

British Journal of General Practice

Accuracy of the NICE Traffic Light system for detecting serious illness in acutely unwell children presenting to general practice: a retrospective cohort study

Clark, Amy; Cannings-John, Rebecca ; Blyth, Megan; Hay, Alastair; Butler, Christopher; Hughes, Kathryn

DOI: <https://doi.org/10.3399/BJGP.2021.0633>

To access the most recent version of this article, please click the DOI URL in the line above.

Received 08-November-2021

Revised 12-January-2022

Accepted 04-February-2022

© 2022 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>). Published by British Journal of General Practice. For editorial process and policies, see: <https://bjgp.org/authors/bjgp-editorial-process-and-policies>

When citing this article please include the DOI provided above.

Author Accepted Manuscript

This is an 'author accepted manuscript': a manuscript that has been accepted for publication in British Journal of General Practice, but which has not yet undergone subediting, typesetting, or correction. Errors discovered and corrected during this process may materially alter the content of this manuscript, and the latest published version (the Version of Record) should be used in preference to any preceding versions

Accuracy of the NICE Traffic Light system for detecting serious illness in acutely unwell children presenting to general practice: a retrospective cohort study.

Amy Clark¹, Rebecca Cannings-John², Megan Blyth³, Alastair D Hay⁴, Christopher C Butler⁵, Kathryn Hughes⁶

¹School of Medicine, Cardiff University, CF14 4YS, Cardiff, Wales, UK.

Amy Clark (BSc.(hons))

Medical Student, Email: clarkAE1@cardiff.ac.uk. ORCID iD [0000-0002-9409-0486](https://orcid.org/0000-0002-9409-0486)

²Centre for Trials Research, Cardiff University, CF14 4YS, Cardiff, Wales, UK.

Dr Rebecca Cannings-John (BSc.(hons), MSc, PhD)

Senior Research Fellow, Email: canningsrl@cardiff.ac.uk. ORCID iD [0000-0001-5235-6517](https://orcid.org/0000-0001-5235-6517)

³Division of Population Medicine, School of Medicine, Cardiff University, CF14 4YS, Cardiff, Wales, UK.

Dr Megan Blyth (MBBCh, MRCP)

Clinical lecturer in general practice, Email: BlythM1@Cardiff.ac.uk. ORCID iD [0000-0002-3800-1143](https://orcid.org/0000-0002-3800-1143)

⁴Centre for Academic Primary Care, Bristol Medical School, University of Bristol, BS8 2PS, Bristol, UK.

Professor Alastair D Hay (MBChB, MD, MRCP, DCH, MRCP, DFFP, FRCGP, FHEA)

Professor of Primary Care, Email: Alastair.Hay@bristol.ac.uk. ORCID iD [0000-0003-3012-375X](https://orcid.org/0000-0003-3012-375X)

⁵Nuffield Department of Primary Care Health Sciences, University of Oxford, OX2 6GG, Oxford, UK.

Professor Christopher C Butler (BA, MBChB, DCH, CCH, MD, FRCGP, FFP, FMedSci)

Professor of Primary Care, Email: christopher.butler@phc.ox.ac.uk. ORCID iD [0000-0002-0102-3453](https://orcid.org/0000-0002-0102-3453)

⁶PRIME Centre Wales, Division of Population Medicine, School of Medicine, Cardiff University, CF14 4YS, Cardiff, Wales, UK.

Dr Kathryn Hughes (MBBCh, MRCP, DCH, PhD)

Senior Clinical Lecturer, Email: hugheska6@cardiff.ac.uk. ORCID iD [0000-0002-8099-066X](https://orcid.org/0000-0002-8099-066X)

Corresponding Author: Dr Kathryn Hughes, School of Medicine, Neuadd Meirionnydd, Heath Park Way, Cardiff University, Heath Park, Cardiff, Wales, CF14 4YS. Email: hugheska6@cardiff.ac.uk

Abstract

Background

The NICE Traffic Light system was created to facilitate the assessment of unwell children in primary care. No studies have validated this tool in UK general practice.

Aim

To evaluate the accuracy of this system for detecting serious illness in children presenting to general practice.

Design and Setting

Retrospective diagnostic accuracy study, using a cohort of acutely unwell children under five years presenting to general practice in England and Wales.

Method

The Traffic Light categories of 6,703 children were linked with hospital data to identify admissions and diagnoses. The sensitivity and specificity of these categories were calculated against the reference standard: a hospital diagnosed serious illness within seven days of GP consultation, measured using ICD-10 codes.

