

# Development and Validation of an EHR-based Algorithm for Identifying *Pneumocystis jirovecii* Pneumonia

YU KITAMURA<sup>1</sup>, MASAMI TSUCHIYA<sup>2</sup>, HIROTOSHI IIHARA<sup>3,4,5</sup>, SHUJI SAKURAI<sup>3</sup>,  
JUNKI ENDO<sup>1</sup>, KEISUKE KUMADA<sup>4</sup> and YUKARI TSUBATA<sup>1,5</sup>

<sup>1</sup>Department of Respiriology, Graduate School of Medicine, Gifu University, Gifu, Japan;

<sup>2</sup>Division of Drug Informatics, Keio University Faculty of Pharmacy, Tokyo, Japan;

<sup>3</sup>Department of Pharmacy, Gifu University Hospital, Gifu, Japan;

<sup>4</sup>Patient Safety Division, Gifu University Hospital, Gifu, Japan;

<sup>5</sup>Cancer Center, Gifu University Hospital, Gifu, Japan

## Abstract

**Background/Aim:** *Pneumocystis jirovecii* pneumonia (PCP) remains a life-threatening opportunistic infection in patients receiving chemotherapy and other immunosuppressive cancer treatments. Accurate identification of true PCP cases within real-world electronic health record (EHR) databases is essential for epidemiological research and optimization of prophylactic strategies in oncology practice. The aim of this study was to develop and validate a practical, EHR-based algorithm for reliably identifying PCP cases.

**Patients and Methods:** This retrospective, single-center validation study used EHR data from a Japanese university hospital between April 2022 and March 2024. Adult patients ( $\geq 20$  years) who were assigned an ICD-10 code for PCP were extracted, and true cases were confirmed by a detailed review of the patient records. Seven candidate algorithms combining diagnostic codes, therapeutic-dose anti-PCP prescriptions, laboratory testing, chemotherapy exposure, and prescription duration were evaluated. The positive predictive value (PPV) and capture rate were then calculated using chart-confirmed PCP as the reference standard.

**Results:** Among 617 ICD-coded patients, 11 (1.8%) were confirmed as true PCP cases. The PPV of diagnostic codes alone was 1.8%. A prescription-enhanced algorithm (A1) identified 12 patients, including 11 true cases (PPV=91.7%; capture rate 100%). Algorithms incorporating  $\beta$ -D-glucan or PCR testing achieved PPVs of 100% with lower capture rates (63.6-81.8%). Incorporation of concurrent chemotherapy also resulted in a PPV of 100% with reduced capture. An algorithm requiring therapeutic-dose prescription for  $\geq 21$  days showed equivalent performance to A1.

*continued*



Masami Tsuchiya, PhD, Division of Drug Informatics, Keio University Faculty of Pharmacy, 1-5-30 Shibakoen, Minato-ku, Tokyo, 105-8512, Japan. E-mail: m-tsuchiya@keio.jp

Received February 6, 2026 | Revised March 8, 2026 | Accepted March 16, 2026



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.

©2026 The Author(s). Anticancer Research is published by the International Institute of Anticancer Research.

**Conclusion:** Prescription-based algorithms substantially improve the accuracy of PCP case identification in EHR data compared with diagnostic codes alone. This straightforward, scalable approach offers a robust framework for real-world oncology research, enabling a more reliable evaluation of PCP incidence and informing future prophylaxis strategies for patients receiving anticancer treatment.

**Keywords:** *Pneumocystis jirovecii* pneumonia, electronic health records, algorithm validation, real-world data, positive predictive value.

## Introduction

*Pneumocystis jirovecii* pneumonia (PCP) is an opportunistic fungal infection that predominantly affects immunocompromised individuals. This population includes individuals receiving chemotherapy or immunosuppressive agents, as well as those with hematologic malignancies and other conditions associated with impaired immunity. Although prophylactic strategies have substantially reduced the incidence of PCP, the infection continues to pose a serious clinical threat to immunocompromised individuals (1). Current international guidelines recommend trimethoprim-sulfamethoxazole (TMP-SMX) prophylaxis for individuals receiving rituximab-based therapy or purine analogs, as these treatments are associated with a substantially increased risk of PCP (2, 3). However, clear recommendations for PCP prophylaxis are lacking for many other immunosuppressive therapies, including most chemotherapy regimens. Recent clinical studies have reported that dose-dense or otherwise intensified chemotherapy regimens, particularly those used for breast cancer and other solid tumors, may increase the risk of PCP (4, 5). These observations indicate that prophylaxis may need to be reconsidered in specific high-risk settings, although clear evidence to guide such decisions remains limited.

