

# Can paediatric early warning scores (PEWS) be used to guide the need for hospital admission and predict significant illness in children presenting to the emergency department? An assessment of PEWS diagnostic accuracy using sensitivity and specificity

Peter J Lillitos,<sup>1,2</sup> Graeme Hadley,<sup>1,2</sup> Ian Maconochie<sup>1,2</sup>

► Additional material is published online. To view please visit the journal (<http://dx.doi.org/10.1136/emered-2014-204355>).

<sup>1</sup>Imperial College Faculty of Medicine, London, UK

<sup>2</sup>Department of Emergency Medicine, St Mary's Hospital, London, UK

## Correspondence to

Peter J Lillitos, Paediatric Emergency Department, Imperial College Healthcare NHS Trust, St Mary's Hospital, Praed Street, London W2 1NY, UK; [plillitos@doctors.org.uk](mailto:plillitos@doctors.org.uk)

Received 23 September 2014

Revised 8 August 2015

Accepted 2 October 2015

Published Online First

3 November 2015

## ABSTRACT

**Objectives** Designed to detect early deterioration of the hospitalised child, paediatric early warning scores (PEWS) validity in the emergency department (ED) is less validated. We aimed to evaluate sensitivity and specificity of two commonly used PEWS (Brighton and COAST) in predicting hospital admission and, for the first time, significant illness.

**Methods** Retrospective analysis of PEWS data for paediatric ED attendances at St Mary's Hospital, London, UK, in November 2012. Patients with missing data were excluded. Diagnoses were grouped: *medical* and *surgical*. To classify diagnoses as *significant*, established guidelines were used and, where not available, common agreement between three acute paediatricians.

**Results** 1921 patients were analysed. There were 211 admissions (11%). 1630 attendances were medical (86%) and 273 (14%) surgical. Brighton and COAST PEWS performed similarly. *hospital admission*: PEWS of  $\geq 3$  was specific (93%) but poorly sensitive (32%). The area under the receiver operating curve (AUC) was low at 0.690. *Significant illness*: for medical illness, PEWS  $\geq 3$  was highly specific (96%) but poorly sensitive (44%). The AUC was 0.754 and 0.755 for Brighton and COAST PEWS, respectively. Both scores performed poorly for predicting significant surgical illness (AUC 0.642). PEWS  $\geq 3$  performed well in predicting significant respiratory illness: sensitivity 75%, specificity 91%.

**Conclusions** Both Brighton and COAST PEWS scores performed similarly. A score of  $\geq 3$  has good specificity but poor sensitivity for predicting hospital admission and significant illness. Therefore, a high PEWS should be taken seriously but a low score is poor at ruling out the requirement for admission or serious underlying illness. PEWS was better at detecting significant medical illness compared with detecting the need for admission. PEWS performed poorly in detecting significant surgical illness. PEWS may be particularly useful in evaluating respiratory illness in a paediatric ED.

## INTRODUCTION

Various paediatric early warning scores (PEWS) have been used since the validation of the Brighton score in 2005<sup>1</sup> and were designed to detect the early deterioration of the hospitalised child.

The UK National Patient Safety Agency and the National Institute for Health and Care Excellence, together with the UK Confidential Enquiry into

## Key messages

### What is already known on this subject?

Previous studies have shown paediatric early warning scores (PEWS) to be specific but not sensitive in predicting admission to hospital from the paediatric emergency department (ED). No study has yet looked at PEWS validity in predicting significant illness in children presenting to the ED.

### What might this study add?

Our study suggests that PEWS is specific but not sensitive for predicting significant illness in the ED and is better at predicting significant medical illness, especially respiratory illness, compared with significant surgical illness. PEWS should, therefore, be taken seriously if abnormal but a low PEWS should not falsely reassure.

Maternal and Childhood report 'Why Children Die', recommended early warning scores to identify children in hospital with developing critical illness.<sup>2,3</sup> Despite their widespread implementation, studies examining their efficacy vary in quality, making validation variable for different scores.<sup>4,5</sup> With the introduction of the 4 h target in UK emergency departments (EDs), there is increasing pressure to make decisions on patient management.<sup>6</sup> PEWS, which include the commonly used Brighton and COAST systems, were designed to reflect trends in physiological condition, thus allowing the early detection of deterioration and hence prompt timely intervention for the hospitalised child. However, the validity of a one-off 'snapshot' of physiological parameters in the ED to predict the need for hospital admission or illness severity is not validated. With PEWS use in the ED increasing,<sup>7</sup> it is a necessity to know the predictive power of this tool to predict (1) hospital admission and (2) significant illness. Only two studies<sup>6,8</sup> have specifically examined PEWS ability to predict hospital admission from the children's ED, using 424 and 1223 study patients, respectively. In both studies, the Brighton score was used. PEWS was found to be specific but not sensitive. No studies to our knowledge have examined their ability to predict serious illness in the children's ED. Two commonly used



CrossMark

**To cite:** Lillitos PJ, Hadley G, Maconochie I. *Emerg Med J* 2016;**33**:329–337.

scoring systems that are advocated in the UK by NHS Improving Quality (previously known as the NHS Institute for Innovation and Improvement)<sup>9</sup> are Brighton and COAST. Both use an abnormal HR, RR, work of breathing, level of consciousness, the need for supplemental oxygen and parental/medical staff concern to create an illness score. COAST was adapted from Brighton PEWS for use specifically in the ED whereby hypoxia, thought to be a more useful marker of illness severity rather than supplemental oxygen, is used to trigger a score. This study's aims are to estimate the diagnostic accuracy of two commonly used PEWS score (Brighton and COAST) (1) to predict hospital admission and (2) to detect significant illnesses among unselected paediatric patients presenting to the ED.

## METHODS

This was a retrospective analysis of routinely collected clinical observations recorded on arrival to the paediatric ED at St Mary's Hospital, London (Imperial College NHS Trust, UK). The patients studied attended in the month of November 2012, during which there were 2261 attendances. November was chosen as this was a busy month for paediatric ED attendances, likely owing to it being winter. Therefore, a larger sample size of patients was possible to study. Previous studies have used study populations between 87 and 1336;<sup>6 8 10–13</sup> therefore, we regarded our sample to be of sufficient size. Data to calculate PEWS were retrieved from the ED *Ascribe Symphony* system and subsequently anonymised. The data for every attendance in the ED were included. All included PEWS were complete. Where data were missing, the patient's electronically scanned notes were examined, and if data were still missing these patients were excluded. Any incomplete scores were excluded as part of the 340 patients excluded.

