauly et al., 2019).

Risk prediction tools have many other uses beyond population-level risk stratification. One such use is estimating an individual's probability of an adverse outcome. Zaccardi et al. (2017) noted that their calculator for in-hospital mortality in patients with diabetes hospitalized for hypoglycemia can be used to inform individual care planning and strategies to reduce mortality risk. Furthermore, risk scores can be used to support clinical decision-making, such as preoperative assessments of whether surgery should be performed, as proposed by Sato et al. (2017) for their ADOPT-LC score. Risk scores can also be used to compare outcomes between hospitals to evaluate the quality of care. Qu et al. (2016) proposed such a use case for their risk score that predicts in-hospital mortality in patients admitted for AMI. In addition, risk scores can be used to adjust for case mix in retrospective observational studies utilizing administrative data, as described by Qu et al. (2016), as well as Wada et al. (2017) for their risk score that predicts in-hospital mortality in patients admitted for trauma.

Clinical and Methodological Limitations

Despite their potential clinical utility, risk prediction tools based on administrative data are intended to support rather than replace clinicians in the clinical decision-making process (Zaccardi et al., 2017; Zhang et al., 2026). Several authors reported performance metrics indicating that their tools could yield a non-negligible number of false negatives and false positives, which could lead to erroneous decisions if the results are not interpreted in the context of the patient's clinical presentation (Zaccardi et al., 2017). This is especially relevant for outcomes such as readmission, for which administrative data may not capture enough information to allow for reliable predictions (Zaccardi et al., 2017).

Clinical risk prediction tools would arguably be most useful if they could predict the risk of adverse outcomes of hospitalization at the time of hospital admission. However, with administrative data, it is often not clear whether certain information was available on admission or whether it was obtained later in the course of hospitalization (Doctoroff and Herzig, 2020; Zhang et al., 2026). In some cases, the authors included variables that would not be known until later, even close to the outcome. For example, Pauly et al. (2019) included discharge destination to predict 30-day readmission and noted that this could preclude predictions early enough to plan protective interventions. In the calculator for predicting in-hospital mortality among patients with type 2 diabetes hospitalized for stroke, Ar Rochmah et al. (2026) included length of stay. Although the authors intended for the calculator to be used throughout hospitalization (Ar Rochmah et al., 2026), this raises concerns because the length of stay is closely tied to its course, and it is unclear how far in advance the calculator could provide an actionable warning of mortality.

The choice of features can also lead to tools that produce biased results. Patient sex was a frequently used predictor and male sex was assigned higher scores in most cases. This could be indicative of greater susceptibility of male patients to certain complications or deterioration, but it could also potentially lead to biased allocation of resources toward male patients. Biased results can also be a concern when race and ethnicity are used as predictors. Zaccardi et al. (2017) included ethnicity as one of the variables; ethnicity other than White was associated with a lower calculated risk of in-hospital mortality, raising the possibility that risk could be underestimated in minority ethnic groups. In contrast, although Doctoroff and Herzig (2020) had access to information on race in their data source, they explicitly stated that they chose not to use it so as to avoid guiding hospital activities by patients' race. Zhang et al. (2026) were the only authors to perform subgroup analyses to ensure that their model performed comparably across different sociodemographic groups.

The quality of administrative data can also be a source of bias, both at the tool development stage and the clinical use stage. Data collected for administrative purposes are often affected by missingness (Doctoroff and Herzig, 2020; Ar Rochmah et al., 2026), as well as inadequate or erroneous coding (Wada et al., 2017; Zaccardi et al., 2017) and local coding practices (Gilbert et al., 2018; Zhang et al., 2026). Diagnoses are typically coded using ICD codes, which have been shown to have poor sensitivity for certain common conditions, such as obesity (Suissa et al., 2021) and chronic kidney disease (Ronksley et al., 2012). Furthermore, conditions that are present on admission but considered relevant to the index hospitalization – and therefore do not directly contribute to hospital reimbursement – may be less likely to be accounted for in the administrative record (Sato et al., 2017; Gilbert et al., 2018). Sato et al. (2017) suggested that this could give rise to predictive bias, as the number of recorded comorbidities for each patient may be more reflective of their clinical state during the index hospitalization than of their true overall comorbidity burden.