Reliable identification of PCP within large-scale EHR or administrative database systems is essential for establishing rational prophylactic strategies and for accurately estimating disease incidence across diverse immunosuppressive treatment settings. However, diagnostic codes alone often fail to distinguish true PCP

from other respiratory diseases, as miscoding and non-specific coding are common limitations of routinely collected EHR and administrative data (6, 7). Recent Japanese claims-based validation research, including the VALIDATE-J study, reported that the positive predictive value (PPV) of PCP algorithms remained insufficient (approximately 20-50%) even when treatment and laboratory tests were incorporated. These findings highlighted the difficulty of accurately identifying PCP cases using claims data alone and underscore the need for alternative approaches using more granular EHR-based information (8). However, although the VALIDATE-J study incorporated medical records for gold-standard adjudication, the algorithms themselves were developed and evaluated using claims-linked hospital data with month-level temporal resolution. Moreover, the absence of well-validated case-identification algorithms for infectious outcomes in administrative claims data may lead to under-ascertainment of treatment-related infections. A recent claims-based study evaluating infectious outcomes associated with anticancer therapies highlighted that reliance on diagnostic codes alone can substantially underestimate infection incidence, underscoring a broader methodological limitation of claims-based outcome definitions (9).

To date, no studies in Japan have validated PCP case-identification algorithms using day-level, granular EHR data that allow precise alignment of diagnostic evaluation and therapeutic-dose prescription information. Therefore, the present study aimed to develop and validate a practical EHR-based algorithm for accurate identification of true PCP cases in a tertiary oncology care setting. By evaluating combinations of diagnostic codes, therapeutic-dose anti-

PCP prescriptions, laboratory testing, and chemotherapy exposure, we sought to establish a reproducible and scalable framework that can support future real-world oncology research and inform PCP prophylaxis strategies in patients receiving anticancer therapies.

## Patients and Methods

*Study design and setting.* This was a retrospective, single-center algorithm validation study conducted at Gifu University Hospital, a tertiary care institution in Japan. The study evaluated the validity of EHR-based algorithms designed to identify PCP cases. All analyses used structured EHR data routinely collected during clinical care. The study period spanned from April 1, 2022, to March 31, 2024.

*Study population.* Adult patients (aged  $\geq 20$  years) assigned an International Statistical Classification of Diseases and Related Health Problems, 10<sup>th</sup> Revision (ICD-10) code for *Pneumocystis jirovecii* pneumonia (B59) during the study period were extracted from the EHR dataset within the Diagnosis Procedure Combination (DPC)-linked hospital information system. Patients coded as “suspected” PCP were excluded. A total of 617 patients met the initial ICD-based criteria. True PCP cases were identified through a detailed review of electronic medical records by respiratory specialists, using clinical presentation, radiologic findings, microbiologic evidence, and response to therapy. The adjudication criteria were aligned with established diagnostic standards and required concordant radiologic evidence, compatible clinical symptoms, and supportive laboratory findings when available (*e.g.*,  $\beta$ -D-glucan elevation or positive *P. jirovecii* PCR). Discrepancies between evaluators were resolved by consensus. Eleven patients were confirmed as true PCP cases by respiratory specialists and constituted the gold-standard cohort for algorithm validation.

*Algorithm definitions.* Candidate algorithms were developed to identify true PCP cases using combinations of diagnostic codes (ICD-10 B59), prescriptions, and

laboratory test results. Algorithms were defined as follows: Algorithm 1: PCP code+therapeutic-dose anti-PCP medication within  $\pm 1$  month; Algorithm 2: Algorithm 1+ $\beta$ -D-glucan test performed; Algorithm 3: Algorithm 1+ $\beta$ -D-glucan  $\geq 6$  pg/ml; Algorithm 4: Algorithm 1+ $\beta$ -D-glucan  $\geq 8.5$  pg/ml; Algorithm 5: Algorithm 1+*P. jirovecii* PCR performed; Algorithm 6: Algorithm 1+concurrent chemotherapy; Algorithm 7: Algorithm 1+prescription duration  $\geq 21$  consecutive days.