Table 1 shows how PEWS were calculated for both Brighton and COAST scores. Each abnormal parameter scores 1 point with a maximum score possible of 6. The data used to calculate PEWS came from the first assessment when the child arrived in the ED.

Both Brighton and COAST score 1 point for parental concern. We automatically assigned a score of '1' for every patient presenting to the ED on the basis that parents/carers, by bringing their child to the ED, were by default concerned.

Each patient outcome, that is, admitted, not admitted or sent to another hospital for further assessment, was recorded. If a parent/carer left the department before being seen by a clinician, this was classified as a non-admission. If a child was transferred to another facility urgently, this counted as an admission.

Diagnoses recorded on *Ascribe Symphony* were extracted for each patient. Owing to the large heterogeneity in diagnoses, diagnoses were grouped into two broad categories of illness: *medical* and *surgical*. *Surgical* included trauma, with *medical*

encompassing all of the other diagnoses. Subdivision of diagnoses involved a 'systems' approach, for example, infection and respiratory. Each illness was then grouped into 'minor' and 'significant'. 'Significant' was defined as a condition of sufficient potential severity that could result in acute morbidity/mortality. Where possible, national guidelines were used to classify diagnoses as minor or significant. We first reviewed the recorded diagnoses before us in the data set, and then searched for guidelines that graded severity. An example of a severity guideline was the British Thoracic Society guideline 2011 on the management of acute asthma.<sup>14</sup> Where there was not an established guideline, common agreement based on the clinical opinion between the three authors, two of whom were consultants in paediatric emergency medicine (IM and GH) and one who was a paediatric registrar (PL), was used to create an illness severity classification. See online supplementary material for the details of the diagnostic grouping and classification. PL and GH each separately assigned 50% of patients to illness severity groups. To see whether PL and GH interpreted the illness classifications similarly, that is, had good agreement, a kappa calculation was performed on a random sample of 200 patients taken from the whole data set (10.4% of the total sample) after illness allocation was completed. PL and GH were blinded as to which patients they had analysed initially from this sample. Kappa was 0.74 (95% CI 0.57 to 0.91), which is regarded as good statistical agreement. Diagnoses were assigned to minor and significant groups before PEWS was calculated from the observational data; therefore, the authors were blinded to the diagnoses relationship to PEWS.

## Methods of analysis

For both Brighton and COAST PEWS, the sensitivity, specificity and positive and negative likelihood ratios for PEWS of  $\geq 2$ ,  $\geq 3$  and  $\geq 4$  were calculated with relation to their predictive power of predicting (1) admission and (2) significant medical and surgical illness. Receiver operating curves (ROCs) and the area under the curves (AUCs) were generated for Brighton and COAST PEWS in relation to admission and illness severity. Data were analysed using Microsoft Excel 2010 and IBM SPSS Statistics V22.0.

## RESULTS

There were 2261 attendances to the paediatric ED in November 2012. Data to calculate PEWS were missing from the *Symphony* system in 565 cases. After examination of the electronic notes, missing data on 225 patients was retrieved and so these patients were included. The remaining 340 patients were excluded. The final data analysis was on 1921 patients (figure 1). Age range was 2 days to 17 years. All patients included in the analysis had their observations recorded upon arrival to the ED triage or within the ED resuscitation area (if taken straight to the resuscitation area). Baseline characteristics are summarised in table 2.

Figure 2 summarises the spread of all significant and minor illness, the breakdown of significant and minor medical and surgical illness, and further breakdown into illness subcategory. There were 211 admissions constituting 11% of attendances. Five patients transferred to another facility for urgent subspecialty care were classified as admissions. Attendances owing to all medical illness totalled 1630 (85%) and 273 (14%) had surgical conditions. The remaining 1% comprised deliberate self-harm with minor cuts, a safe-guarding case, and neonates with minor feeding difficulties. Of the total medical attendances, 268 (16.4%) were significant; and of the total surgical illnesses, 30 (11%) were significant.

**Table 1** Brighton and COAST scoring systems

	Score 1 for each abnormal parameter
Reduced consciousness	
Abnormal HR (for age)	
Abnormal RR (for age)	
Moderate-severe respiratory distress	
Supplemental oxygen (Brighton) or hypoxia sats $\leq 9$ 2% in air (COAST)	
Family/healthcare professional concern	
Total score (maximum 6)	

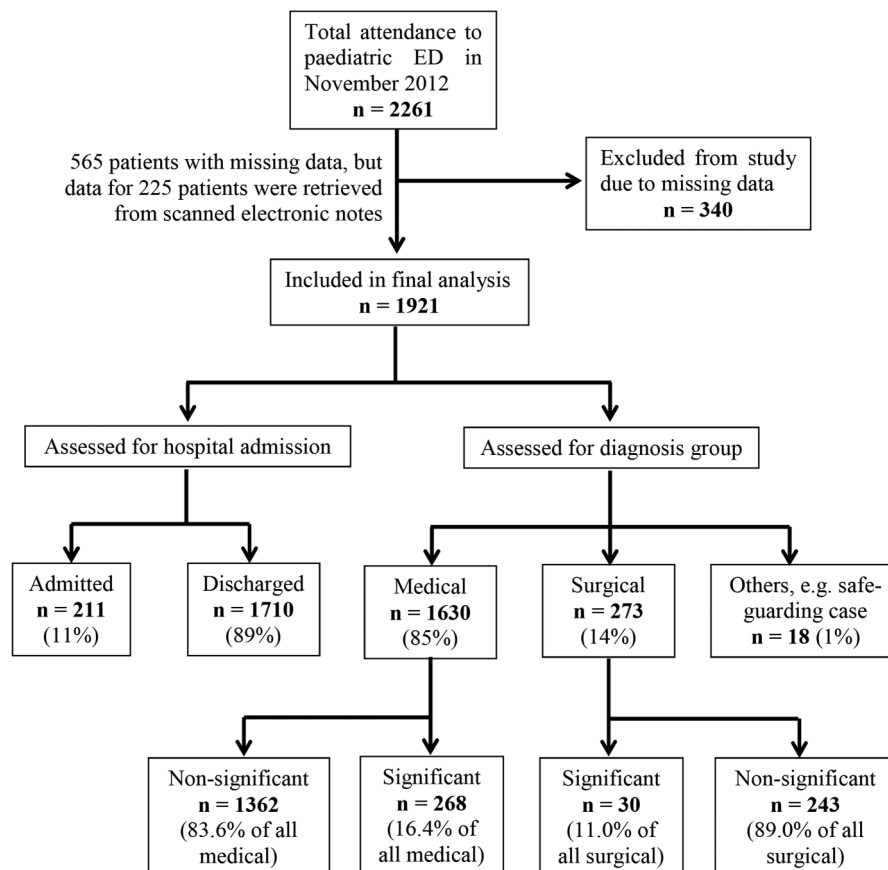
**Figure 1** Enrolment flow diagram. ED, emergency department.**AUCs for PEWS in relation to hospital admission and predicting significant illness**

