BMJ Open

Performance of scoring systems in selecting short stay medical admissions suitable for assessment in same day emergency care: an analysis of diagnostic accuracy in a UK hospital setting

Catherine Atkin ,[1] Suzy Gallier,[2] Elizabeth Wallin,[3] Vinay Reddy-Kolanu,[3] Elizabeth Sapey [1,3]

To cite: Atkin C, Gallier S, Wallin E, *et al*. Performance of scoring systems in selecting short stay medical admissions suitable for assessment in same day emergency care: an analysis of diagnostic accuracy in a UK hospital setting. *BMJ Open* 2022;**12**:e064910. doi:10.1136/bmjopen-2022-064910

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-064910).

Received 17 May 2022
Accepted 02 December 2022

For numbered affiliations see end of article.

Correspondence to
Dr Catherine Atkin;
c.atkin@nhs.net

ABSTRACT

Objectives To assess the performance of the Amb score and Glasgow Admission Prediction Score (GAPS) in identifying acute medical admissions suitable for same day emergency care (SDEC) in a large urban secondary centre.

Design Retrospective assessment of routinely collected data from electronic healthcare records.

Setting Single large urban tertiary care centre.

Participants All unplanned admissions to general medicine on Monday–Friday, episodes starting 08:00–16:59 hours and lasting up to 48 hours, between 1 April 2019 and 9 March 2020.

Main outcome measures Sensitivity, specificity, positive and negative predictive value of the Amb score and GAPS in identifying patients discharged within 12 hours of arrival.

Results 7365 episodes were assessed. 94.6% of episodes had an Amb score suggesting suitability for SDEC. The positive predictive value of the Amb score in identifying those discharged within 12 hours was 54.5% (95% CI 53.3% to 55.8%). The area under the receiver operating characteristic curve (AUROC) for the Amb score was 0.612 (95% CI 0.599 to 0.625).

42.4% of episodes had a GAPS suggesting suitability for SDEC. The positive predictive value of the GAPS in identifying those discharged within 12 hours was 50.5% (95% CI 48.4% to 52.7%). The AUROC for the GAPS was 0.606 (95% CI 0.590 to 0.622).

41.4% of the population had both an Amb and GAPS score suggestive of suitability for SDEC and 5.7% of the population had both and Amb and GAPS score suggestive of a lack of suitability for SDEC.

Conclusions The Amb score and GAPS had poor discriminatory ability to identify acute medical admissions suitable for discharge within 12 hours, limiting their utility in selecting patients for assessment within SDEC services within this diverse patient population.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study compared performance of the Amb score and Glasgow Admission Prediction Score in identifying patients likely to be discharged within 12 hours of admission using real-world outcome data.
⇒ Scores were calculated based on routinely collected electronic healthcare data, reflecting potential use in clinical practice, however this meant some data fields had higher rates of missing data.
⇒ Analysis of score performance incorporated National Early Warning Score 2, reflecting current clinical practice.
⇒ Patients admitted for >48 hours were not included, therefore score performance may be an overestimate if applied to all medical admissions.

INTRODUCTION

The increase in emergency medical admissions to hospital places a significant demand on acute care and inpatient services within secondary care.[1] Same day emergency care (SDEC) has been proposed as a care model to reduce hospital admission. Here, patients admitted with a medical emergency are reviewed within working hours with investigations and treatments instigated, with the facility for patients to return for further investigations on subsequent days as needed, without admission to a hospital bed. In the UK, SDEC has been highlighted as a priority within the National Health Service (NHS),[2] including the NHS Long Term Plan, which provides a suggested target that a third of medical patients be managed without overnight admission.[3] Currently, it is unclear how best to structure SDEC services to deliver care most effectively to those that may benefit.[4] A key criterion is the correct selection of patients for SDEC as soon as possible following presentation, with those

Table 1 Scoring systems to identify medical admissions potentially suitable for discharge from hospital without admission >12 hours

Amb score			GAPS		
Sex	Female	0	NEWS		1 point per point on NEWS score
	Male	−0.5			
Age	<80	0	Age		1 point per decade
	≥80	−0.5			
Access to personal transport/can take public transport	Agree	2	Triage category	3 2 (or 2+) 1	5 10 20
	Disagree	0			
Intravenous treatment not anticipated	Agree	2	Referred by GP		5
	Disagree	0			
Not acutely confused	Agree	2	Arrived in ambulance		5
	Disagree	0			
MEWS=0	Agree	1	Admitted <1 year ago		5
	Disagree	0			
Not discharged from hospital within previous 30 days	Agree	1			
	Disagree	0			

Amb score[5] and GAPS.[7] Amb score of 5 or more indicates likely discharge within 12 hours; GAPS of 16 or more suggests patient likely to be admitted to hospital.
GAPS, Glasgow Admissions Prediction Score; GP, general practitioner; MEWS, Modified Early Warning Score; NEWS, National Early Warning Score.

likely to be discharged within 12 hours directed through SDEC services, and those requiring admission (lasting >12 hours) assessed within acute medical units (AMUs).