These algorithms were designed to be reproducible using structured EHR or claims data and to reflect practical diagnostic and therapeutic workflows. The  $\pm 1$ -month window was adopted to account for the limited temporal resolution of diagnosis coding in claims-linked EHR systems, where disease codes are recorded at the month level, and to reflect real-world variability in the timing between code assignment and the initiation of therapeutic-dose treatment. This window, therefore, captures clinically relevant events even when such timing discrepancies occur. Because  $\beta$ -D-glucan testing and PCR are often performed during the diagnostic workup for suspected PCP, we evaluated algorithms that incorporate “test performed” as an indicator of diagnostic intent, separate from those that require a positive result.

*Definition of therapeutic doses.* Therapeutic doses of anti-PCP medications were defined as follows: TMP-SMX (oral): 5-12 tablets/day or 36-48 mini-tablets/day; TMP-SMX (intravenous): 15-20 mg/kg/day of trimethoprim, divided into three daily infusions; Atovaquone: 750 mg (5 ml) twice daily for 21 days; Pentamidine isethionate: 4 mg/kg/day intravenously.

These dose definitions are consistent with standard PCP treatment regimens described in national and international clinical guidelines (10).

*Data extraction and analysis.* Information on ICD codes, prescriptions, laboratory testing ( $\beta$ -D-glucan and *P. jirovecii* PCR performed), and chemotherapy administration was extracted from the DPC-based EHR system. For each algorithm, we calculated the number of identified cases, the

Table I. Performance of case-identification algorithms for true *Pneumocystis jirovecii* pneumonia (PCP) cases.

Algorithm	Definition	No. of identified cases	True PCP cases	PPV, % (95% CI)	Capture rate (% among 11 true cases)
A1	PCP code+therapeutic-dose anti-PCP prescription within $\pm 1$ month	12	11	91.7 (64.6-98.5)	100
A2	A1+ $\beta$ -D-glucan test performed	12	11	91.7 (64.6-98.5)	100
A3	A1+ $\beta$ -D-glucan $\geq 6$ pg/ml	9	9	100 (70.1-100)	81.8
A4	A1+ $\beta$ -D-glucan $\geq 8.5$ pg/ml	7	7	100 (64.6-100)	63.6
A5	A1+Pneumocystis PCR test performed	9	9	100 (70.1-100)	81.8
A6	A1+concurrent chemotherapy	7	7	100 (64.6-100)	63.6
A7	A1+prescription duration $\geq 21$ days	12	11	91.7 (64.6-98.5)	100

PPV: Positive predictive value; A1-A7: Algorithms 1-7; CI: confidence interval.

number of chart-confirmed true PCP cases, the PPV, and the capture rate among the 11 confirmed cases. The 95% confidence intervals (CIs) for PPVs were calculated using Wilson's score method for binomial proportions, which provides more reliable interval estimates than the normal approximation, particularly in small samples or when PPVs approach 0 or 1.

Because the primary objective of this study was to validate case-identification algorithms among patients with PCP-related diagnostic codes, individuals without PCP codes were not sampled. Therefore, sensitivity, specificity, and negative predictive value could not be determined.

**Ethical considerations.** This study was conducted in accordance with the Ethical Guidelines for Medical and Biological Research Involving Human Subjects in Japan and the Declaration of Helsinki (1964) and its later amendments. The study protocol was reviewed and approved by the Institutional Review Board of Gifu University Graduate School of Medicine (Approval No. 2024-107), with unified ethical oversight extending to the collaborating institution, Keio University. Because the study used only pre-existing data extracted from the EHR and involved no direct contact with patients, the requirement for written informed consent was waived. Instead, information regarding the study and procedures for opting out was posted on the hospital website to ensure that patients were given the opportunity to decline participation. All data were fully de-identified before analysis to protect patient confidentiality.

## Results

**Study population.** During the study period, 617 adult patients were assigned an ICD code for PCP in the EHR-based DPC dataset at Gifu University Hospital. Of these, 11 patients (1.8%) were confirmed as true PCP cases through chart review by respiratory specialists, while the remaining 606 patients (98.2%) did not meet the diagnostic criteria for PCP.

**Algorithm performance.** Table I summarizes the performance of each candidate algorithm in identifying true PCP cases. When only diagnostic codes were used, the PPV was 1.8%.