Table 3 summarises the AUCs for Brighton and COAST PEWS in relation to predicting hospital admission and significant

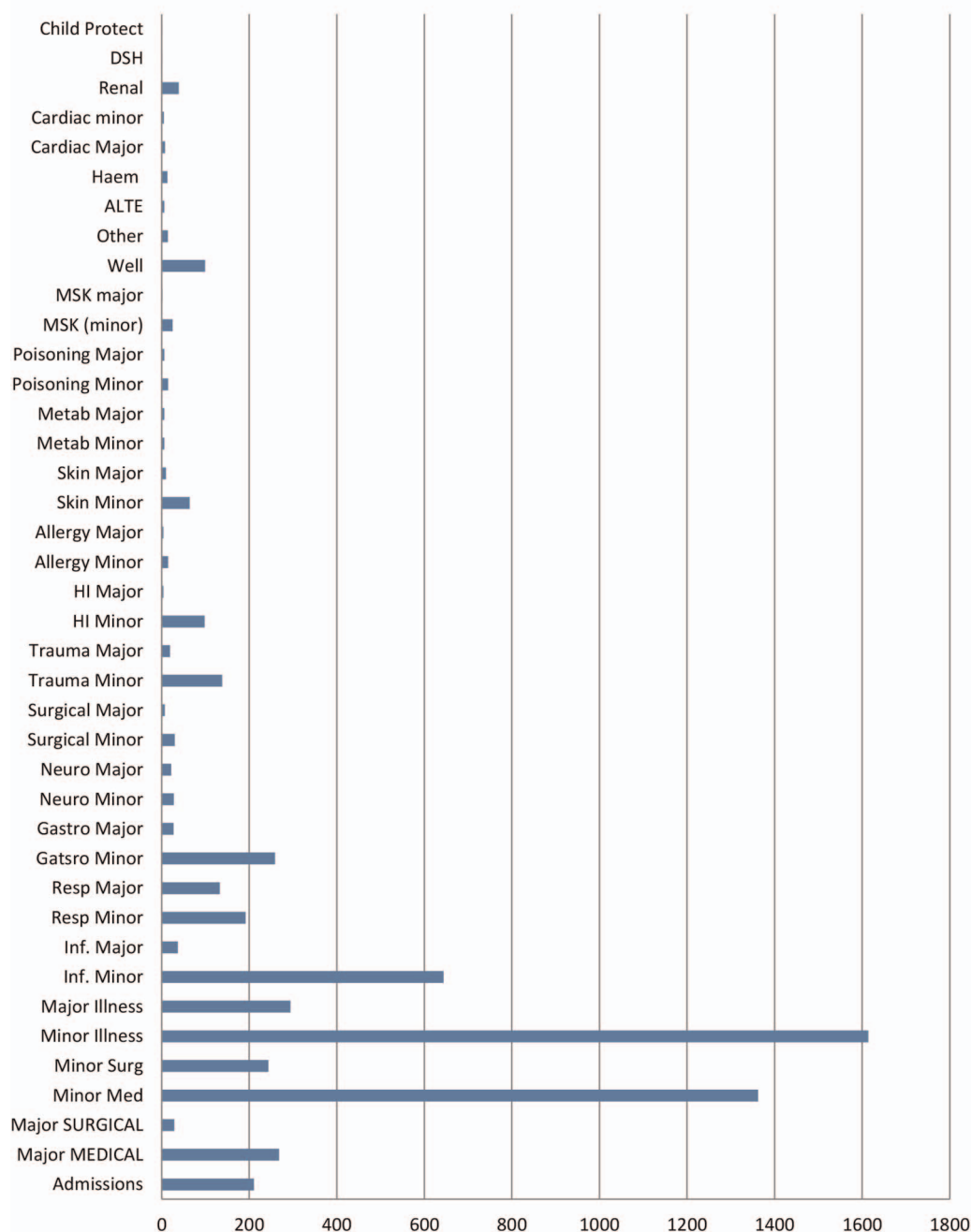
medical and surgical illness. An illness subclassification is also displayed. The AUC (0.690) was identical for both Brighton and COAST PEWS with regard to admission, and so the Brighton ROC is displayed (figure 3). For hospital admission, PEWS

**Table 2** Baseline characteristics of study population

Characteristics	Analysed patients (n=1921)	Missing data due to patients relatively well (n=298)	Missing data due to patients very ill at triage (n=8)	Missing data with no diagnosis (n=34)
Gender (female)	897 (46.7%)			
Age, median (IQR)	3.1 (6.1)	7.82 (9.3)	10.9 (11.3)	3.4 (7.4)
Admitted	211 (11.0%)	0	4 (50%)	1 (3.0%)
Diagnosis group				
Medical (significant)	268 (14.0%)	0	4 (50%)	
Medical (minor)	1362 (70.9%)	105 (35.2%)	0	
Surgical (significant)	30 (1.6%)	0	4 (50%)	
Surgical (minor)	243 (12.6%)	192 (64.4%)	0	
Others	18 (0.94%)	1 (0.34%)	0	
Brighton PEWS score				
1	1242 (64.7%)			
2	497 (25.9%)			
3	122 (6.4%)			
4	47 (2.4%)			
≥5	13 (0.7%)			
COAST PEWS score (%)				
1	1250 (65.1%)			
2	489 (25.5%)			
3	126 (6.6%)			
4	39 (2.0%)			
≥5	17 (0.9%)			

PEWS, paediatric early warning scores.

