Article:
Haeusler, Gabrielle M., Thursky, Karin A., Slavin, Monica A. et al. (5 more authors) (2018) External Validation of Six Pediatric Fever and Neutropenia Clinical Decision Rules. Pediatric Infectious Disease Journal. pp. 329-335. ISSN 1532-0987

https://doi.org/10.1097/INF.0000000000001777

The Pediatric Infectious Disease Journal Publish Ahead of Print

DOI: 10.1097/INF.0000000000001777

External Validation of Six Pediatric Fever and Neutropenia Clinical Decision Rules

Gabrielle M. Haeusler, MBBS,[1-5] Karin A. Thursky, MD,[2,5-8] Monica A. Slavin, MD,[2,5,6,8] Francoise Mechinaud, MD,[9] Franz E. Babl, MD,[10-12] Penelope Bryant, PhD,[11,13] Richard De Abreu Lourenco, PhD[14], and Robert Phillips, PhD[15,16]

Corresponding author: Dr Gabrielle M. Haeusler, Department of Infectious Diseases, Peter MacCallum Cancer Centre, 305 Grattan Street, Melbourne, Australia, 3000, P: +61 3 9656 5853 F: +61 3 9656 1185, E: gabrielle.haeusler@petermac.org

Abbreviated Title: Validation of Six Fever and Neutropenia Clinical Decision Rules

Running Head: Fever and Neutropenia Validation Study

1. The Paediatric Integrated Cancer Service, Parkville, Victoria, Australia, 3052

2. Department of Infectious Diseases, Peter MacCallum Cancer Centre, Victorian Comprehensive Cancer Centre, Melbourne, Victoria, Australia, 3000

3. Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria, Australia, 3010

4. Department of Infection and Immunity, Monash Children's Hospital, Department of Paediatrics, Monash University, Clayton, Victoria, Australia, 3168

5. NHMRC National Centre for Infections in Cancer, Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria, Australia, 3010

6. Department of Medicine, University of Melbourne, Parkville, Victoria, Australia, 3010

7. NHMRC National Centre for Antimicrobial Stewardship, The Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia, 3000

8. Victorian Infectious Diseases Service, The Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia, 3000

9. Children's Cancer Centre, Royal Children's Hospital, Parkville, Victoria, Australia, 3052

10. Department of Emergency Medicine, Royal Children's Hospital, Parkville, Victoria, Australia, 3052

11. Murdoch Children's Research Institute, Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia, 3052

12. Department of Paediatrics, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, Victoria, Australia, 3010

13. Infectious Diseases Unit, Royal Children's Hospital, Parkville, Victoria, Australia, 3052

14. Centre for Health Economics Research and Evaluation, University of Technology Sydney, Ultimo, New South Wales, Australia, 2007

15. Centre for Reviews and Dissemination, University of York, Heslington, York, UK, YO10 5DD

16. Leeds Children's Hospital, Leeds General Infirmary, Leeds LS1 3EX, UK

Conflicts of interest and sources of funding: GMH was supported by a National Health and Medical Research Council post-graduate scholarship (GNT1056158). FEB was part funded by a grant from the Royal Children's Hospital Foundation, Melbourne, Australia and an NHMRC Practitioner Fellowship. RP was funded by a Post-Doctoral Research Fellow grant from the NIHR (UK) Grant number PDF-10872. There are no conflicts of interest to declare.

Key words: fever and neutropenia, validation, risk stratification, child, low-risk

Background

Fever and neutropenia (FN) clinical decision rules (CDRs) are recommended to help distinguish children with cancer at high and low risk of severe infection. The aim of this study was to validate existing pediatric FN CDRs, designed to stratify children with cancer at high or low risk of serious infection or medical complication.

Methods

Pediatric CDRs suitable for validation were identified from a literature search. Relevant data were extracted from an existing dataset of 650 retrospective FN episodes in children with cancer. The sensitivity and specificity of each of the CDR were compared with the derivation studies to assess reproducibility.

Results

Six CDRs were identified for validation: two were designed to predict bacteremia and four to predict adverse events. Five CDRs exhibited reproducibility in our cohort. A rule predicting bacteremia had the highest sensitivity (100%; 95% confidence interval (CI) 93-100%) although poor specificity (17%) with only 15% identified as low risk. For adverse events, the highest sensitivity achieved was 84% (95% CI, 75-90%) with specificity of 29% and 27% identified as low risk. A rule intended for application after a 24-hour period of inpatient observation yielded a sensitivity of 80% (95% CI, 73-86) and specificity of 46%, with 44% identified as low-risk.

Conclusion

Five CDRs were reproducible although not all can be recommended for implementation because of either inadequate sensitivity or failure to identify a clinically meaningful number of low-risk patients. The 24-hour rule arguably exhibits the best balance between sensitivity and specificity in our population.