Two scoring systems have been proposed for UK health services, the Amb score (Ambs) and Glasgow Admission Prediction Score (GAPS) (see table 1). The Ambs[5] has been recommended by the Royal College of Physicians (RCP),[6] with a score of 5 points or more indicating a patient will likely be discharged from hospital within 12 hours. The Ambs was derived in a rural patient cohort, with the validatory study using retrospective data testing the score's ability to discriminate between patients with admissions of <12 hours or >48 hours. That study excluded patients who remained in hospital for 12–48 hours.

GAPS has also been suggested as a scoring system to identify patients who are likely to require admission to hospital.[7] The score was derived in Scotland and was designed to predict a dichotomous outcome of discharge from hospital versus admission. This score is used in some centres to aid selection of patients for SDEC services. A predefined cut-off score identifying those likely to be admitted to hospital is not provided, as it is recommended that this be adjusted to local patient populations, however a score of 16 or more predicted admission to hospital in the original study.

To enable effective flow through hospitals, patients suitable for SDEC should be selected early and accurately, so SDEC areas are not filled with patients who later need admission, and AMU beds are not filled by patients who are quickly discharged home.

This retrospective health data study was conducted to determine the performance of the Ambs and GAPS for selecting SDEC patients in a diverse urban centre in the UK, assessing in particular the scores' ability to discriminate between acute medical admissions suitable for SDEC and those requiring admission for at least 12–48 hours.

METHODS

This data study was conducted in collaboration with PIONEER, a Health Data Research Hub in Acute Care.

Retrospective data were collected for patients admitted to Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Trust (UHB) between the period of 1 April 2019 and 9 March 2020.

UHB is one of the largest Trusts nationally, covering 4 NHS hospital sites, treating over 2.2 million patients per year and housing the largest single critical care unit in Europe. The AMU contains 68 inpatient beds, with a physically distinct SDEC area consisting of 5 cubicles for assessment and 15 chairs.

UHB is a paperless hospital with all health data and noting captured within UHB's in-house electronic health record (EHR) called Prescribing Information and Communication System. Admission episodes starting in the emergency department are also recorded within Oceano (CSE Healthcare).

All patients aged ≥16 years with an emergency admission under acute or general medicine services lasting up to 48 hours were included. Longer admissions were not

included, as this analysis focused on patients likely to be managed within acute medicine services, without admission to specialty medicine inpatient wards.

Length of stay was measured from initial arrival time to hospital, including any period of care under emergency medicine. All admission episodes within the censor period were included with the end date chosen to align with detection of the first confirmed SARS-CoV-2 case in UHB, to minimise the impact on the analysis of changes in patient admission patterns and patient pathways during the COVID-19 pandemic. During this time period, the acute medicine service delivered same day emergency care through a dedicated ambulatory area, without use of a standardised scoring system.

Patient and public involvement

This project was discussed with a patient and public advisory group who highlighted the importance of minimising wait times in acute services, and of options for treatment that avoid hospital admission. This group co-agreed the data fields included in this analysis and have helped write a lay summary about the project.

Data included patient demographics (age, sex and self-assigned ethnicity), time stamps related to arrival to and discharge from hospital, method of arrival to hospital, referral source, patient location within hospital and comorbidities. The first recorded set of observations after arrival was included, with early warning scores calculated from this set of observations. Previous attendance to UHB within 30 days and 12 months of each episode was included. Primary diagnosis for the admission and comorbidities were assessed from recorded SNOMED (Systematized Nomenclature of Medicine Clinical Terms) and mapped International Classification of Diseases-10 codes. For episodes initiated in the emergency department, the initial triage problem, as recorded into the EHR on patient arrival to hospital, and the coded primary diagnosis at exit from the emergency department, representing the suspected diagnosis at this point, were included. Triage category was available for admissions starting in the emergency department.