## Number of conditions (significant & minor) and hospital admissions for November 2012



**Figure 2** Number of minor and significant conditions by illness category and total number of admissions. ALTE, acute life threatening event; DSH, deliberate self harm; MSK, musculoskeletal.

would be regarded as a 'poor' test for diagnostic accuracy based on this AUC. With respect to significant medical illness, again Brighton and COAST performed very similarly with AUCs of 0.754 and 0.755, respectively. This would be regarded as a 'fair' test for diagnostic accuracy. As the ROCs were almost identical, only the Brighton ROC is displayed in [figure 4](#). Overall, both Brighton and COAST scores performed poorly as a predictive tool for detecting significant surgical illness, with an identical AUC of 0.642. The Brighton ROC is displayed in [figure 5](#). The AUCs for significant illness subcategories are displayed in [table 3](#). The only significant illness subcategory in which PEWS performed as a 'good' test was respiratory illness with the AUC

being 0.900 and 0.866 for Brighton and COAST scores, respectively. The Brighton ROC for respiratory illness is displayed in [figure 6](#).

### Sensitivity, specificity and likelihood ratios results for admission and significant illness

A summary of values is given for Brighton and COAST scores in [tables 4](#) and [5](#), respectively. Brighton and COAST performed very similarly in all measures tested. A PEWS score of  $\geq 3$  was highly specific (93%) for admission, but only 32% sensitive. For significant medical illness, a PEWS of  $\geq 3$  was 96% specific but only 44% sensitive and with relation to surgical illness PEWS

**Table 3** Area under the curves (AUCs) for Brighton and COAST paediatric early warning scores (PEWS) with respect to admission and significant illness

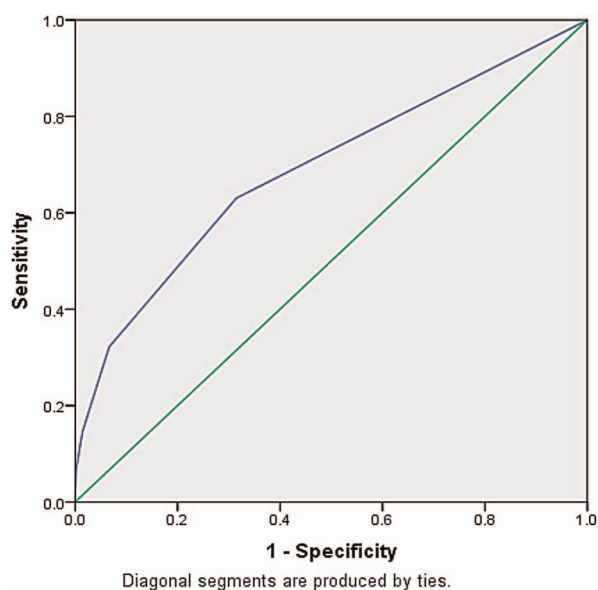
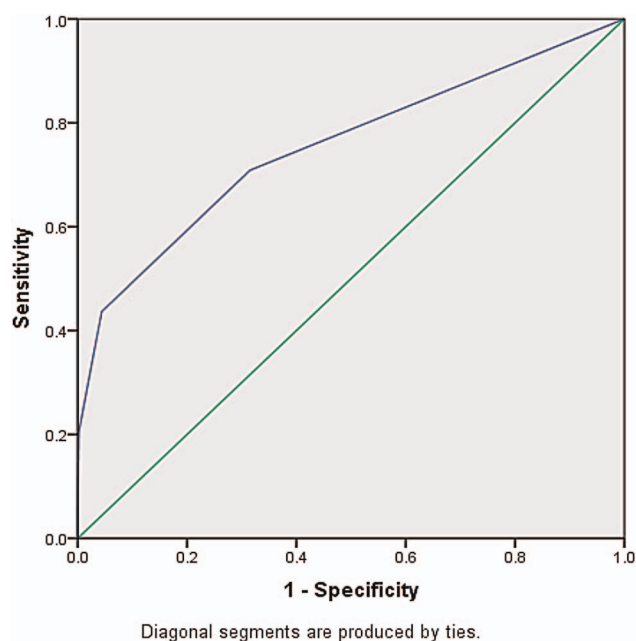
Test characteristics for diagnosing	AUC Brighton PEWS (95% CI)	AUC COAST PEWS (95% CI)
Admission (n=211)	0.690 (0.648 to 0.733)	0.690 (0.648 to 0.733)
Significant medical illness (n=268)	0.754 (0.717 to 0.791)	0.755 (0.719 to 0.792)
Significant infection (n=37)	0.605 (0.509 to 0.701)	0.605 (0.508 to 0.700)
Significant respiratory (n=133)	<b>0.900 (0.866 to 0.934)</b>	<b>0.866 (0.840 to 0.910)</b>
Significant gastrointestinal (n=27)	0.701 (0.516 to 0.815)	0.701 (0.516 to 0.815)
Significant neurological (n=22)	0.481 (0.313 to 0.648)	0.471 (0.305 to 0.637)
Significant allergy (n=4)	0.700 (0.364 to 1.000)	0.700 (0.364 to 1.000)
Significant skin (n=10)	0.544 (0.340 to 0.748)	0.544 (0.340 to 0.748)
Significant metabolic (n=6)	0.778 (0.499 to 1.000)	0.778 (0.499 to 1.000)
Significant poisoning (n=6)	0.750 (0.477 to 1.000)	0.750 (0.477 to 1.000)
Significant musculoskeletal (n=2)	0.400 (0.046 to 0.740)	0.400 (0.046 to 0.740)
Significant cardiac (n=8)	0.725 (0.439 to 1.000)	0.725 (0.439 to 1.000)
Renal/ALTE/haematology account for remaining 13 patients. receiver operating curves not done as all these diagnoses were regarded as significant, eg, glomerulonephritis, sickle crisis, acute life-threatening event		
Significant surgical illness (n=30)	0.642 (0.522 to 0.761)	0.642 (0.522 to 0.761)
Significant general surgery (n=7)	0.560 (0.311 to 0.808)	0.560 (0.311 to 0.808)
Significant trauma (n=19)	0.652 (0.503 to 0.801)	0.652 (0.503 to 0.801)
Head injury (n=4)	0.702 (0.385 to 1.000)	0.702 (0.385 to 1.000)