INTRODUCTION

The risk of infection in the setting of chemotherapy-induced neutropenia, heralded by fever, remains an unavoidable complication of the treatment of childhood cancer. Treatment strategies for fever and neutropenia (FN) that are tailored to an individual's likelihood of severe infection, by incorporating risk stratification, are well described.(1) To help differentiate children at low and high risk of severe infection, pediatric FN clinical decision rules (CDRs) have been recommended as an important adjunct to the risk stratification process.(2) Children accurately identified as low-risk may benefit from reduced intensity antibiotic therapy and early hospital discharge, while additional supportive care measures and heightened vigilance may avoid clinical deterioration in high-risk patients.(3)

There are four key components to CDR development: derivation, internal validation, external validation and implementation and impact analysis.(4) Before a CDR, especially one targeting pediatric FN, can be implemented into practice it should undergo evaluation in a population external to the derivation dataset to ensure it is safe and reliable.(5) While many of the pediatric FN CDRs that have undergone external validation show some reproducibility, most result in lower sensitivity compared to the derivation study.(6-10) This highlights the importance of detailed local external validation to provide clinicians with a realistic expectation of the predictive performance of a CDR in their own population. Such validations will identify CDR limitations and should guide implementation of low-risk treatment programs that incorporate safeguards against potential failures of the CDR.

Using an existing local dataset of consecutive episodes of outpatient onset FN, retrospectively collected to validate the Prediciting Infectious ComplicatioNs in Children with Cancer (PICNICC) CDR, the aim of this study was to validate additional published pediatric FN CDRs

designed to stratify children with cancer or hematologic malignancy at high or low risk of serious infection or medical complication.(10) The sensitivity, specificity, positive predictive value and negative predictive value of each of these rules applied to our retrospective dataset was compared with the derivation studies.

METHODS

Identification of clinical decision rules for validation

A list of published pediatric CDRs was compiled from two systematic reviews.(6, 11) A PubMed search for relevant pediatric CDRs published since these reviews using the search terms: (fever OR febrile OR sepsis) AND (neutropenia or neutropenic) AND (child OR children OR paediatric OR pediatric) was also conducted (non-English studies and abstracts were excluded). The date of the search was 18th April 2016. Studies were excluded if there was insufficient information available from the existing retrospective dataset to validate the rule or if no rule was described. Rules that included presence of CVAD as predictor of outcome were also excluded as 95% of children in the existing dataset had a CVAD and this was deemed *a priori* as non-discriminatory.(10)

Data collection

External validation was performed using an existing local dataset of retrospectively identified episodes of outpatient-onset FN in children and adolescents with cancer or hematological malignancy.(10) This local dataset will be herein described as validation cohort. Detailed methodology for patient episode identification and data collection is described elsewhere.(10) Briefly, consecutive episodes of outpatient-onset FN in children and adolescents (age <19 years) with cancer and receiving chemotherapy or hematopoetic stem cell transplant (HSCT) at The Royal Children's Hospital (RCH), Melbourne were included in the study (November 2011 to

June 2015). Demographic, FN episode and clinical outcome data were obtained from electronic records and entered into REDCap database.(12) Data were collected by a research assistant blinded to the CDRs included in this analysis. Patients were excluded if they were already receiving empiric or targeted treatment antibiotics or onset of the FN episode occurred in hospital.

Definitions

Fever was defined as a single tympanic temperature greater than, or equal to, 38 degrees Celsius and neutropenia was defined as an absolute neutrophil count less than 1000/mm^3. Bacteremia was defined as a recognized pathogen (including viridans group streptococci in the setting of mucosal barrier injury or neutropenia) cultured from one or more blood cultures or common commensal bacteria cultured from two or more blood cultures drawn on separate occasions.(13) For validation, the variable or outcome definition used in the derivation study was applied to our validation cohort. An exception to this was 'bacteremia,' where the above international consensus definition was applied to avoid incorrectly attributing single positive blood culture with a common commensal as a true bacteremia. Where no definition was provided, variables or outcomes followed international consensus recommendations.(13, 14) The date and time bacteremia episodes were known were extracted from the electronic pathology database. For all other clinical and microbiologically defined infections (MDIs) and for medical complications such as intensive care unit (ICU) admission, the date and time the infection or event was documented in the medical record were used.

Statistical analysis

The sensitivity, specificity, positive predictive value (PPV) and negative predictive values (NPV) for each rule were calculated in our validation cohort using both the inclusion and exclusion

criteria from our existing dataset(10) and those criteria restricted to that described by the derivation studies. Confidence intervals around sensitivity and specificity were calculated using hybrid Wilson/Brown method. To ensure consistency, confidence intervals from the derivation studies were recalculated from available data. For rules that stratified patients into more than 2 groups (ie low, intermediate and high-risk), we combined intermediate and low risk into a single low-risk group. The sensitivity and specificity of the Swiss Pediatric Oncology Group (SPOG) rule was determined by combining the information on episodes with MDI known at day 2 with the results of prediction on the remaining episodes.(9)

Continuous data were presented as median and interquartile range. Fisher's exact test was used to estimate P-values for categorical data, including comparison of sensitivity and specificity between derivation and validation cohorts. The Newcombe-Wilson test with continuity correction was used for difference between proportions. A CDR was considered reproducible if there was no significant difference in either the sensitivity or specificity between the derivation and validation cohorts. All tests were 2-tailed, and a P value of <0.05 was considered to be statistically significant.