In contrast, Algorithm 1, which combines PCP codes with therapeutic-dose anti-PCP prescriptions, identified 12 cases, 11 of which were true PCP, yielding a PPV of 91.7% and a capture rate of 100%. Adding  $\beta$ -D-glucan testing (Algorithm 2) did not change performance because all patients identified by Algorithm 1 had undergone the test. When  $\beta$ -D-glucan positivity thresholds were applied (Algorithms 3 and 4), the PPV increased to 100%, but the number of true cases captured decreased to nine (81.8%) and seven (63.6%), respectively. Similarly, incorporating PCR testing (Algorithm 5) or concurrent chemotherapy (Algorithm 6) resulted in perfect PPVs but reduced capture rates (81.8% and 63.6%). Adding a prescription duration criterion of  $\geq 21$  days (Algorithm 7) maintained high accuracy and complete case capture, with performance identical to Algorithm 1 (PPV = 91.7%, capture rate = 100%).

## Discussion

This validation study demonstrated that adding prescription information to case-identification algorithms markedly improves the accuracy of identifying true PCP cases in EHR data. The positive predictive value of disease coding alone was only 1.8%, indicating that the PCP diagnostic code in routinely collected EHR or administrative data is highly nonspecific. In contrast, algorithms combining PCP codes with therapeutic-dose anti-PCP prescriptions demonstrated markedly higher accuracy. These prescription-enhanced algorithms achieved PPVs exceeding 90% and captured all chart-confirmed PCP cases. Notably, Algorithm 1 identified one false-positive case. This likely reflected empirical therapeutic-dose treatment initiated for clinically suspected PCP prior to diagnostic exclusion. These findings underscore a major limitation of database-based studies that rely solely on diagnostic codes to identify infection-related outcomes. Such misclassification can lead to biased effect estimates and undermine the validity of epidemiologic studies using routinely collected health data. For conditions such as PCP, which frequently coexist with malignancies or other immunosuppressive states, coding inaccuracies are particularly common and can lead to substantial misclassification bias. Previous validation research has shown that the PPV of PCP diagnosis codes in Japanese claims data ranges from approximately 20% to 50%, depending on the gold standard used (8). Compared to previous results, the significantly higher PPV observed in this study likely reflects the greater granularity of EHR-derived information. Specifically, distinguishing therapeutic from prophylactic doses and aligning prescription and diagnosis timing to the day level are key methodological advantages. This improvement allows for a more precise differentiation between empirical treatment, suspected cases, and confirmed cases, which is crucial for evaluating infection-related outcomes in oncology data. However, these tests are not uniformly administered and may yield false-negative results, potentially reducing the overall case capture rate. Although several algorithms achieved a PPV

of 100%, these estimates were based on very few cases and should therefore be interpreted with caution. These PPV estimates were derived from a small number of cases, and the corresponding Wilson 95% confidence intervals were wide, further underscoring the need for cautious interpretation. Algorithms incorporating test performance (Algorithm 2 and 5) capture the clinician's diagnostic intent, whereas algorithms requiring positive test results (Algorithm 4) trade off sensitivity for higher PPV. Therefore, a prescription-based algorithm offers a practical balance between diagnostic accuracy and applicability to large-scale real-world datasets. The clinical implications of this approach extend beyond algorithm validation, as it provides a scalable strategy for reliably identifying PCP in diverse epidemiologic settings. These findings contrast with those of the VALIDATE-J study, which reported that claims-based PCP algorithms achieved PPVs of only 20-50% even when incorporating treatment or diagnostic tests. This discrepancy likely reflects the greater granularity of EHR data, particularly the ability to distinguish therapeutic-dose treatment from prophylaxis and to capture contemporaneous testing patterns. Our study, therefore, provides an updated, more clinically compatible approach to PCP case identification within the Japanese EHR system.

In Japan, PCP prophylaxis is currently recommended for patients receiving rituximab-based therapy or purine analogs, but not for most other chemotherapy regimens (2). However, recent studies have suggested that dose-dense or otherwise intensified chemotherapy protocols, particularly those used for breast cancer and other solid tumors, may increase the risk of PCP (4, 5, 9, 12). Thus, reliable database-based identification of PCP is essential for accurately estimating disease incidence across different chemotherapy regimens and for informing future prophylaxis guidelines. In addition to its clinical relevance, this study aligns with methodological recommendations for validating case-identification algorithms in routinely collected health data. Our use of chart-confirmed diagnoses as a reference standard and the evaluation of multiple algorithmic definitions are consistent with best practices

described in prior intra-database validation work and in recent methodological guidance for validation studies using real-world data (12).