Length of admission was grouped into 12-hour intervals; for evaluation of scoring systems, admissions lasting 12–48 hours were grouped. Additional outcomes assessed were death within 30 days of admission, and reattendance within 7 and 30 days.

Analysis of score performance was restricted to episodes beginning between 08:00 and 16:59 hours, Monday to Friday ('normal working day' (NWD)), to reflect common opening hours of SDEC services and highest access to diagnostic investigations and specialist pathways that would facilitate SDEC.

The Amb score[5] and GAPS[7] were calculated for each episode, using the score as outlined in the original derivation studies (table 1). For the Amb score, a Modified Early Warning Score (MEWS) was calculated[5]; when calculating the score, all patients received 2 points for access to transport as UHB provides transport to any patient if required. Intravenous treatment was taken as not being anticipated where patients did not receive an intravenous therapy within 6 hours of arrival. A score of 5 or more was used to indicate suitability for SDEC and likely discharge within 12 hours, as per the original study. For the GAPS, a National Early Warning Score (NEWS) was calculated.[8] A GAPS of 16 or more, used as a binary cut-off in the original study, was used to indicate likelihood of admission, making a patient unsuitable for SDEC. For both scores, patients were only included where all components could be assessed from the EHR data.

The NEWS2 is currently used in clinical practice and recommended by the RCP.[9] The first NEWS2 on arrival was calculated; this was substituted into the Amb score (replacing MEWS) and GAPS (replacing NEWS) to reflect how these scores would perform in clinical practice using NEWS2. Comparison of score performance with the original early warning score and NEWS2 is shown.

Statistical analysis was performed using Stata/SE V.15.1. Cell counts containing fewer than 10 patients were suppressed, due to reporting requirements. For univariate analysis of factors influencing likelihood of discharge within 12 hours, ORs for variables included in the original Amb score or GAPS derivation studies were assessed using a mixed-effects logistic regression, with patient included as a random effect, as patients could appear in the dataset more than once. Multivariable analysis of the Amb score and GAPS components was also performed using mixed-effects logistic regression, with patient as a random effect, to demonstrate the performance of components within the score and allow an evaluation of whether score components were associated with length of stay in this cohort. Receiver operator characteristic (ROC) curves were calculated for each scoring system, and the area under the receiver operating characteristic curve (AUROC) calculated. Subgroup analysis was performed in prespecified groups based on previous research.[10] Comparison of proportions between those correctly identified by the GAPS or Amb score was performed using χ^2. A p value of <0.05 was used to signify statistical significance throughout. Rates of reattendance were assessed at 7 days and at 30 days, with a sensitivity analysis of readmissions for episodes not associated with another episode in the preceding 30 days.

To evaluate likely impact on patient pathway, an average of 100 total admission per day to acute medical services was assumed, reflecting admission numbers through UHB acute medical services, with 50% of patients remaining in hospital <48 hours, based on previous research.[10]

RESULTS

A total of 14 314 acute medical inpatient episodes lasting up to 48 hours were identified during the censor period. These episodes were from 12 587 patients with 11 229 patients having one episode in this time period. Patients

Table 2 Demographics and characteristics of patients with emergency medical admissions lasting up to 48 hours