RESULTS

A total of 21 potentially relevant studies describing pediatric FN CDRs or risk factors for severe infection were identified in published systematic reviews(6, 11) and a further six were identified in our search of the literature.(15-20) Of these 27 studies, six described CDRs that were suitable for validation in our dataset.(9, 21-25) Eleven could not be validated, as there was insufficient information available from the existing dataset.(16, 21, 26-34) A further eight studies described individual variables for infection or adverse outcome in the absence of a defined rule.(17-20, 35-

38) In the remaining two, a CVAD was used as a predictor of outcome(39) and validation of the PICNICC CDR using this dataset had already been completed.(10)

Details of study design and demographic data from the validation cohort and the included derivation studies are available in Table, Supplemental Digital Content 1, http://links.lww.com/INF/C843. Where sufficient data were available for comparison, there was no significant difference in sex, proportion with relapsed disease or death in the validation cohort compared to any of the derivation studies. There were significantly more patients with hematologic malignancy in the validation cohort compared to the Alexander *et al* and Rackoff *et al* derivation studies.(22, 25) Bacteremia occurred in significantly fewer FN episodes in our validation cohort compared to the Swiss Pediatric Oncology Group (SPOG), Baorto *et al* and Rackoff *et al* derivation studies.(9, 23, 25)

Table 1 provides details of the inclusion and exclusion criteria as well as description of the CDR variables and predicted outcomes. A different definition of fever, albeit slightly, was used in all six studies and almost all excluded patients with HSCT. The number of clinical variables included in the CDRs ranged from one to nine (1 variable in 2 CDR, 2 in 1, 4 in 2 and 9 in 1). Two CDRs were designed to predict bacteremia only, of which a definition was provided for only one.(23, 25) The remaining four CDRs predicted composite outcomes encompassing a varying combination of microbiological infection, sepsis, pneumonia, severe medical complication or death.

Results of the sensitivity, specificity, PPV and NPV analyses for each CDR are shown in Table 2. The clinical impact of each rule was calculated using both the existing dataset inclusion criteria (validation cohort) and the derivation study inclusion criteria (restricted validation cohort). Notably, for each of the six rules, there was very little difference in the validation results

obtained using the different inclusion criteria with broad overlap of the 95% confidence intervals for all results (Figures 1 and 2). Bacteremia was observed in significantly fewer FN episodes in our validation cohort compared to the two derivation studies that specifically investigated this outcome.(23, 25) (Table 2) For the remainder, there was no significant difference in the proportion of patients with the specific outcome investigated.

A direct comparison of sensitivity and specificity between the derivation studies and both the validation cohort and restricted validation cohort is also shown in Table 2. There was a significant difference in both sensitivity and specificity between the Hakim *et al* derivation and both validation cohorts.(21) For the remaining five CDRs, there was no difference between sensitivity for one CDR (Klaassen), specificity in one (SPOG) or both sensitivity and specificity in three (Rackoff, Baorto and Alexander).(22, 23, 25) The CDR with the highest sensitivity in the local validation cohort was developed by Baorto *et al*, followed by the Klaassen and SPOG CDRs.(9, 23, 24) Of these three, the SPOG CDR had the greatest specificity at 46%, correctly identifying 86% of low-risk patients in the local validation cohort.(9)

DISCUSSION

Using a pre-existing dataset we were able to externally validate six pediatric CDRs designed to predict bacteremia or adverse outcomes in children with cancer and FN. Reproducibility was observed in five of the CDRs, with no significant difference in both sensitivity and specificity between the derivation and validation cohorts in three of these.(22, 23, 25) Although often attributed as a cause of discordant derivation and validation study results, we also showed that using inclusion and exclusion criteria that varied to that of the derivation study appeared to have little impact.

Three of the reproducible rules were designed to predict a composite outcome of 'adverse outcome' or 'serious infection.'(9, 22, 24) While the definition for these composite outcomes varied between studies, all three included at least bacteremia, other bacterial infection and death. The CDR by Klaassen *et al* had the highest sensitivity in validation cohort (84%) but the lowest specificity (28%).(24) The inclusion of the subjective outcome of 'life-threatening complication as judged by the treating physician,' may have contributed to the lower sensitivity observed in the validation results of the SPOG and Alexander CDR.(9, 22) Despite this, the SPOG rule, in particular, produced a sensitivity of up to 80% with a specificity of 46%, correctly identifying 86% of low-risk patients. This CDR is unique in that it is applied after a 24-hour period of inpatient observation.

The remaining two reproducible CDRs in our validation cohort were designed to predict bacteremia.(23, 25) Notably, the proportion of bacteremia episodes in the local validation cohort (9%) was significantly lower than both these derivation studies. This difference can be attributed to the strict exclusion of common commensals identified on single blood cultures to avoid incorrectly attributing these as a true bacteremia.(35). The rule developed by Baorto *et al* produced the highest sensitivity and PPV in the validation cohort, approaching 100%. Not surprisingly the specificity was poor, with very few episodes being identified as low risk. These data suggest that implementation of this CDR in our population, while reassuring given the very high sensitivity, would be difficult to justify as only 33 out of 177 patients per year would qualify as low risk and therefore appropriate for consideration of reduced intensity therapy.