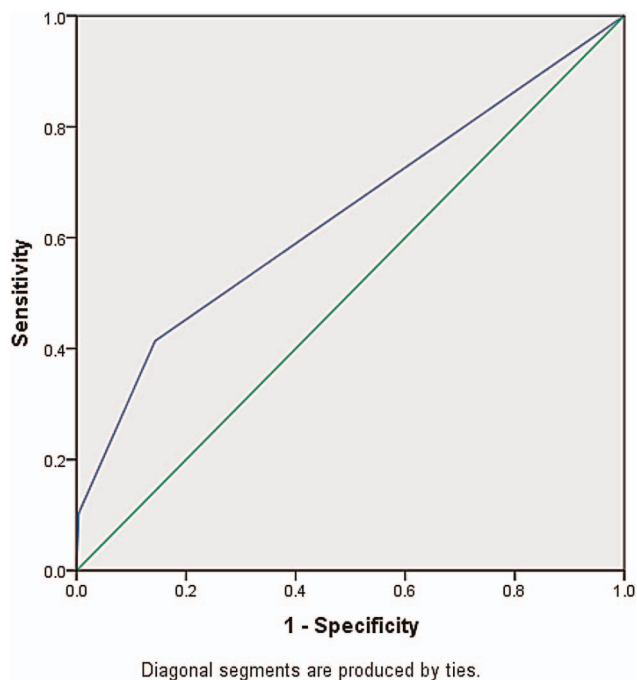
$\geq 3$  was 100% specific and 10% sensitive. No patient had a PEWS  $\geq 4$  in the COAST group among those with surgical illness. For significant respiratory illness, Brighton was 74% sensitive and 90% specific and COAST was 75% sensitive and 91% specific for a PEWS  $\geq 3$ . For all outcomes measured, positive likelihood ratios rose exponentially as the PEWS increased though negative likelihood ratios stayed close to 1.

## DISCUSSION

PEWS are specific but not sensitive in predicting hospital admission and significant illness. ROCs demonstrated PEWS was a poor tool for predicting hospital admission and significant

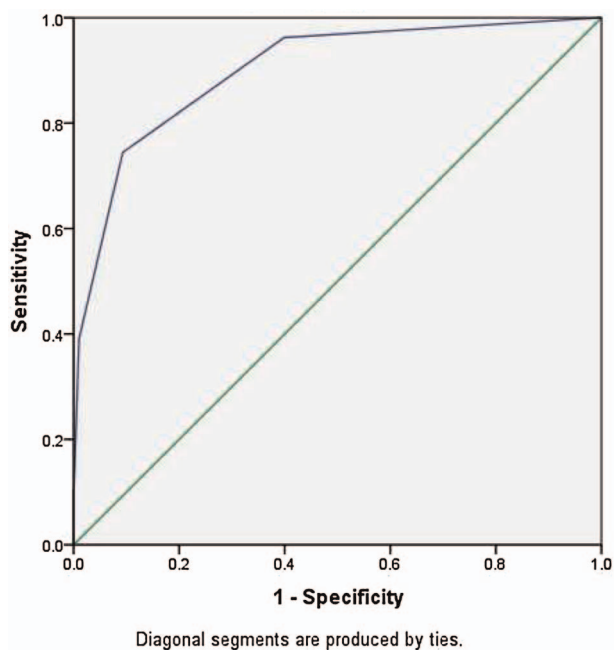
surgical illness, and fair at predicting significant medical illness. Based on these data, for all the outcomes measured (admission and significant medical and surgical illness) a PEWS threshold of  $\geq 3$  is the point closest to the top-left corner of all the plotted ROC, hence representing the optimal trade-off between sensitivity and specificity, with a specificity of  $>90\%$  for admission and significant medical and surgical illness, and a sensitivity 32% for admission, 44% for medical illness and 10% for surgical illness. A PEWS threshold of  $<3$  has poor specificity while a threshold of  $\geq 4$  has little increase in specificity with a decline in sensitivity owing to only a fewer number of patients with such high PEWS scores.

**Figure 3** Brighton receiver operating curve for admission.**Figure 4** Brighton receiver operating curve for significant medical illness.



**Figure 5** Brighton receiver operating curve for significant surgical illness.

Previous work has examined illness severity in already hospitalised children in relation to PEWS, and a recent review of 10 studies encompassing a range of PEWS scoring systems included a total of 17 943 children.<sup>1 11–13 15–21</sup> Area under the ROCs for predicting hospitalisation was poor to moderate (range 0.56–0.68) with sensitivity and specificity being 36.4–85.7% and 27.1–90.5%, respectively. Only two of the studies included ED patients: 46 patients admitted from the ED to paediatric intensive care department (PICU),<sup>17</sup> and another study that examined the incidence of cardiorespiratory arrests before and after the



**Figure 6** Brighton receiver operating curve for significant respiratory illness.

introduction of a medical emergency team and thus did not examine PEWS' direct relationship to general hospital admission.<sup>18</sup> Scoring systems are used frequently within the paediatric ED with regard to triage such as the Manchester Scoring System.<sup>22 23</sup> The Pediatric Risk of Admission Score exists to predict the risk of hospitalisation in children and has a sensitivity >80%,<sup>24–26</sup> but is not widely used in the UK.

Specifically focusing on Brighton PEWS, Bradman in 2008 examined 424 children with any medical problem attending a paediatric ED and found the Brighton PEWS sensitivity and specificity for hospital admission to be 24% and 96%, respectively, for a score  $\geq 4$ .<sup>6</sup> In 2012, Bradman and colleagues<sup>8</sup> prospectively studied 1223 children attending the ED. In addition to assessing Brighton PEWS validity in predicting hospital admission, they also examined triage nurse (TN) opinion on whether the child needed admission. TN opinion most accurately predicted hospital admission having a prediction accuracy of 87.7%, followed by an elevated PEWS with a prediction accuracy of 82.9%. PEWS  $\geq 4$  had poor sensitivity (14%) but good specificity (98%). This study included a relatively even split of medical, surgical and injured patients, but subgroup analyses of these groups and diagnostic severity were not examined. The recently completed Paediatric Observation Priority Score (POPS) study evaluated a novel aggregate scoring system designed for use in the ED.<sup>27</sup> Examining 2068 patients under 16 years, the AUC for predicting hospital admission was 0.73 for medical patients and 0.69 for trauma patients, with sensitivity and specificity for a POPS  $\geq 3$  being 36% and 93%, respectively. With regard to predicting admission, our study showed similar results to the above studies, that is, good specificity but poor sensitivity. Therefore, as a screening tool for admission, the test performs increasingly poorly as the measure of illness severity increases. The worsening of sensitivity with a rising PEWS may be owing to diminishing population size in ascending PEWS categories. Unlike Egdell and Chaikyulsil,<sup>10 17</sup> our study did not examine the association with PEWS and admission to PICU from the ED because the number of PICU admissions from a single centre in a 1-month period would be too small to yield informative data.