Results

2,116 (31·6%) children were categorised as red; 4,204 (62·7%) amber; and 383 (5·7%) green. 139 (2·1%) children were admitted to hospital within seven days of consultation, of whom 17 (12·2%; 0·3% overall) had a serious illness. The sensitivity of the 'red' category (vs. amber/green) was 58·8% (95% confidence interval: 32·9% to 81·6%) and the specificity 68·5% (67·4% to 69·6%). The sensitivity and specificity of 'red' and 'amber' combined (vs. green) was 100% (80·5% to 100%) and 5·7% (5·2% to 6·3%) respectively.

Conclusion

The Traffic Light system cannot accurately detect children admitted with a serious illness, nor those not seriously ill who can be managed at home. This system is not suitable for use as a clinical tool in general practice. Further research is required to update or replace the Traffic Light system.

Keywords: General Practice, Child Health, Clinical Prediction Rules

How this fits in

The NICE Traffic Light system is widely used in general practice for the assessment of unwell children; however, the majority of previous studies validating this tool have been conducted in secondary care settings. No studies have validated this tool within UK general practice. This study found that the Traffic Light system cannot accurately detect or exclude serious illness in

children presenting to UK general practice with an acute illness. We conclude that it cannot be relied upon by clinicians for the assessment of acutely unwell children and that it is unsuitable for use as a clinical decision tool.

INTRODUCTION

Children with acute illnesses are a common presentation in general practice, and can be a diagnostic challenge for clinicians.^{1,2} This is primarily due to their presentation at an early stage of illness, where the signs and symptoms of non-serious and serious illnesses can appear similar.³ The majority of acute illnesses are self-limiting viral infections, although a minority of children may have a serious illness such as pneumonia or meningitis.⁴ The estimated incidence of these illnesses in children presenting to general practice is 1% per year.^{3,5} Despite this, fear and anxiety in parents with an unwell child are common.^{6,7}

In an effort to help primary care clinicians confidently assess unwell children, the National Institute for Health and Care Excellence (NICE) created the 'Traffic Light' system.⁸ This tool categorises children as either green, amber, or red depending on their consulting clinical features; corresponding to a low, intermediate, or high risk of serious illness respectively. According to the tool, 'green' children can be managed at home. Children in the 'amber' category can either be referred to hospital for assessment or sent home with safety-net advice. Children in the 'red' category should be referred urgently for assessment in hospital.

The recent relaxation of restrictions during the coronavirus disease 2019 (COVID-19) pandemic has led to a dramatic increase in the prevalence of respiratory illness in young children, with emergency departments experiencing high demand from children, many of which are not seriously ill.^{9,10} This illustrates the importance of an accurate primary care tool to identify those needing secondary care assessment.

Validation of the Traffic Light system within general practice has been performed once previously in 2013, using a small Dutch cohort of 506 febrile children.¹¹ Additional studies have evaluated this tool within emergency department settings.^{4,11-13} However, no studies have evaluated the NICE Traffic Light system in UK general practice.

The aim of this study is to evaluate the accuracy of the NICE Traffic Light system for predicting serious illness in acutely unwell children aged under five presenting to UK general practice.

METHODS

The Standards for Reporting of Diagnostic Accuracy (STARD) guidelines were followed in the reporting of this study.¹⁴

Sample Formation

This retrospective cohort study involved the secondary analysis of a dataset collected for the Diagnosis of Urinary Tract Infection in Young Children (DUTY) study.¹⁵ The DUTY study was a prospective cohort study of acutely unwell children aged under five years in primary care, which evaluated the presenting signs and symptoms of urinary tract infections (UTIs). Children from 233 sites in England and Wales (general practices, walk-in centres, and emergency departments) were consecutively recruited between 2010 and 2012 if they were constitutionally unwell due to any acute illness and/or presenting with a UTI symptom described by NICE. The

eligibility criteria for recruitment into the DUTY study are presented in Supplementary Table 1. The study presented here only includes the children from the DUTY study who presented to general practice (GP).

Data Processing

The children's clinical features at the time of their GP presentation were mapped retrospectively to equivalent variables within the Traffic Light system as part of a separate study.¹⁶ Children were categorised as 'red' if they had at least one red feature, 'amber' if they had at least one amber feature (and no red), and 'green' if they had no amber or red features. Children with missing data, preventing Traffic Light categorisation, were excluded from our analyses.

Routinely collected hospital data in England and Wales were accessed to identify admissions; provided by Hospital Episode Statistics (collected by NHS Digital) and the Patient Episode Database for Wales, respectively. The data collected from children during the DUTY study were linked to hospital data using the Secure Anonymised Information Linkage (SAIL) Databank.