While this is the first study to validate these international CDRs in Australia, five have previously undergone validation in populations external to the derivation studies.(9, 22-25) Most of these validation studies demonstrate at least some degree of overlap in confidence intervals for

either sensitivity or specificity suggesting validity.(7-9, 40-42) The three rules derived in the USA(22, 23, 25) have been shown to be effective in Europe and the UK(7, 9, 40, 42) and the Canadian rule(24) in Europe and USA.(7, 9, 41) However, until now, the SPOG rule has not been tested outside of Europe.(8)

The CDRs included in this study were validated using an existing dataset designed to validate the PICNICC CDR.(15) This rule was developed from an individual participant data meta-analysis and included data from four of the rules validated in this study.(9, 21, 22, 24) For the prediction of MDI, the recalibrated PICNICC rule did not perform as well in our population as compared to the derivation study with a sensitivity and specificity of 78.4% and 39.8%, respectively.(10) However when using methodology described by Ammann *et al* for the SPOG rule, the sensitivity of the PICNICC rule improved to 88% further highlighting the importance of an overnight period of observation.(9)

Although this study was performed using an existing, retrospective dataset, it includes a contemporary cohort of consecutive episodes of FN. Given the reliance on the existing dataset, sample size calculations were not performed. For validation of a CDR in a new population, a sample size that includes 100 outcome events and 100 non-outcome events has been recommended.(43) Based on this, an appropriate sample size was achieved for validation of three of the six CDRs, of which two predicted adverse outcome(9, 22) and one predicted significant bacterial infection.(24) With regard to validation of the SPOG rule, it is possible that the date and time that non-bacteraemia microbiologically and clinically documented infections, as well as medical complications such as admission to ICU, were known were earlier than what was documented in the medical records. This would have underestimated the sensitivity of the SPOG rule that takes into account the number of infections or medical complications known at time of

assessment. Our study is unique in that it compared validation results using differing inclusion criteria. The similarities in results between the two validation cohorts suggests that the impact of differences in inclusion and exclusion criteria between derivation and validation studies may have been overstated, although this will vary study to study.

The rationale for risk prediction in FN is to tailor treatment strategies according to the likelihood of having a documented infection or adverse outcome. This will avoid over treatment of children with viral illness or non-infective causes of fever and, conversely, enable targeted treatment and observation strategies for high-risk patients to avoid severe complication such as late onset sepsis, ICU admission and death. A systematic review of oral and outpatient antibiotic regimens for children with low-risk FN which analyzed data from 13 randomized controlled trials and 24 prospective observational studies concluded that both oral antibiotics and outpatient therapy are safe alternatives to standard care.(1) The rate of modification from reduced intensity therapy back to standard inpatient care appears to be affected by the time of discharge with a significant reduction in requirements for pathway modifications when patients were discharged after 48 hours compared to immediately (2.2% versus 14%). Deviations from low-risk treatment were also significantly less frequent in centers using stringent risk tools compared to centers using unnamed and unvalidated tools (7% versus 19.1%).(1) Although none of the validated rules included in this study have been subject to formal implementation and impact analysis, it is conceivable that a rule such as SPOG rule, which requires a 24 hours period of observation, may result in less failure.(9)

When implemented at the time of hospital presentation with FN, the sensitivity of a CDR to predict infection or adverse event is considered to be of greatest importance. While this is traditionally at the expense of specificity, the ability of the test to correctly identify those without

the disease needs to be sufficiently high to make implementation of the rule worthwhile. Our validation study has identified five internationally derived pediatric FN CDRs that are reproducible. However, although reproducibility was observed in these studies, not all can be recommended for implementation based on either inadequate sensitivity or failure to identify a sufficient number of patients that are low risk. Of the rules validated in this study, the SPOG rule arguably exhibits the best balance between sensitivity and specificity in our population and may facilitate the implementation of a low-risk FN program that is safe, practical and will avoid the over treatment of as many children as possible. Further research is required to assess the clinical, psychosocial and economic impact of such a program and to ensure the strengths and the weakness of the CDR continue to be evaluated.

REFERENCES

1. Morgan J, Cleminson J, Atkin K, Stewart L, Phillips R. Systematic review of reduced therapy regimens for children with low risk febrile neutropenia. *Support Care Cancer*. 2016;24:2651-2660.

2. Lehrnbecher T, Robinson P, Fisher B, et al. Guideline for the Management of Fever and Neutropenia in Children With Cancer and Hematopoietic Stem-Cell Transplantation Recipients: 2017 Update. *J Clin Oncol*. 2017;35:2082-2094.

3. Haeusler G, Sung L, Ammann R, Phillips B. Management of fever and neutropenia in paediatric cancer patients: room for improvement? *Curr Opin Infect Dis*. 2015;28:532-538.

4. McGinn TG GG, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000;284:e84.

5. Phillips B. Clinical decision rules: how to build them. *Arch Dis Child Educ Pract Ed*. 2010;95:83-87.

6. Phillips RS, Lehrnbecher T, Alexander S, Sung L. Updated systematic review and meta-analysis of the performance of risk prediction rules in children and young people with febrile neutropenia. *PLoS ONE*. 2012;7:e38300.

7. Macher E, Dubos F, Garnier N, et al. Predicting the risk of severe bacterial infection in children with chemotherapy-induced febrile neutropenia. *Pediatr Blood Cancer*. 2010;55:662-667.

8. Miedema KGE, de Bont ESJM, Oude Nijhuis CSM, van Vliet D, Kamps WA, Tissing WJE. Validation of a new risk assessment model for predicting adverse events in children with fever and chemotherapy-induced neutropenia. *J Clin Oncol*. 2011;29:e182-184; author reply e185.