Similar to hospital admission, we demonstrated that PEWS was specific for predicting significant illness, and for a PEWS of  $\geq 3$  suspicion of underlying significant illness should be taken seriously. Illness subgroup analysis demonstrated that PEWS performed well as a tool for detecting significant respiratory illness, with ROCs for Brighton and COAST scores yielding AUCs of 0.900 and 0.866, respectively. The robustness of this result maybe owing to this illness group having the largest population size. The timing of the year (winter) will account for the high proportion of respiratory illnesses. Interestingly, this AUC was similar to the finding by Breslin *et al*,<sup>28</sup> who found that Brighton PEWS had an AUC of 0.80 for those who were admitted to hospital with respiratory illnesses, but the AUC was only 0.63 (95% CI 0.57 to 0.69) for those admitted with all other illnesses. The Breslin study calculated an ROC specifically for respiratory illness. A collective ROC was done for all other illnesses combined, and a classification of illness severity was not done. Nonetheless, it is interesting that PEWS seems to have a relationship with respiratory illness and perhaps future work should examine tailoring early warning scores to individual illnesses.

Our study has shown that two commonly used PEWS (Brighton and COAST) perform similarly. If this finding is replicated in future studies, then perhaps there is advantage in using a single type of PEWS among the paediatric patients in ED.

**Table 4** Brighton paediatric early warning scores (PEWS) characteristics

Test characteristics for diagnosing	PEWS $\geq 2$	PEWS $\geq 3$	PEWS $\geq 4$
<i>Hospital admission</i>			
Total admitted=211	N admitted=131	N admitted=67	N admitted=35
Total discharged=1710	N discharged=548	N discharged=115	N discharged=25
Sensitivity (95% CI)	62% (55.1% to 68.6%)	32% (25.6% to 38.6%)	17% (12.0% to 22.5%)
Specificity (95% CI)	68% (65.7% to 70.2%)	93% (92.0% to 94.4%)	99% (97.8% to 99.0%)
Positive likelihood ratio (95% CI)	1.94 (1.71 to 2.20)	4.72 (3.62 to 6.16)	11.3 (6.93 to 18.6)
Negative likelihood ratio (95% CI)	0.56 (0.47 to 0.66)	0.73 (0.67 to 0.80)	0.85 (0.80 to 0.90)
<i>Significant medical illness</i>			
Total significant=268	N significant=190	N significant=117	N significant=55
Total minor=1362	N minor=429	N minor=59	N minor=3
Sensitivity (95% CI)	71% (65.0% to 76.2%)	44% (37.7% to 49.8%)	21% (16.0% to 26.0%)
Specificity (95% CI)	69% (65.9% to 71.0%)	96% (94.4% to 96.7%)	100% (99.3% to 100%)
Positive likelihood ratio (95% CI)	2.25 (2.02 to 2.51)	10.1 (7.58 to 13.4)	93.2 (29.4 to 296)
Negative likelihood ratio (95% CI)	0.42 (0.35 to 0.51)	0.59 (0.53 to 0.65)	0.80 (0.75 to 0.85)
<i>Significant surgical illness</i>			
Total significant=30	N significant=12	N significant=3	N significant=1
Total minor=243	N minor=35	N minor=1	N minor=0
Sensitivity (95% CI)	40% (22.7% to 59.4%)	10% (2.11% to 26.5%)	3% (0.1% to 17.2%)
Specificity (95% CI)	86% (80.5% to 89.5%)	100% (97.7% to 100%)	100% (98.0% to 100%)
Positive likelihood ratio (95% CI)	2.78 (1.63 to 4.74)	24.3 (2.61 to 226)	Incalculable
Negative likelihood ratio (95% CI)	0.70 (0.52 to 0.94)	0.90 (0.80 to 1.02)	0.97 (0.90 to 1.03)
<i>Significant respiratory illness</i>			
Total significant=133	N significant=128	N significant=99	N significant=52
Total minor=192	N minor=76	N minor=18	N minor=2
Sensitivity (95% CI)	96 (91.4% to 98.8%)	74 (66.15% to 81.6%)	39% (30.8% to 47.9%)
Specificity (95% CI)	60% (53.1% to 67.4%)	90% (85.6% to 94.4%)	99% (96.3% to 99.9%)
Positive likelihood ratio (95% CI)	2.43 (2.03 to 2.90)	7.94 (5.06 to 12.46)	38 (9.3 to 151.4)
Negative likelihood ratio (95% CI)	0.06 (0.03 to 0.15)	0.28 (0.21 to 0.38)	0.62 (0.54 to 0.71)

This will help reduce confusion of having different scoring systems for different hospitals.