This study has several limitations that should be acknowledged. First, this analysis was conducted at a single center and included only patients who had been assigned a PCP code, which may limit the generalizability of our findings. Differences in diagnostic practices, laboratory test availability, and coding behaviors across institutions may influence algorithm performance. As demonstrated in the VALIDATE-J study, multi-institutional variability can substantially affect PPV estimates, underscoring the need for multicenter EHR-based validation. Second, the small number of confirmed PCP cases reflects the rarity of the condition and introduces statistical uncertainty into our estimates. Third, because patients without PCP codes were not included in the cohort, sensitivity and specificity could not be calculated. Finally, although clinical plausibility was maintained, the temporal relationships among diagnosis, testing, and treatment were not fully modeled, which may have introduced additional uncertainty. However, this limitation is substantially smaller than in prior claims-based validation studies, in which diagnostic and treatment information was available only at the claim-month level. The primary objective of this study is to improve the accuracy of identifying positive cases, which is crucial for assessing outcomes in real-world oncology research.

Despite its limitations, the study offers a clinically grounded and feasible approach to improving PCP case identification in EHR data. Aligning algorithm design with real-world treatment patterns bridges the gap between epidemiological methods and oncology clinical practice. This approach could serve as a foundation for future studies evaluating infection risk, treatment-related immunosuppression, and prevention strategies in patients receiving anticancer therapy.

## Conclusion

This study demonstrated that diagnostic codes alone are insufficient for accurately identifying true PCP cases in

EHR data. Incorporating therapeutic-dose prescriptions of anti-PCP agents, particularly when combined with prescription duration criteria, substantially improved diagnostic accuracy while remaining feasible for use in real-world databases. The proposed algorithm provides a practical and scalable approach for future epidemiologic studies assessing PCP incidence and prophylaxis needs across diverse chemotherapy and immunosuppressive treatment settings. By addressing limitations identified in prior claims-based validation studies, this EHR-based algorithm offers a contemporary and clinically aligned strategy for accurate PCP identification in Japan. Multicenter validation will be essential to confirm the generalizability of this algorithm and support its application in nationwide research.

## Conflicts of Interest

The Authors declare no conflicts of interest in relation to this study.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Authors' Contributions

*Study conception and design:* YK, HI, and MT conceived the study concept and designed the research plan; *Data acquisition:* SS extracted and prepared the electronic health record dataset. HI also supported data handling and preliminary dataset checks; *Data evaluation:* YK, HI, and MT evaluated and organized the electronic health record data for research use. YK, JE, and YT contributed to the clinical adjudication of PCP cases; *Statistical analysis:* MT performed the statistical analyses and contributed to the interpretation of algorithm performance; *Manuscript drafting:* YK, HI, and MT prepared the first draft of the manuscript; *Critical revision of the manuscript:* KK and YT contributed to strengthening the scientific content and

critically revised the manuscript for important intellectual improvements. All authors revised the manuscript critically and provided important intellectual contributions; *Approval of the final manuscript*: All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

## Artificial Intelligence (AI) Disclosure

During the preparation of this manuscript, a large language model (ChatGPT, OpenAI) was used solely for language editing and stylistic improvements in select paragraphs. No sections involving the generation, analysis, or interpretation of research data were produced by generative AI. All scientific content was created and verified by the authors. Additionally, no figures or visual data were generated or modified using generative AI or machine learning-based image enhancement tools.