9. Ammann RA, Bodmer N, Hirt A, et al. Predicting adverse events in children with fever and chemotherapy-induced neutropenia: the prospective multicenter SPOG 2003 FN study. *J Clin Oncol*. 2010;28:2008-2014.

10. Haeusler GM, Thursky KA, Mechinaud F, et al. Predicting Infectious ComplicatioNs In Children with Cancer: an external validation study. *Brit J Cancer*. 2017.

11. Phillips B, Wade R, Stewart LA, Sutton AJ. Systematic review and meta-analysis of the discriminatory performance of risk prediction rules in febrile neutropaenic episodes in children and young people. *Eur J Cancer*. 2010;46:2950-2964.

12. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377-381.

13. Haeusler G, Phillips R, Lehrnbecher T, Thursky K, Sung L, Ammann R. Core Outcomes and Definitions for Pediatric Fever and Neutropenia Research: A Consensus Statement From an International Panel. *Pediatr Blood Cancer*. 2015;62:483-489.

14. Goldstein B, Giroir B, Randolph A, International Consensus Conference on Pediatric S. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med*. 2005;6:2-8.

15. Phillips R, Sung L, Ammann R, et al. Predicting microbiologically defined infection in febrile neutropenic episodes in children: global individual participant data multivariable meta-analysis. *Bri J Cancer*. 2016;114:623-630.

16. Miedema K, Tissing W, Abbink F, et al. Risk-adapted approach for fever and neutropenia in paediatric cancer patients – A national multicentre study. *Eur J Cancer*. 2016;53:16-24.

17. Delebarre M, Garnier N, Macher E, et al. Which Variables Are Useful for Predicting Severe Infection in Children With Febrile Neutropenia?
. *J Pediatr Hematol Oncol*. 2015;37:e468-474.
18. Prasad M, Chinnaswamy G, Arora B, Vora T, Hawaldar R, Banavali S. Risk predictors for adverse outcome in pediatric febrile neutropenia: Single center experience from a low and middle-income country. *Indian J Cancer*. 2014;51:432-437.
19. Hazan G, Ben-Shimol S, Fruchtman Y, et al. Clinical and laboratory parameter dynamics as markers of blood stream infections in pediatric oncology patients with fever and neutropenia
. *J Pediatr Hematol Oncol*. 2014;36:e275-279.
20. Bothra M, Seth R, Kapil A, Dwivedi S, Bhatnagar S, Xess I. Evaluation of predictors of adverse outcome in febrile neutropenic episodes in pediatric oncology patients.
. *Indian J Pediatr*. 2013;80:297-302.
21. Hakim H, Flynn PM, Srivastava DK, et al. Risk prediction in pediatric cancer patients with fever and neutropenia. *Pediatr Infect Dis J*. 2010;29:53-59.
22. Alexander SW, Wade KC, Hibberd PL, Parsons SK. Evaluation of risk prediction criteria for episodes of febrile neutropenia in children with cancer. *J Pediatr Hematol Oncol*. 2002;24:38-42.
23. Baorto EP, Aquino VM, Mullen C, Buchanan GR, DeBaun MR. Clinical Parameters Associated with Low Bacteremia Risk in 1100 Pediatric Oncology Patients with Fever and Neutropenia. *Cancer*. 2001;92:909-913.
24. Klaassen R, Goodman T, Pham B, Doyle J. "Low-risk" prediction rule for pediatric oncology patients presenting with fever and neutropenia. *J Clin Oncol*. 2000;18:1012-1019.

25. Rackoff W, Gonin R, Robinson C, Kreissman S, Breitfeld P. Predicting the risk of bacteremia in childen with fever and neutropenia. *J Clin Oncol*. 1996;14:919-924.

26. Adcock K, Akins R, Farrington E. Evaluation of empiric vancomycin therapy in children with fever and neutropenia. *Pharmacotherapy*. 1999;19:1315-1320.

27. Ammann RA, Hirt A, Luthy AR, Aebi C. Identification of children presenting with fever in chemotherapy-induced neutropenia at low risk for severe bacterial infection. *Med Pediatr Oncol*. 2003;41:436-443.

28. Ammann RA, Hirt A, Luthy AR, Aebi C. Predicting bacteremia in children with fever and chemotherapy-induced neutropenia. *Pediatr Infect Dis J*. 2004;23:61-67.

29. Jones G, Konsler G, SN P. Infection risk factors in febrile, neutropenic children and adolescents. *Pediatr Hematol Oncol*. 1996;13:217-229.

30. Lucas K, Brown A, Armstrong D, Chapman D, Heller G. The identification of febrile, neutropenic children with neoplastic disease at low risk for bacteremia and complications of sepsis. *Cancer*. 1996;77:791-798.

31. West DC, Marcin JP, Mawis R, Jongsong He MS, Nagle A, Dimand R. Children with cancer, fever and treatment induced neutropenia. *Pediatr Emerg Care*. 2004;20:79-84.

32. Paganini H, Aguirre C, Puppa G, et al. A Prospective, Multicentric Scoring System to Predict Mortality in Febrile Neutropenic Children With Cancer. *Cancer*. 2007;109.

33. Santolaya ME, Alvarez AM, Becker A, et al. Prospective, multicenter evaluation of risk factors associated with invasive bacterial infection in children with cancer, neutropenia, and fever. *J Clin Oncol*. 2001;19:3415-3421.