Outcome Definitions

The primary outcome of interest was an unplanned hospital admission with a serious illness, within seven days of initial GP consultation.

A 'hospital admission' was defined as a spell in hospital under the care of a consultant; assessment in the emergency department was not recorded as an admission unless the treating team decided to admit them. To ensure a strong association between consulting clinical features and any subsequent hospital admission, a seven-day follow-up period was chosen.^{4,12,17-19}

Our definition of 'serious illness' was created by identifying the NICE definitions used during the creation of the Traffic Light system and exploring previous literature.^{2,4,13,20-22} The International Classification of Diseases: 10th revision (ICD-10) codes for these serious illnesses were used to create a reference table (Supplementary Table 2). The principal serious illnesses included: sepsis, pneumonia, meningitis, and UTI.

Statistical Analysis

To evaluate the test performance of the NICE Traffic Light system, sensitivity and specificity were calculated separately for two thresholds of test positivity: designation of 'red' category and designation of 'red' or 'amber' category. This allowed comparison of test performance between 'high-risk' and 'intermediate to high-risk' populations. Further calculations included 95% confidence intervals (CIs) and positive and negative predictive values.

Two sensitivity analyses were conducted. The first assessed the ability of the Traffic Light system to detect children admitted to hospital (with or without a serious illness), to reduce the impact of incorrect diagnoses coding within the routine data. The second evaluated the Traffic Light system when applied to a cohort of febrile children, chosen because the Traffic Light system was created by NICE to assess febrile children specifically. The NICE definitions of fever were used to identify this subgroup: "measured or perceived elevation of body temperature above the normal daily variation ($\geq 38^{\circ}\text{C}$) by a parent or clinician".⁸

Analyses were performed using SPSS v.26.²³ To comply with SAIL regulations, data with frequencies less than five were suppressed and presented as '<5' or rounded to the nearest five.

RESULTS

During the DUTY study period, 7,374 children were recruited from primary care. Children presenting to care providers outside of general practice were excluded, in addition to those without available Traffic Light and hospital admissions data. Overall, 6,703 (91.0%) children were eligible for inclusion in this study (Figure 1). This population of children presenting to general practice will be referred to as the 'DUTY cohort'.

Characteristics of the DUTY Cohort

Nearly a quarter of the 6,703 children attending general practice were between 1 and 2 years old ($n=1,577$, 23.5%, Table 1). The majority of children were categorised as 'amber' ($n=4,204$, 62.7%), followed by 'red' ($n=2,116$, 31.6%) and then 'green' ($n=383$, 5.7%). The 'green' children were older, with a median age of 2.8 years (25th to 75th centiles: 1.4 to 3.8 years) compared to 2.0 years for both 'amber' (0.9 to 3.4 years) and 'red' (1.1 to 3.3 years) children, respectively. The variable which triggered 'amber' and 'red' labels most commonly was '*impaired social cues (moderate or severe)*', responsible for '97.0%' and '61.2%' of these categorisations respectively. Abnormal vital signs including temperature, pulse rate, oxygen saturation, and respiratory rate contributed to an 'amber' classification in 28.8% of 'amber' children and a 'red' classification in 1.7% of 'red' children. Further analyses of the Traffic Light categorisations are available as part of a separate paper.¹⁶

Hospital Admissions

Linkage to hospital admissions data was achieved for 98.7% of children with available Traffic Light categories (6,703/6,791). To ensure there was no selection bias between children who could not be linked to hospital data ($n=88$) and children who could ($n=6,703$), a series of descriptive statistics were undertaken (Supplementary Table 3). These confirmed no clinically significant differences in median age, duration of illness, or distribution of Traffic Light categories.

Within seven days of presenting to general practice, 139 of 6,703 children (2.1%) were admitted to hospital. The children admitted to hospital were younger than the children who were not admitted (Table 1). Additionally, more children were 'red' at initial presentation to general practice (Table 1).

The median duration between general practice consultation and admission was 1 day (25th to 75th centiles: 0 to 3 days), with just under half of cases admitted the same day ($n=57$, 41.0%). Children were most commonly discharged on the same day ($n=81$, 58.3%) or the day after ($n=36$, 25.9%). The median length of hospital stay was 0 days (0 to 1 day). The most common diagnosis in hospital was an unspecified viral infection (14.4%, Supplementary Table 4).