## References

- 1 Takeuchi K, Yakushijin Y: Pneumocystis jirovecii pneumonia prophylaxis for cancer patients during chemotherapy. *Pathogens* 10(2): 237, 2021. DOI: 10.3390/pathogens10020237
- 2 Taplitz RA, Kennedy EB, Bow EJ, Crews J, Gleason C, Hawley DK, Langston AA, Nastoupil LJ, Rajotte M, Rolston KV, Strasfeld L, Flowers CR: Antimicrobial prophylaxis for adult patients with cancer-related immunosuppression: ASCO and IDSA clinical practice guideline update. *J Clin Oncol* 36(30): 3043-3054, 2018. DOI: 10.1200/JCO.18.00374
- 3 Maertens J, Cesaro S, Maschmeyer G, Einsele H, Donnelly JP, Alanio A, Hauser PM, Lagrou K, Melchers WJ, Helweg-Larsen J, Matos O, Bretagne S, Cordonnier C, 5th European Conference on Infections in Leukaemia (ECIL-5): ECIL guidelines for preventing Pneumocystis jirovecii pneumonia in patients with haematological malignancies and stem cell transplant recipients. *J Antimicrob Chemother* 71(9): 2397-2404, 2016. DOI: 10.1093/jac/dkw157
- 4 Waks AG, Tolaney SM, Galar A, Arnaout A, Porter JB, Marty FM, Winer EP, Hammond SP, Baden LR: Pneumocystis jirovecii pneumonia (PCP) in patients receiving neoadjuvant and adjuvant anthracycline-based chemotherapy for breast cancer: incidence and risk factors. *Breast Cancer Res Treat* 154(2): 359-367, 2015. DOI: 10.1007/s10549-015-3573-2
- 5 Barbounis V, Aperis G, Gambletsas E, Koumakis G, Demiris M, Vassilomanolakis M, Efremidis A: Pneumocystis carinii pneumonia in patients with solid tumors and lymphomas: Predisposing factors and outcome. *Anticancer Res* 25(1B): 651-655, 2005.
- 6 Senécal J, Smyth E, Del Corpo O, Hsu JM, Amar-Zifkin A, Bergeron A, Cheng MP, Butler-Laporte G, McDonald EG, Lee TC: Non-invasive diagnosis of Pneumocystis jirovecii pneumonia: a systematic review and meta-analysis. *Clin Microbiol Infect* 28(1): 23-30, 2022. DOI: 10.1016/j.cmi.2021.08.017
- 7 Yamana H, Konishi T, Yasunaga H: Validation studies of Japanese administrative health care data: A scoping review. *Pharmacoepidemiol Drug Saf* 32(7): 705-717, 2023. DOI: 10.1002/pds.5636
- 8 Hase R, Suzuki D, de Luise C, Chen H, Nonnenmacher E, Higuchi T, Katayama K, Kinjo M, Jinno S, Morishima T, Sugiyama N, Tanaka Y, Setoguchi S: Validity of claims-based diagnoses for infectious diseases common among immunocompromised patients in Japan. *BMC Infect Dis* 23(1): 653, 2023. DOI: 10.1186/s12879-023-08466-8
- 9 Dote S, Inose R, Goto R, Kobayashi Y, Muraki Y: Risk of a second cancer and infection in patients with indolent B-cell lymphoma exposed to first-line bendamustine plus rituximab: A retrospective analysis of an administrative claims database. *Hematol Oncol* 41(3): 354-362, 2023. DOI: 10.1002/hon.3128
- 10 Haensel L, Sprute R, Grothe J, Robert-Gangneux F, Gangneux JP, Meis JF, Cornely OA, Koehler P: The 2025 EQUAL Pneumocystis Score-an ECMM tool to measure QUALity in Pneumocystis pneumonia management. *JAC Antimicrob Resist* 7(5): dlaf165, 2025. DOI: 10.1093/jacamr/dlaf165
- 11 Ehrenstein V, Hellfritzsch M, Kahlert J, Langan SM, Urushihara H, Marinac-Dabic D, Lund JL, Sørensen HT, Benchimol EI: Validation of algorithms in studies based on routinely collected health data: general principles. *Am J Epidemiol* 193(11): 1612-1624, 2024. DOI: 10.1093/aje/kwae071
- 12 Shiiba R, Himeji D, Matsumoto R, Tanaka GI, Otomo N: Pneumocystis jirovecii pneumonia in three patients with breast cancer receiving neoadjuvant dose-dense chemotherapy. *Cureus* 14(2): e21812, 2022. DOI: 10.7759/cureus.21812
- 13 Thurin NH, Bosco-Levy P, Blin P, Rouyer M, Jové J, Lamarque S, Lignot S, Lassalle R, Abouelfath A, Bignon E, Diez P, Gross-Goupil M, Soulié M, Roumiguié M, Le Moulec S, Debouverie M, Brochet B, Guillemin F, Louapre C, Maillart E, Heinzlef O, Moore N, Droz-Perroteau C: Intra-database validation of case-identifying algorithms using reconstituted electronic health records from healthcare claims data. *BMC Med Res Methodol* 21(1): 95, 2021. DOI: 10.1186/s12874-021-01285-y