34. Agyeman P, Aebi C, Hirt A, et al. Predicting bacteremia in children with cancer and fever in chemotherapy-induced neutropenia: results of the prospective multicenter SPOG 2003 FN study. *Pediatr Infect Dis J*. 2011;30:e114-119.

35. Hann I, Viscoli C, Paesmans M, Gaya H, Glauser M. A comparison of outcome from febrile neutropenic episodes in children compared with adults: results from four EORTC studies. *Brit J Haem*. 1997;99:580-588.

36. Riikonen P, Jalanko H, Hovi L, Saarinen UM. Fever and neutropenia in children with cancer: diagnostic parameters at presentation. *Acta Paediatr*. 1993;82:271-275.

37. Tezcan G, Kupesiz A, Ozturk F, et al. Episodes of fever and neutropenia in children with cancer in a tertiary care medical center in Turkey. *Pediatr Hematol Oncol*. 2006;23:217-229.

38. Badiei Z KM, Alami MH, Kianifar HR, Banihashem A, Farhangi H, Razavi AR. Risk factors associated with life-threatening infections in children with febrile neutropenia: a data mining approach. *J Pediatr Hematol Oncol*. 2011;33:e9-e12.

39. Rondinelli PIP, Ribeiro KdCB, de Camargo B. A proposed score for predicting severe infection complications in children with chemotherapy-induced febrile neutropenia. *J Pediatr Hematol Oncol*. 2006;28:665-670.

40. Dommett R, Geary J, Freeman S, et al. Successful introduction and audit of a step-down oral antibiotic strategy for low risk paediatric febrile neutropaenia in a UK, multicentre, shared care setting. *Eur J Cancer*. 2009;45:2843-2849.

41. Madsen K, Rosenman M, Hui S, Breitfeld P. Value of electronic data for model validation and refinement: bacteremia risk in children with fever and neutropenia. . *J Pediatr Hematol Oncol*. 2002;24:256-262.

42. Arif T, Sutcliffe R, Hewitt M, et al. Validation of two risk stratification guidelines in a one year cohort of febrile admissions in paediatric oncology patients in a UK centre. *Arch Dis Child*. 2014;99:A1-A212.

43. Vergouwe Y, Steyerberg E, Eijkemans M, Habbema J. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005;58:475-483.

Figure 1. Forest plot showing sensitivity with 95% confidence intervals for derivation studies, validation cohort and restricted validation cohort

Figure 2. Forest plot showing specificity with 95% confidence intervals for derivation studies, validation cohort and restricted validation cohort

Table 1. Comparison of clinical decision rule inclusion and exclusion criteria, variables and outcomes

Rule	**Inclusion criteria**	**Exclusion criteria**	**High risk criteria**	**High risk outcome**
Validation cohort(10)	Cancer or haematological malignancy; fever ≥ 38.0°C once; ANC ≤ 1000cells/mm^3; outpatient	Receiving antibiotics; inpatient onset FN	NA	NA
SPOG(9)	Cancer or haematological malignancy; fever ≥ 38.5°C once or ≥ 38.0°C during ≥2 hours^; ANC ≤ 500cells/mm^3; outpatient	Myeloablative chemotherapy; AE known at presentation	Applied after 24 hours. Total score ≥ 9 = high risk of AE. Score for preceding chemotherapy more intensive than ALL maintenance =4; Hb≥90=5;	Adverse outcome – defined as a SMC (death, complication requiring ICU and potentially life-threatening complication as judged by the treating physician) as a result of infection, MDI (positive bacterial or fungal culture from a

			leukocyte count < 0.3 G/L=3; platelet <50 G/L=3	normally sterile site and detection of a viral antigen by PCR) and radiologically confirmed pneumonia. Bacteraemia not defined
Hakim(21)	Cancer or haematological malignancy; fever ≥ 38.3°C or ≥ 38.0°C for ≥1 hour^; ANC ≤ 500cells/mm^3; outpatient	HSCT; inpatient onset FN	Total score ≥ 24 = high risk of invasive bacterial infection. Score for cancer diagnosis: AML=20, ALL/lymphoma=7, solids=0 points; Clinical presentation serious unwell or toxic = 14 points; Fever ≥39°C at presentation = 11 points; ANC<100 = 10 points	Proven invasive bacterial infection – defined as isolation of a pathogen from a sterile body site or as proven by histology. Culture-negative sepsis – defined as a systemic response to a possible infection because of hemodynamic instability, focal or multiple organ involvement or altered mental status or lethargy. Bacteraemia defined as a recognized

				pathogen cultured from one or more blood cultures or common commensals cultured from two or more blood cultures.
Alexander (22)	Cancer or haematological malignancy, fever >38.5°C at presentation or within 6h; ANC ≤ 500cells/mm^3; outpatient	Previous HSCT, inpatient onset FN	Any of following = high risk AE. AML, Burkitt lymphoma, ALL in induction, progressive or relapsed disease; Hypotension, tachypnea/hypoxia 94%; new CXR changes; altered mental status; severe mucositis; vomiting or abdominal pain; focal infection; other clinical reason for in-patient treatment	Adverse outcome – defined as identification of a pathogen (bacteraemia not defined)* or where there was a SMC* or death