Serious Illnesses

The prevalence of serious illness in our cohort was 0.3% (17/6,703, 95% CI: 0.2% to 0.4%). The majority of serious illnesses were cases of pneumonia ($n=8$, 47.1%). No cases of sepsis or meningitis were reported. Information on all diagnosed serious illnesses cannot be disclosed due

to small numbers. Of the children diagnosed with a serious illness, 10 (58.8%) were categorised as 'red' at presentation to general practice (Figure 2). The median duration between general practice consultation and hospital admission in this group was 2 days (25th to 75th centiles: 0 to 2 days) and the median length of hospital stay was 1 day (0 to 2 days). Children were most commonly discharged on the same day (n=7, 41.2%) or two days later (n<5).

Performance of the NICE Traffic Light system

The 'red' category had a sensitivity of 58.8% (95% CI: 32.9% to 81.6%) and specificity of 68.5% (67.4% to 69.6%) for detecting children admitted to hospital with a serious illness (Table 2). Combining the 'red' and 'amber' categories improved the sensitivity to 100% (80.5% to 100%) but reduced the specificity to 5.7% (5.2% to 6.3%).

The performance of the Traffic Light system for detecting any admission to hospital was also calculated (Table 3). The 'red' category had a sensitivity of 61.9% (53.3% to 70.0%) and specificity of 69.1% (67.9% to 70.2%). Combining 'red' and 'amber' improved the sensitivity to 97.8% but reduced specificity to 5.8%. The cross-tabulations used to calculate these measures of test performance are displayed in Supplementary Table 5.

Sensitivity Analysis for Febrile Children

In our sample, 5,032 children (75.1%) were febrile. Of these, 120 (2.4%) were hospitalised and 16 (0.3%) were diagnosed with a serious illness. Therefore, 86.3% (120/139) of children admitted to hospital and 94.1% (16/17) of children with a serious illness were febrile at GP presentation. The Traffic Light system had improved sensitivity but poorer specificity for detecting serious illness when used in this population (Table 4).

DISCUSSION

Summary

We found that the prevalence of serious illness in 6,703 acutely unwell young children presenting to UK general practice was 0.3% and that the NICE Traffic Light System categorised 31.6% of all children as red; 62.7% as amber; and 5.7% as green. Overall, 139 (2.1%) children were admitted within seven days of their initial presentation. The Traffic Light tool had a sensitivity of 58.8% and specificity of 68.5% for the identification of children admitted to hospital with a serious illness, when comparing 'red' with 'amber'/'green' categories. Changing the threshold to include 'red'/'amber' categories combined, compared to 'green', improved the sensitivity to 100% but worsened the specificity to 5.7%. The results were robust to detecting hospital admissions for any reason, or when applied to febrile children only.

Strengths and Limitations

To our knowledge, this is the first study to validate the Traffic Light system in UK general practice. We used a dataset of 6,703 children, representing one of the largest and most detailed characterisations of clinical features among acutely unwell young children presenting to general practice in the UK, linked to hospital admissions data. Although the NICE guidelines and Traffic Light system were designed for febrile children, our inclusion of all children with an acute

illness better represents the variety of children requiring assessment by GPs in clinical practice. Furthermore, we included all illnesses defined by NICE as 'serious', which matches the illnesses that the Traffic Light system was designed to detect. We were also able to provide a current prevalence estimate of serious illness in children presenting to general practice.

There are several limitations to this study. The prospective data were not collected specifically to answer this research question. Consequently, not all clinical features present in the Traffic Light system could be matched to the DUTY dataset during the assignment of Traffic Light categories.¹⁶ The variables which could not be mapped were mostly within the 'red' category and principally involved neurological features such as neck stiffness, or orthopaedic signs such as limb swelling. If these features had been matched, potentially more children would have been 'red', and the sensitivity of the Traffic Light system may have improved. However, most of the key constitutional features of the Traffic Light system were captured by the DUTY study, and 64% of data fields were mapped.¹⁶

Our 'serious illness' reference standard was dependant on diagnostic codes within routinely collected hospital data; reference standards such as clinical, laboratory, and radiological data were not available to assess the evidence supporting a final diagnosis of a serious illness. It is possible that some children with a 'serious illness' were not identified in our study, due to incorrect coding of their diagnoses, thus underestimating our disease prevalence and tool sensitivity. However, the coding in our dataset represents the diagnoses recorded on discharge by a clinician with access to all investigations.