Future Directions

Prospective validation studies should be conducted in clinical settings and among different demographic and socioeconomic groups, including age, sex, race, ethnicity, and socioeconomic status, in order to minimize the risk of producing biased predictions (Gilbert et al., 2018; Zhang et al., 2026). Future papers should also investigate whether more advanced machine learning algorithms could improve the performance of the tools originally developed using logistic regression. Interestingly, Jaotombo et al. (2020) – which include Pauly and several coauthors of the risk score developed to predict 30-day readmission after acute hospitalization – later used the same dataset to test several tree-based algorithms and neural networks in order to determine whether they could improve its performance. They found that random forest improved the AUC to 0.790, compared with 0.740 for logistic regression and the original risk score.

Considering the significant variation in the content of administrative health data across different healthcare systems, prediction tools intended for use in Poland should ideally be developed using Polish data sources. In Poland, two of the most comprehensive sources of administrative health data that could potentially be used to predict hospitalization outcomes, such as in-hospital mortality and length of stay, are the National Health Fund's claims and settlement data and the Nationwide General Hospital Morbidity Study, maintained by the National Institute of Public Health – National Institute of Hygiene (Poznańska et al., 2019). Other national sources, such as the e-Health System (P1) (Wierzbicki et al., 2024), could be useful for assessing other outcomes such as readmission. The nationwide scope of these sources could support the development of tools to be applied across the entire Polish healthcare system. Nevertheless, successful development and implementation would require close cooperation between data administrators, researchers, hospital administrators, and hospital staff.

Limitations of This Review

The main limitation of this review was its narrow scope. The goal was to evaluate studies describing the development of actual tools for risk prediction of hospitalization-related outcomes. Therefore, studies that did not provide a readily usable tool were excluded. This means that the review likely omitted many relevant studies on risk prediction models that could, in principle, be converted into calculators, assuming that the model coefficients or artifacts were provided. In addition, it only included studies that explicitly identified their data source as administrative. This limitation was partially mitigated by the fact that the search strategy was not limited to the term "administrative data" but incorporated other terms that could refer to administrative health data sources. Nevertheless, it is possible that a more thorough manual review of data sources might have yielded additional eligible studies. Lastly, the selection of studies may have been inadvertently biased due to the author's subjective decisions, as the article screening process was not standardized and was conducted by a single researcher.

Conclusion

In conclusion, risk prediction tools based on administrative data can be valuable for stratifying patients by risk of adverse hospitalization outcomes. In the past decade, several tools have been proposed for a range of patient populations and outcomes, and some were shown to outperform existing tools. However, further research is needed to validate these tools before they can be implemented in clinical practice. The findings of this review may be valuable to hospital administrators and clinicians seeking to understand how risk prediction tools based on administrative data could support routine clinical work and improve patient outcomes, as well as to public health researchers and policymakers investigating which hospitalization-related outcomes may be feasibly predicted using routinely collected administrative data.

Research funding: This publication did not receive any external financial support.

Statement from the relevant ethics committee: The study was not reviewed by an ethics committee. Participation in the study was voluntary. All participants were informed about the purpose of the study, confidentiality policies, and the option to withdraw from participation at any stage. Informed consent was obtained from the participants for their participation in the study.

Conflict of interest: The authors declare that there is no conflict of interest related to the publication of this study.