### Limitations

Study limitations include this work being a single-centre inner-city retrospective analysis of patient data in a paediatric ED over a 1-month period during the winter where respiratory illnesses predominate. We calculated PEWS based on the data recorded at the first assessment in the ED. Therefore, the potential for evolving physiological data over time was not accounted for. Patients with missing data tended to be either relatively well with minor injuries or were judged to be very ill at triage and so taken straight to the resuscitation area. Therefore, a bias was created by excluding a number of very well or very sick patients from our study. Nonetheless, complete data were retrieved from 225 out of the 565 with missing data fields. Owing to the difficulty in retrospectively delineating parental/carer concern, we automatically assigned a PEWS of 1 to every patient on the basis that parents/carers, by bringing their child to the ED, were by default 'concerned' about their child. This scoring approach, however, may not be universal in clinical practice. Owing to the heterogeneity in diagnoses presenting to the ED, a classification system was devised. To classify diagnoses as *minor* or *significant*, we used existing guidelines where possible to minimise subjectivity bias. Though disease management guidelines exist for many of the other conditions, illness severity classification into significant and non-significant is not a component of the guidance as the majority of these conditions such as meningococcal sepsis, meningitis, sickle crisis and status epilepticus are by

definition significant illnesses. Therefore, we classified such diseases as significant. Where grey areas existed regarding what was significant in the absence of defined guidelines, explanation is given in the online supplementary table next to the diagnosis about our definition. After completing the allocations of illness severity for the whole sample, we then randomly selected 200 patients for PL and GH to perform an analysis of investigator agreement. Though the statistical agreement was good (kappa 0.74), we acknowledge that this is still imperfect agreement and we did not extend out agreement analysis outside of the 200 random sample. Imperfect agreement could affect sensitivity/specificity especially in the surgical group where the number of patients was small. We also did not perform age subgroup analysis. All patients with missing observational data were excluded; therefore, we did not compare the ROC of PEWS calculated from complete and missing observational data.

### CONCLUSION

In keeping with previous work, this study has found that a PEWS score of  $\geq 3$  is specific but not sensitive in predicting hospital admission and for the first time has found this PEWS score be specific but not sensitive in predicting significant illness in the children's ED. There was no difference in the performance of Brighton and COAST scores in all parameters studied.

The implication from this study is that a high PEWS (we propose  $\geq 3$ ) has few false positives and must prompt thought for hospital admission and the investigation of significant illness, but a low PEWS should not be taken to exclude significant illness or the need for admission. Those using Brighton and COAST

**Table 5** COAST paediatric early warning scores (PEWS) characteristics

Test characteristics for diagnosing	PEWS $\geq 2$	PEWS $\geq 3$	PEWS $\geq 4$
<i>Hospital admission</i>			
Total admitted=211	N admitted=132	N admitted=67	N admitted=35
Total discharged=1710	N discharged=540	N discharged=116	N discharged=25
Sensitivity (95% CI)	63% (55.7% to 69.2%)	32% (25.5% to 38.5%)	14% (10.0% to 20.0%)
Specificity (95% CI)	68% (65.7% to 70.2%)	93% (91.8% to 94.3%)	99% (97.8% to 99.0%)
Positive likelihood ratio (95% CI)	1.99 (1.75 to 2.25)	4.64 (3.56 to 6.06)	9.42 (5.68 to 15.61)
Negative likelihood ratio (95% CI)	0.55 (0.46 to 0.65)	0.73 (0.67 to 0.80)	0.87 (0.82 to 0.92)
<i>Significant medical illness</i>			
Total significant=268	N significant=191	N significant=117	N significant=52
Total minor=1362	N minor=430	N minor=61	N minor=3
Sensitivity (95% CI)	71% (65.5% to 76.6%)	44% (37.6% to 49.8%)	19% (14.8% to 24.7%)
Specificity (95% CI)	69% (65.9% to 71.0%)	96% (94.3% to 96.6%)	100% (99.4% to 100%)
Positive likelihood ratio (95% CI)	2.26 (2.02 to 2.52)	9.75 (7.36 to 12.9)	88.1 (27.7 to 280)
Negative likelihood ratio (95% CI)	0.42 (0.35 to 0.51)	0.59 (0.53 to 0.66)	0.81 (0.76 to 0.86)
<i>Significant surgical illness</i>			
Total significant=30	N significant=12	N significant=3	N significant=0
Total minor=243	N minor=35	N minor=1	N minor=0
Sensitivity (95% CI)	40% (22.7% to 59.4%)	10% (2.11% to 26.53%)	0% (0.0% to 11.6%)
Specificity (95% CI)	86% (80.5% to 89.5%)	100% (97.7% to 100%)	100% (98.5% to 100%)
Positive likelihood ratio (95% CI)	2.78 (1.63 to 4.74)	24.3 (2.61 to 226)	Incalculable
Negative likelihood ratio (95% CI)	0.70 (0.52 to 0.94)	0.90 (0.80 to 1.02)	1.00 (1.00 to 1.00)
<i>Significant respiratory illness</i>			
Total significant=133	N significant=129	N significant=100	N significant=50
Total minor=192	N minor=76	N minor=18	N minor=2
Sensitivity (95% CI)	97% (92.5% to 99%)	75% (67.0% to 82.3%)	38% (29.4% to 46.4%)
Specificity (95% CI)	60% (53.1% to 67.4%)	91% (85.6% to 94.3%)	99% (96.3% to 99.9%)
Positive likelihood ratio (95% CI)	2.45 (2.05 to 2.93)	8.02 (5.11 to 12.58)	36.1 (8.94 to 145.8)
Negative likelihood ratio (95% CI)	0.05 (0.02 to 0.13)	0.27 (0.20 to 0.37)	0.63 (0.55 to 0.72)

PEWS within the ED should also note that they were not designed for use within the ED and instead were validated to detect early deterioration in patients who are hospitalised. Being that 15% of our initial sample had irretrievable missing observational data and were subsequently excluded (leading to exclusion bias), future work should include prospective multicentre studies and these should take place over a year to reduce seasonal variability in pathology and increase the heterogeneity of the study population. Age subgroup analysis should also be performed. A comparison of PEWS against TN opinion and established ED triage systems would be useful as suggested by Seiger *et al*,<sup>21</sup> as should comparison with the new ED-specific POPS scoring system that has recently been demonstrated to be a better predictor for hospital admission compared with the Manchester children's early warning system.<sup>29</sup> Correlation with patient disposition and length of stay in the hospital would also be useful to incorporate into future analysis. Finally, we propose that future work should examine tailoring early warning scores to individual illnesses based on our findings with respiratory illness.

**Acknowledgements** The authors thank Dr Jill Warner and Karen Harrison White for their help in the review of the manuscript.

**Funding** NIHR Imperial BRC funding, P46153.

**Competing interests** IM is in receipt of a grant from National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London.

**Ethics approval** The local research compliance body classified this study as audit since we were assessing established illness scoring systems within their remit, that is, to score illness severity, using anonymised retrospective data. Therefore, ethical approval was not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Unpublished raw data on PEWS in relation to age subgroups are available and in possession of the main author PJL.