Finally, our sample only included children who fulfilled the DUTY study eligibility criteria. Consequently, our cohort may not be entirely representative of all acutely unwell children within this age group. The criteria did not include children who were constitutionally well unless they had symptoms suggestive of a urinary tract infection. Children with focal illnesses or mild respiratory infections may have been excluded, and children with urinary tract infections may have been overrepresented due to the aim of the DUTY study. However, the Traffic Light system was principally designed for constitutionally unwell children; therefore, we believe our cohort is adequately representative of this. There is a possibility that some of the children presenting to general practice were too ill to be recruited for the original study. However, our low disease prevalence is similar to previous estimates of serious illnesses presenting to general practice.⁵

Comparison With Existing Literature

To our knowledge, only one other study has validated the Traffic Light system within general practice.¹¹ This was a retrospective cohort study conducted by Verbakel et al in the Netherlands, using a sample of 506 febrile children aged 3 months to 6 years. They reported a sensitivity of 100% and specificity of 1% for the identification of serious bacterial infections, using the presence of any 'red' or 'amber' features as a positive test.¹¹ These results are similar to ours when 'red/amber' categories were combined. However, it was not clear how the authors defined 'serious infection'; either as clinical judgement, hospital admission, or investigations performed in secondary care. Furthermore, information regarding the designation of Traffic Light categories was not provided.

Further studies have assessed the Traffic Light system in emergency department settings, reporting sensitivities between 85% and 99%, and specificities between 2% and 29% (using 'red' or 'amber' combined as a positive test).^{4,11-13} Our results for 'red' and 'amber' combined were similar, with a particularly poor specificity. This may be due to the lower prevalence of serious illnesses seen in a GP setting. Notably, these studies limited their outcome to serious bacterial infections only, and included children up to the age of 16 years in some cases.^{11,12}

Implications for Research and Practice

We conclude that the NICE Traffic Light system is not able to accurately detect or exclude serious illness in acutely unwell children presenting to general practice when the 'red' category is used as a positive threshold. If our cohort's Traffic Light classifications had been followed by GPs, one-third of children (categorised as 'red') would have been urgently referred to hospital. Additionally, using the 'red' category as a threshold for hospital referral would have missed 41.2% of children with a serious illness who were 'amber', although NICE does recommend that clinicians should refer 'amber' children if indicated. Combining 'red' and 'amber' categories improved the sensitivity of the Traffic Light system, such that all seriously ill children were identified. This threshold would allow GPs to be confident sending 'green' children home, but at the cost of referring a substantial number of children to hospital; 94.3% of patients were categorised as 'red' or 'amber'. Moreover, GPs would only be able to confidently exclude serious illness in a minority (5.7%) of children classified as 'green'.

The Traffic Light system was created to help GPs confidently assess unwell children, aiding their decisions about who to refer and who to send home by identifying those at risk of serious illness (thus prioritising sensitivity over specificity), but we have shown that it is unable to accurately achieve this. This is an important finding in light of the current strain experienced by primary care services as a consequence of the COVID-19 pandemic. Secondary care can only function to assess and treat patients with serious illnesses provided there is effective functioning of primary care in serving the remainder.

Research is required to derive an updated tool for the assessment of acutely unwell children presenting to general practice. This assessment tool should correctly identify the most unwell children, while preventing unnecessary hospital referrals for children who are more likely to have a self-limiting illness. It should be derived and validated using data from UK general practice or primary care and must be suitable for use in a typical general practice consultation. This may require combining multiple datasets, due to the low prevalence of serious illness within this population. Future research could also assess whether 'point-of-care' markers of illness, such as C-reactive Protein (CRP) and Procalcitonin (PCT), combined with clinical features in a single assessment tool could provide a more accurate indication of illness severity in children.

FUNDING

No funding was granted for the completion of this study. The lead author was granted a monetary award from the Wolfson Foundation (Wolfson Intercalated Award) to help their completion of this research as part of their intercalated Bachelor of Science degree. This funding was not associated with the study design, result analyses, write-up, or the decision to submit this

article for publication. Additionally, no members of the Wolfson Foundation were involved in the completion of this research.

ETHICAL APPROVAL

Access to the data required for the completion of this study was approved by the Health and Care Research Wales Research Ethics Committee and the Health Research Authority CAG committee (Confidentiality Advisory Group, ref: 16/WA/0166).

DECLARATION OF INTERESTS

We declare no competing interests.

ACKNOWLEDGEMENTS

Many thanks to all those involved in the DUTY study. This study makes use of the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers who make anonymised data available for research. The Centre for Trials Research and PRIME Centre Wales is funded by Health & Care Research Wales and Cancer Research UK.

COPYRIGHT STATEMENT

NHS Digital data re-use statement: Copyright © (2019), the Health and Social Care Information Centre. Re-used with the permission of the Health and Social Care Information Centre. All rights reserved.

DATA SHARING STATEMENT

The patient datasets analysed in this study cannot be made available due to confidentiality and data-sharing agreements in place.