Klaassen(24)	Cancer or haematological malignancy, fever > 38.5°C once or > 38.0°C during 12 hour period^; ANC ≤ 500 or between 0.5 and 1.0 cells/mm^3 and expected to fall, outpatient	New diagnosis cancer, HSCT within 6 months, comorbidity on presentation inc severe mucositis and pneumonia	AMC < 100 cells/m^3	Significant bacterial infection – defined as blood or urine culture positive for bacteria, interstitial or lobar consolidation on CXR, or unexpected death from infection (patient not palliative)
Baorto(23)	Cancer or haematological malignancy, fever ≥ 38.0°C; ANC ≤ 500cells/mm^3	Age <1y, previous HSCT	AMC < 155 cells/m^3	Bacteraemia (not defined)*
Rackoff(25)	Cancer or haematological malignancy; fever ≥ 38.5°C once or ≥ 38.0°C 3x during a 24h period^; ANC <	Inpatient onset FN	AMC < 100 cells/m^3 and temperature ≥39°C. Low risk = AMC ≥ 100 cells/m^3;	Bacteraemia – defined as a positive blood culture* .

	500cells/mm^3; outpatient		intermediate risk = AMC <100 cells/m^3 and temperature <39°C;	

HSCT, haematopoietic stem cell transplant; SMC, serious medical complication; ICU, intensive care unit; MDI, microbiologically defined infection; PCR, polymerase chain reaction; AMC, absolute monocyte count

*international consensus definition used for validation(13); ^due to available data this definition was modified for validation to ≥ 38.0°C once as per existing dataset

Table 2. Sensitivity, specificity, positive predictive value and negative predictive value of derivation study (d) and validation cohort (v) and restricted validation cohort (Rv).

Rule	Epi-sodes	Out-come n (%)	LR n (%)	True pos	True neg	False pos	False neg	Sensitivity		Specificity		PPV, % (95% CI)	NPV, % (95% CI)
								% (95% CI)	Dif from derivation % (p value)	% (95% CI)	Dif from derivation % (p value)		
Rules predicting infection and adverse outcome (refer to table 1 for specific definitions)													
d-SPOG(9)	423	122 (28.2)	165 (39)	112*	155	146	10	91.8 (85.6-95.5)		51.1 (45.9-57.1)		43.3 (37.5-49.5)	93.9 (89.2-96.7)
v-SPOG	650	168 (25.8)	289 (44.4)	131*	223	259	37	78.0 (71.1-83.6)	13.8 (0.002)	46.3 (41.9-50.7)	5.2 (0.16)	33.6 (29.1-38.4)	85.8 (81.0-89.5)
Rv-SPOG	561	149	244	119*	188	224	30	79.9 (72.7-	11.9	45.6 (40.9-	5.9 (0.13)	34.7 (29.9-	86.2 (81-

		(26.6)	(43.5)					85.6)	(0.006)	50.5)		39.9)	90.2)
d-Hakim(21)	323	47 (14.6)	223 (69)	35	211	65	12	74.5 (60.5-84.7)		76.4 (71.1-81.1)		35 (26.4-44.7)	94.6 (90.8-96.9)
v-Hakim	650	90 (13.8)	565 (86.9)	30	505	55	60	33.3 (24.5-43.6)	41.1 (<0.001)	90.2 (87.4-92.4)	13.7 (<0.001)	35.3 (26-45.9)	89.4 (86.6-91.7)
Rv-Hakim	542	78 (14.4)	462 (85.2)	28	412	52	50	35.9 (26.1-47)	38.6 (<0.001)	88.8 (85.6-91.4)	12.3 (<0.001)	35 (25.5-45.9)	89.2 (86-91.7)
d-Alexander(22)	104	22 (21.2)	55 (53)	20	53	29	2	90.9 (72.2-98.4)		64.6 (53.8-74.1)		40.8 (28.2-54.8)	96.4 (87.7-99.4)
v-Alexander	650	162 (24.9)	307 (47.2)	114	259	229	47	70.8 (63.4-77.3)	20.1 (0.07)	53.1 (48.6-57.5)	11.6 (0.06)	33.2 (28.5-38.4)	84.6 (80.2-88.2)
Rv-	342	96 (28)	160	69	133	113	27	71.9 (62.2-	19 (0.10)	54.1 (47.8-	10.6 (0.12)	37.9 (31.2-	83.1 (76.6-

Alexander			(46.8)					79.9)		60.2)		45.1)	88.1)
d-Klaassen(24)	227	43 (18.9)	83 (36.6)	36	76	107	7	83.7 (70-91.9)		41.5 (34.6-48.8)		25.3 (18.8-32.9)	91.6 (83.6-95.9)
v-Klaassen	650	108 (16.6)	169 (26)	91	152	390	17	84.3 (76.2-89.9)	0.5 (>0.99)	28.0 (24.4-32.0)	13.5 (<0.001)	18.9 (15.7-22.7)	89.9 (84.5-93.6)
Rv-Klaassen	634	104 (16.4)	168 (26.5)	87	151	379	17	83.7 (75.4-89.5)	0.1 (>0.99)	28.5 (24.8-32.5)	13.0 (<0.001)	18.7 (15.4-22.5)	89.9 (84.4-93.6)
Rules predicting bacteraemia													
d-Baorto(23)	1171	189 (16.1)	164 (14)	179	154	828	10	94.7 (90.5-97.1)		15.7 (13.5-18.1)		17.8 (15.5-20.3)	93.9 (89.1-96.7)
v-Baorto	650	61 (9.4)^	122 (18.8)	59	120	469	2	96.7 (88.8-99.4)	2.0 (0.74)	20.4 (17.3-23.8)	4.7 (0.02)	11.2 (8.7-14.1)	98.4 (94.2-99.7)
Rv-Baorto	535	54	83	54	83	398	0	100 (93.4-	5.3 (0.12)	17.3 (14.1-	1.6 (0.45)	11.9 (9.3-	100 (95.6-