References

Ar Rochmah, M., Nugroho, D. B., Gofir, A., Ikhsan, M. R., Hanif, F., Khairani, A. F., Chandra, L. A., & Setyopranoto, I. (2026). In-hospital mortality predictors of stroke patients with diabetes mellitus. *Journal of Neurosciences in Rural Practice*, *17*(1), 64–72. Retrieved from https://doi.org/10.25259/JNRP_85_2025

Doctoroff, L., & Herzig, S. J. (2020). Predicting patients at risk for prolonged hospital stays. *Medical Care*, *58*(9), 778. Retrieved from https://doi.org/10.1097/MLR.0000000000001345

Emerson, S. D., McLinden, T., Sereda, P., Yonkman, A. M., Trigg, J., Peterson, S., Hogg, R. S., Salters, K. A., Lima, V. D., & Barrios, R. (2024). Secondary use of routinely collected administrative health data for epidemiologic research: Answering research questions using data collected for a different purpose. *International Journal of Population Data Science*, *9*(1). Retrieved from https://doi.org/10.23889/ijpds.v9i1.2407

Gilbert, T., Neuburger, J., Kraindler, J., Keeble, E., Smith, P., Ariti, C., Arora, S., Street, A., Parker, S., Roberts, H. C., Bardsley, M., & Conroy, S. (2018). Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: An observational study. *The Lancet*, *391*(10132), 1775–1782. Retrieved from https://doi.org/10.1016/S0140-6736(18)30668-8

Jaotombo, F., Pauly, V., Auquier, P., Orleans, V., Boucekine, M., Fond, G., Ghattas, B., & Boyer, L. (2020), Machine-learning prediction of unplanned 30-day rehospitalization using the French hospital medico-administrative database. *Medicine*, *99*(49), e22361. Retrieved from https://doi.org/10.1097/MD.0000000000022361

Librero, J., Peiró, S., & Ordiñana, R. (1999). Chronic comorbidity and outcomes of hospital care: Length of stay, mortality, and readmission at 30 and 365 days. *Journal of Clinical Epidemiology*, *52*(3), 171–179. Retrieved from https://doi.org/10.1016/S0895-4356(98)00160-7

Lim, W. S., Eerden, M. M. van der, Laing, R., Boersma, W. G., Karalus, N., Town, G. I., Lewis, S. A., Macfarlane, J. T. (2003). Defining community acquired pneumonia severity on presentation to hospital: An international derivation and validation study. *Thorax*, *58*(5), 377–382. Retrieved from https://doi.org/10.1136/thorax.58.5.377

Pauly, V., Mendizabal, H., Gentile, S., Auquier, P., & Boyer, L. (2019). Predictive risk score for unplanned 30-day rehospitalizations in the French universal health care system based on a medico-administrative database. *PLOS ONE*, *14*(3), e0210714. Retrieved from https://doi.org/10.1371/journal.pone.0210714

Poznańska, A., Goryński, P., Seroka, W., Stokwiszewski, J., Radomski, P., & Wojtyniak, B. (2019). Nationwide General Hospital Morbidity Study as a source of data about Polish population health. *Przegląd Epidemiologiczny – Epidemiological Review*, *73*(1), 69–80. Retrieved from https://doi.org/10.32394/pe.73.08

Qu, Z., Zhao, L. P., Ma, X., & Zhan, S. (2016). Building a patient-specific risk score with a large database of discharge summary reports. *Medical Science Monitor*, 22, 2097–2104. Retrieved from https://doi.org/10.12659/MSM.899262

Ronksley, P. E., Tonelli, M., Quan, H., Manns, B. J., James, M. T., Clement, F. M., Samuel, S., Quinn, R. R., Ravani, P., Brar, S. S., Hemmelgarn, B. R., & Alberta Kidney Disease Network. (2012). Validating a case definition for chronic kidney disease using administrative data. *Nephrology Dialysis Transplantation*, *27*(5), 1826–1831. Retrieved from https://doi.org/10.1093/ndt/gfr598