REFERENCES

1. Hay AD, Heron J, Ness A, The Alspac study team. The prevalence of symptoms and consultations in pre-school children in the Avon Longitudinal Study of Parents and Children (ALSPAC): a prospective cohort study. *Fam Pract.* 2005;**22**(4):367-74.
2. National Collaborating Centre for Women's and Children's Health. National Institute for Health and Care Excellence: Clinical Guidelines. Feverish Illness in Children: Assessment and Initial Management in Children Younger Than 5 Years. London: Royal College of Obstetricians & Gynaecologists (UK); 2013.
3. Van den Bruel A, Hall H, Aertgeerts B, et al. Signs and symptoms for diagnosis of serious infections in children: a prospective study in primary care. *Br J Gen Pract.* 2007;**57**(540):538-46.
4. De S, Williams GJ, Hayen A, et al. Accuracy of the "traffic light" clinical decision rule for serious bacterial infections in young children with fever: a retrospective cohort study. *BMJ.* 2013;**346**:f866.

5. Van den Bruel A, Bartholomeeusen S, Aertgeerts B, et al. Serious infections in children: an incidence study in family practice. *BMC Fam Pract.* 2006;**7**(1):23.
6. Betz MG, Grunfeld AF. 'Fever phobia' in the emergency department: a survey of children's caregivers. *Eur J Emerg Med.* 2006;**13**(3).
7. Crocetti M, Moghbeli N, Serwint J. Fever Phobia Revisited: Have Parental Misconceptions About Fever Changed in 20 Years? *Pediatrics.* 2001;**107**(6):1241.
8. National Institute for Health and Care Excellence. Fever in under 5s: assessment and initial management NICE Guidelines2019 [Available from: <https://www.nice.org.uk/guidance/ng143>.
9. Royal College of Paediatrics and Child Health. A&Es see dramatic rise in number of young children – but it's not COVID: RCPCH; 2021 [Available from: <https://www.rcpch.ac.uk/news-events/news/aes-see-dramatic-rise-number-young-children-its-not-covid>.
10. Public Health England. Health chiefs issue warning as childhood respiratory infections rise ahead of winter: UK Government; 2021 [Available from: <https://www.gov.uk/government/news/health-chiefs-issue-warning-as-childhood-respiratory-infections-rise-ahead-of-winter>.
11. Verbakel JY, Van den Bruel A, Thompson M, et al. How well do clinical prediction rules perform in identifying serious infections in acutely ill children across an international network of ambulatory care datasets? *BMC Med.* 2013;**11**(1):10.
12. Thompson M, Coad N, Harnden A, et al. How well do vital signs identify children with serious infections in paediatric emergency care? *Arch Dis Child.* 2009;**94**(11):888-93.
13. Yao SHW, Ong GY-K, Maconochie IK, et al. Analysis of emergency department prediction tools in evaluating febrile young infants at risk for serious infections. *Emerg Med J.* 2019;**36**(12):729-35.
14. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Clin Chem.* 2015;**61**(12):1446-52.
15. Hay AD, Sterne JAC, Hood K, et al. Improving the Diagnosis and Treatment of Urinary Tract Infection in Young Children in Primary Care: Results from the DUTY Prospective Diagnostic Cohort Study. *Ann Fam Med.* 2016;**14**(4):325.
16. Blyth MH, Cannings-John R, Hay AD, et al. Is the NICE traffic light system fit-for-purpose for children presenting with undifferentiated acute illness in primary care? *Arch Dis Child.* 2021:archdischild-2021-322768.
17. de Vos-Kerkhof E, Nijman RG, Vergouwe Y, et al. Impact of a clinical decision model for febrile children at risk for serious bacterial infections at the emergency department: a randomized controlled trial. *PLoS ONE.* 2015;**10**(5):e0127620.
18. Nijman RG, Vergouwe Y, Thompson M, et al. Clinical prediction model to aid emergency doctors managing febrile children at risk of serious bacterial infections: diagnostic study. *BMJ.* 2013;**346**:f1706.
19. Brent AJ, Lakhanpaul M, Thompson M, et al. Risk score to stratify children with suspected serious bacterial infection: observational cohort study. *Arch Dis Child.* 2011;**96**(4):361-7.
20. Bustinduy AL, Chis Ster I, Shaw R, et al. Predictors of fever-related admissions to a paediatric assessment unit, ward and reattendances in a South London emergency department: the CABIN 2 study. *Arch Dis Child.* 2017;**102**(1):22.
21. Kerkhof E, Lakhanpaul M, Ray S, et al. The Predictive Value of the NICE "Red Traffic Lights" in Acutely Ill Children. *PLOS ONE.* 2014;**9**(3):e90847.
22. Van den Bruel A, Haj-Hassan T, Thompson M, et al. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet.* 2010;**375**(9717):834-45.
23. IBM Corp. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY2019.