## REFERENCES

- 1 Monaghan A. Detecting and managing deterioration in children. *Paediatr Nurs* 2005;17:32–5.
- 2 Pearson G. *Why children die: a pilot Study 2006, England, Wales and Northern Ireland*. London: CEMACH, 2008:1.
- 3 Roland D. Paediatric early warning scores: Holy Grail and Achilles' heel. *Arch Dis Child Educ Pract Ed* 2012;97:208–15.
- 4 Chapman SM, Grocott MP, Franck LS. Systematic review of paediatric alert criteria for identifying hospitalised children at risk of critical deterioration. *Intensive Care Med* 2010;36:600–11.
- 5 Winberg H, Nilsson K, Aneman A. Paediatric Rapid Response Systems: a literature review. *Acta Anaesthesiol Scand* 2008;52:890–6.
- 6 Bradman K, Maconochie I. Can paediatric early warning score be used as a triage tool in paediatric accident and emergency? *Eur J Emerg Med* 2008;15: 359–60.
- 7 Roland D, Oliver A, Edwards ED, *et al*. Use of paediatric early warning systems in Great Britain: has there been a change of practice in the last 7 years? *Arch Dis Child* 2014;99:26–9.
- 8 Bradman K, Borland M, Pascoe E. Predicting patient disposition in a paediatric emergency department. *J Paediatr Child Health* 2014;50:E39–44.
- 9 [http://www.institute.nhs.uk/safer\\_care/paediatric\\_safer\\_care/pews.htm](http://www.institute.nhs.uk/safer_care/paediatric_safer_care/pews.htm) (accessed 2 May 2015).
- 10 Chaiyakulsi C, Pandee U. Validation of pediatric early warning score in pediatric emergency department. *Pediatr Int* 2015;57:694–8.
- 11 Duncan H, Hutchison J, Parshuram CS. The Pediatric Early Warning System score: a severity of illness score to predict urgent medical need in hospitalized children. *J Crit Care* 2006;21:271–8.
- 12 Edwards ED, Powell CV, Mason BW, *et al*. Prospective cohort study to test the predictability of the Cardiff and Vale paediatric early warning system. *Arch Dis Child* 2009;94:602–6.



- 13 Haines C, Perrott M, Weir P. Promoting care for acutely ill children-development and evaluation of a paediatric early warning tool. *Intensive Crit Care Nurs* 2006;22:73–81.
- 14 British Thoracic Society. *British guideline on the management of Asthma*. 2011. <https://www.brit-thoracic.org.uk/document-library/clinical-information/asthma/btssign-asthma-guideline-2011/>
- 15 Akre M, Finkelstein M, Erickson M, *et al*. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics* 2010;125:e763–9.
- 16 Skaletzky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. *Clin Pediatr (Phila)* 2012;51:431–5.
- 17 Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J* 2008;25:745–9.
- 18 Brilll RJ, Gibson R, Luria JW, *et al*. Implementation of a medical emergency team in a large pediatric teaching hospital prevents respiratory and cardiopulmonary arrests outside the intensive care unit. *Pediatr Crit Care Med* 2007;8:236–46; quiz 47.
- 19 Parshuram CS, Bayliss A, Reimer J, *et al*. Implementing the Bedside Paediatric Early Warning System in a community hospital: A prospective observational study. *Paediatr Child health* 2011;16:e18–22.
- 20 Tibballs J, Kinney S, Duke T, *et al*. Reduction of paediatric in-patient cardiac arrest and death with a medical emergency team: preliminary results. *Arch Dis Child* 2005;90:1148–52.
- 21 Seiger N, Maconochie I, Oostenbrink R, *et al*. Validity of different pediatric early warning scores in the emergency department. *Pediatrics* 2013;132:e841–50.
- 22 Mackway-Jones K. *Emergency triage: Manchester Triage Group*. London: BMJ Publishing Group, 1997.
- 23 Mackway-Jones K, Marsden J, Windle J. *Emergency Triage*. London: BMJ (Clinical research ed.), 2006(xiii):178.
- 24 Chamberlain JM, Patel KM, Ruttimann UE, *et al*. Pediatric risk of admission (PRISA): a measure of severity of illness for assessing the risk of hospitalization from the emergency department. *Ann Emerg Med* 1998;32:161–9.
- 25 Chamberlain JM, Patel KM, Pollack MM. The Pediatric Risk of Hospital Admission score: a second-generation severity-of-illness score for pediatric emergency patients. *Pediatrics* 2005;115:388–95.
- 26 Gravel J, Gouin S, Amre D, *et al*. Evaluation of the pediatric risk of admission score in a pediatric emergency department. *Ann Emerg Med* 2003;41:630–8.
- 27 Kelly J, Rowland AG, Cotterill S, *et al*. The Paediatric observation priority score (POPS): a useful tool to predict likelihood of admission from the emergency department. *Emerg Med J* 2013;30:877–8.
- 28 Breslin K, Marx J, Hoffman H, *et al*. Pediatric early warning score at time of emergency department disposition is associated with level of care. *Pediatr Emerg Care* 2014;30:97–103.
- 29 Rowland A, Cotterill S, Lees H, *et al*. The Paediatric Observation Priority Score (POPS): a more accurate predictor of admission risk from the Emergency Department than the Manchester Children's Early Warning System (ManChEWS). *Arch Dis Child* 2014;99(Suppl 1):A8.



# Can paediatric early warning scores (PEWS) be used to guide the need for hospital admission and predict significant illness in children presenting to the emergency department? An assessment of PEWS diagnostic accuracy using sensitivity and specificity

Peter J Lillitos, Graeme Hadley and Ian Maconochie

*Emerg Med J* 2016 33: 329-337 originally published online November 3, 2015

doi: 10.1136/emered-2014-204355

---

Updated information and services can be found at:  
<http://emj.bmj.com/content/33/5/329>

*These include:*

## Supplementary Material

Supplementary material can be found at:  
<http://emj.bmj.com/content/suppl/2015/11/03/emered-2014-204355.DC1.html>

## References

This article cites 24 articles, 11 of which you can access for free at:  
<http://emj.bmj.com/content/33/5/329#BIBL>

## Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

## Topic Collections

Articles on similar topics can be found in the following collections

[TB and other respiratory infections](#) (158)

---

## Notes

---

To request permissions go to:  
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:  
<http://group.bmj.com/subscribe/>