Sato, M., Tateishi, R., Yasunaga, H., Horiguchi, H., Matsui, H., Yoshida, H., Fushimi, K., & Koike, K. (2017). The ADOPT-LC score: A novel predictive index of in-hospital mortality of cirrhotic patients following surgical procedures, based on a national survey. *Hepatology Research*, *47*(3), E35–E43. Retrieved from https://doi.org/10.1111/hepr.12719

Singh, H., Mhasawade, V., & Chunara, R. (2022). Generalizability challenges of mortality risk prediction models: A retrospective analysis on a multi-center database. *PLOS Digital Health*, *1*(4), p. e0000023. Retrieved from https://doi.org/10.1371/journal.pdig.0000023

Suissa, K., Schneeweiss, S., Lin, K. J., Brill, G., Kim, S. C., & Patorno, E. (2021). Validation of obesity-related diagnosis codes in claims data. *Diabetes, Obesity and Metabolism*, *23*(12), 2623–2631. Retrieved from https://doi.org/10.1111/dom.14512

Uematsu, H., Yamashita, K., Kunisawa, S., & Imanaka, Y. (2021). Prediction model for prolonged length of stay in patients with community-acquired pneumonia based on Japanese administrative data. *Respiratory Investigation*, *59*(2), 194–203. Retrieved from https://doi.org/10.1016/j.resinv.2020.08.005

Wada, T., Yasunaga, H., Yamana, H., Matsui, H., Matsubara, T., Fushimi, K., & Nakajima, S. (2017). Development and validation of a new ICD-10-based trauma mortality prediction scoring system using a Japanese national inpatient database. *Injury Prevention*, *23*(4), 263–267. Retrieved from https://doi.org/10.1136/injuryprev-2016-042106

Walraven, C. van, Dhalla, I. A., Bell, C., Etchells, E., Stiell, I. G., Zarnke, K., Austin, P. C., & Forster, A. J. (2010). Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *Canadian Medical Association Journal*, *182*(6), 551–557. Retrieved from https://doi.org/10.1503/cmaj.091117

Welch, J., Dean, J., & Hartin, J. (2022). Using NEWS2: An essential component of reliable clinical assessment. *Clinical Medicine*, *22*(6), 509–513. Retrieved from https://doi.org/10.7861/clinmed.2022-0435

Wierzbicki, M. P., Jantos, B. A., & Tomaszewski, M. (2024). A review of approaches to standardizing medical descriptions for clinical entity recognition: Implications for artificial intelligence implementation. *Applied Sciences*, *14*(21), 9903. Retrieved from https://doi.org/10.3390/app14219903

Yelamanchi, R. (2023). The APACHE II scoring systems and the ICU. In *Biomarkers in Trauma, Injury and Critical Care*. Springer, pp. 1073–1086. Retrieved from https://doi.org/10.1007/978-3-031-07395-3_52.

Zaccardi, F., Webb, D. R., Davies, M. J., Dhalwani, N. N., Gray, L. J., Chatterjee, S., Housley, G., Shaw, D., Hatton, J. W., & Khunti, K. (2017). Predicting hospital stay, mortality and readmission in people admitted for hypoglycaemia: Prognostic models derivation and validation. *Diabetologia*, *60*(6), 1007–1015. Retrieved from https://doi.org/10.1007/s00125-017-4235-1

Zhang, X., Bustamante, J., Stefani, H., Lee, H., Xiao, N., Xiao, L., Soni, A., & Danpanichkul, P. (2026). Development and temporal validation of an interpretable point score for in-hospital mortality in CLL/SLL using the U.S. National Inpatient Sample, 2016–2022. *Leukemia Research*, 164, 108199. Retrieved from https://doi.org/10.1016/j.leukres.2026.108199

Zimmerman, J. E., Kramer, A. A., McNair, D. S., Malila, F. M., & Shaffer, V. L. (2006). Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV*. *Critical Care Medicine*, *34*(10), 2517. Retrieved from https://doi.org/10.1097/01.CCM.0000240233.01711.D9