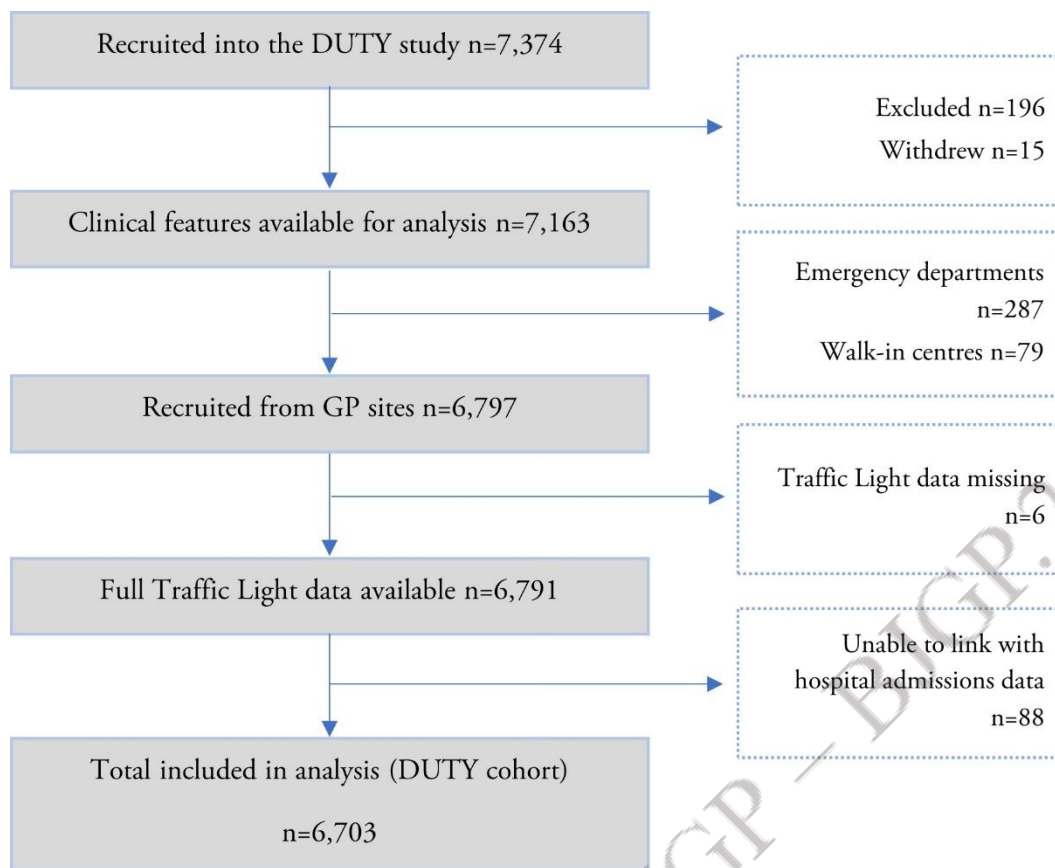


Figure 1: Flow chart of children included in analysis.⁽¹⁵⁾

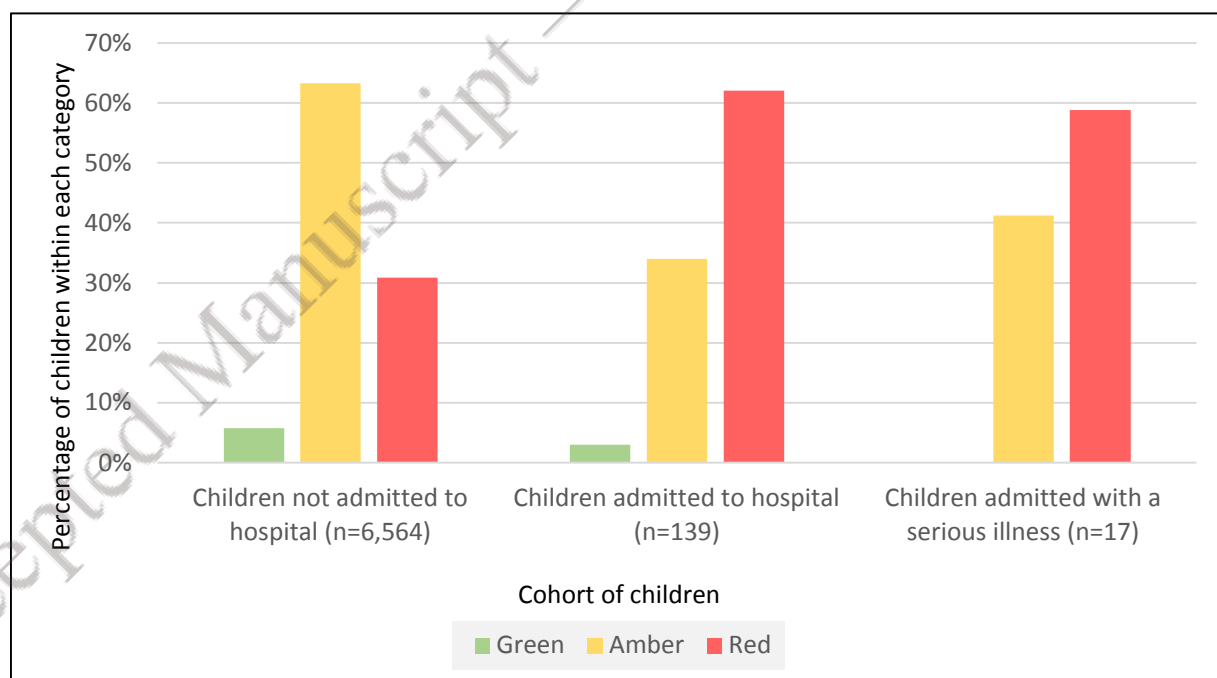


Figure 2: Bar chart demonstrating the Traffic Light categories for children admitted and not admitted to hospital, and those with a serious illness. Note: these figures have been rounded to the nearest five to comply with the 'small data' reporting requirements of the SAIL databank.

Tables

Table 1: Characteristics of the DUTY children presenting to general practice by hospital admission status within seven days. All results are number (%) unless otherwise specified.

Characteristic	DUTY Cohort (n=6,703)	DUTY children not admitted to hospital (n=6,564)*	DUTY children admitted to hospital (n=139)*
Age at recruitment (months):			
< 3	198 (3.0)	191 (2.9)	7 (5.0)
3 – 5	409 (6.1)	395 (6.0)	14 (10.1)
6 – 11	1,065 (15.9)	1,041 (15.9)	24 (17.3)
12 – 23	1,577 (23.5)	1,541 (23.5)	36 (25.9)
24 – 35	1,263 (18.8)	1,248 (19.0)	15 (10.8)
36 – 47	1,257 (18.8)	1,227 (18.7)	30 (21.6)
≥ 48	934 (13.9)	921 (14.0)	13 (9.4)
Median age (years) (25th to 75th centiles)	2.1 (1.0 to 3.4)	2.1 (1.0 to 3.4)	1.6 (0.9 to 3.1)
Sex:			
Male	3,282 (49.0)	3,213 (48.9)	69 (49.6)
Female	3,421 (51.0)	3,351 (51.1)	70 (50.4)
Traffic Light Category:			
Green	383 (5.7)	378 (5.8)**	5 (3.4)**
Amber	4,204 (62.7)	4,154 (63.3)**	50 (34.5)**