		(10.1)^	(15.5)					100)		20.9)		15.3)	100)
d-Rackoff(25)	115	24 (20.9)	94 (81.7)	10	80	11	14	41.7 (24.5-61.2)		87.9 (79.6-93.1)		47.6 (28.3-67.6)	85.1 (76.5-90.9)
v-Rackoff	650	61 (9.4)^	524 (80.6)	20	483	106	41	32.8 (22.3-45.3)	8.9 (0.46)	82.0 (78.7-84.9)	5.9 (0.18)	15.9 (10.5-23.2)	92.2 (89.6-94.2)
Rv-Rackoff	556	57 (10.3)^	444 (79.9)	19	406	93	38	33.3 (22.5-46.3)	8.3 (0.61)	81.4 (77.7-84.5)	6.5 (0.18)	17 (11.1-25)	91.4 (88.5-93.7)

d, derivation study; v, validation using inclusion/exclusion criteria from existing dataset(10); Rv, validation restricted to inclusion/exclusion criteria from derivation study; LR, low risk; pos, positive; neg, negative; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; Dif, difference.

*includes episodes with adverse event known at reassessment; ^statistically significant difference in outcome as compared to derivation study.

Figure 1

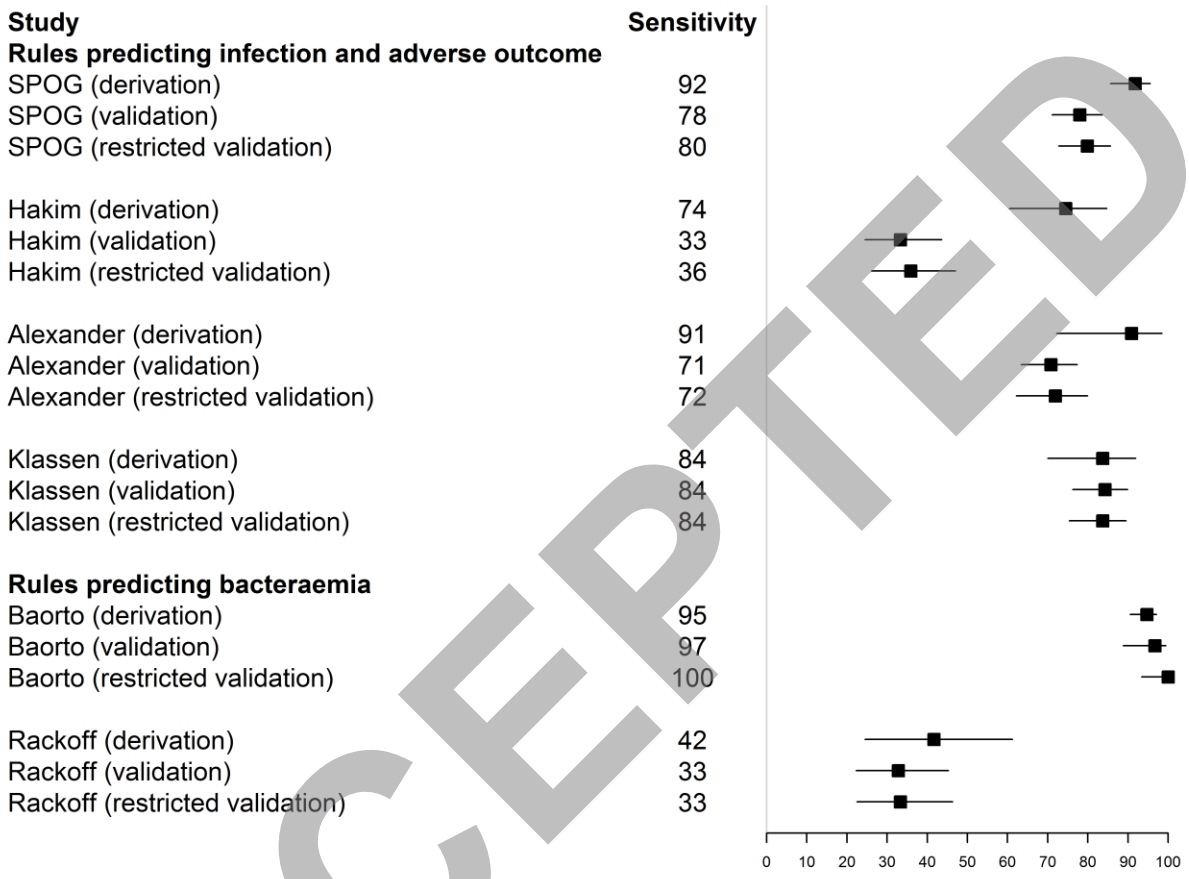

Figure 2