Red	2,116 (31.6)	2,026 (30.9)**	90 (62.1)**

*Within seven days of the GP consultation

**Figures have been rounded and adjusted to comply with SAIL 'small data' reporting regulations.

Table 2: Prognostic performance of the NICE Traffic Light system for detecting serious illness.

Test performance for detecting children with a serious illness	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value	Negative Predictive Value
Red category only	58.8% (32.9% to 81.6%)	68.5% (67.4% to 69.6%)	0.5%	99.8%
Red or amber category	100% (80.5% to 100%)	5.7% (5.2% to 6.3%)	0.3%	100%

Table 3: Prognostic performance of the NICE Traffic Light system for detecting hospital admission.

Test performance for detecting children admitted to hospital	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value	Negative Predictive Value
Red category only	61·9% (53·3% to 70·0%)	69·1% (67·9% to 70·2%)	4·1%	98·8%
Red or amber category	97·8% (93·8% to 99·6%)	5·8% (5·2% to 6·4%)	2·2%	99·2%

Table 4: Prognostic performance of the NICE Traffic Light system for detecting febrile children with a serious illness.

Test performance for detecting serious illness in febrile children	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value	Negative Predictive Value
Red category only	62·5% (35·4% to 84·8%)	64·0% (62·7% to 65·3%)	0·6%	99·8%
Red or amber	100% (79·4% to 100%)	3·0% (2·5% to 3·5%)	0·3%	100%