	All episodes N=14314		Normal working day episodes N=7365		Episodes starting outside normal working day N=6949		
	Frequency (%)		Frequency (%)		Frequency (%)		P value
Age (years)							
16–19	444	(3.1%)	172	(2.3%)	272	(3.9%)	<0.001
20–29	1585	(11%)	724	(10%)	861	(12%)	
30–39	1677	(12%)	826	(11%)	851	(12%)	
40–49	1776	(12%)	909	(12%)	867	(13%)	
50–59	2308	(16%)	1255	(17%)	1053	(15%)	
60–69	2000	(14%)	1063	(14%)	937	(14%)	
70–79	2202	(15%)	1205	(16%)	997	(14%)	
80–89	1749	(12%)	941	(13%)	808	(12%)	
90+	573	(4.0%)	270	(3.7%)	303	(4.4%)	
Under 70	9790	(68%)	4949	(67%)	4841	(70%)	0.001
Over 70	4524	(32%)	2416	(33%)	2108	(30%)	
Gender							
Female	8305	(58%)	4246	(58%)	4059	(58%)	0.36
Ethnicity							
Asian	2259	(16%)	1084	(15%)	1175	(17%)	0.001
Black	655	(4.6%)	332	(4.5%)	323	(4.6%)	
Unknown	1623	(11%)	816	(11%)	807	(12%)	
Mixed	260	(1.8%)	124	(1.7%)	136	(2.0%)	
Other	403	(2.8%)	199	(2.7%)	204	(2.9%)	
White	9114	(64%)	4810	(65%)	4304	(62%)	
Previous attendance in last 30 days	1805	(13%)	963	(13%)	842	(12%)	0.28
Referral source							
ED	9344	(65%)	4346	(59%)	4998	(72%)	<0.001
GP	4970	(35%)	3019	(41%)	1951	(28%)	
Length of stay (hours)							
0–12	6394	(45%)	4053	(55%)	2341	(34%)	<0.001
12–24	4196	(29%)	1590	(22%)	2606	(38%)	
24–36	2248	(16%)	1271	(17%)	977	(14%)	
36–48	1476	(10%)	451	(6%)	1025	(15%)	
Death (within 30 days)	35	(0.2%)	15	(0.2%)	20	(0.3%)	0.31
Readmission							
7 days	1047	(7.3%)	479	(6.5%)	568	(8.2%)	<0.001
14 days	1544	(11%)	681	(9%)	863	(12%)	<0.001
30 days	2268	(16%)	1033	(14%)	1235	(18%)	<0.001

For whole cohort, and for patients arriving in a normal working day (08:00–16:59 hours, Monday to Friday). P values shown for χ^2 comparison of normal working day episodes with episodes starting outside normal working day.
ED, emergency department; GP, general practitioner.

were included if they presented during a NWD, reflecting SDEC opening hours, leaving 7365 episodes in the analysis. The whole cohort and those presenting within a NWD are shown in table 2.

Eighteen per cent of episodes occurred on a weekend. Overall, 62% of patients arrived between 08:00 and 16:59 hours (figure 1); 63% of weekday episodes started between these times.

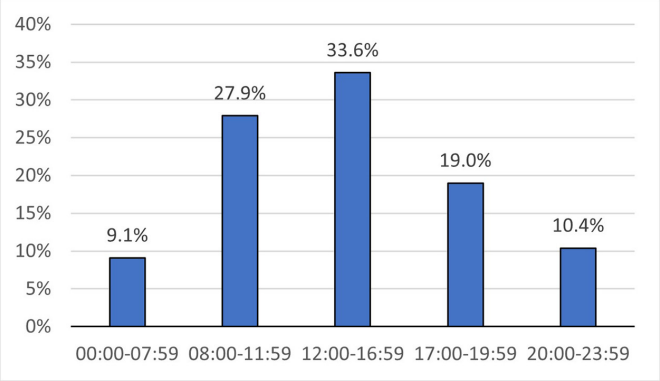

Figure 1 Arrival time for medical attendances lasting up to 48 hours.

There were 11 244 episodes that had an associated emergency department triage code, with 108 different triage codes used. The most common triage problem was chest pain (34% of episodes) (see online supplemental table 1); 6394 episodes (44%) had a length of stay of <12 hours.

Normal working day arrivals

There were 7365 episodes in 6848 patients with an arrival time between 08:00 and 16:59 hours on a weekday (NWD). The triage problem was available for 5272 NWD episodes (72%). The most common triage problem was chest pain (37%) (online supplemental table 1).

There were 4053 episodes (55%) that had a length of stay of <12 hours and 3312 (45%) were discharged after 12–48 hours. Patients arriving in NWD hours were more likely to be discharged within 12 hours than those arriving outside of these hours (55% vs 34%, χ^2 p<0.005).

There were <10 deaths (<0.2%) in those discharged in <12 hours and <10 deaths (<0.2%) in those discharged between 12 and 48 hours.

Compared with patients discharged within 12–48 hours, patients discharged within 12 hours had lower rates of readmission in the next 7 days (5.8% vs 7.4%, p=0.005), 14 days (8.2% vs 16.3%, p=0.001) and 30 days (12.2% vs 16.3%, p<0.005, χ^2 for all).

Factors affecting likelihood of discharge within 12 hours

Univariable comparison of the variables assessed within the original Amb score and GAPS derivation in NWD admissions is shown in table 3. Age ≥80 years and anticipated need for intravenous therapy were associated with an increased risk of admission lasting >12 hours. Absence of confusion, normal conscious level and absence of new neurological deficit were all associated with increased likelihood of discharge within 12 hours. Normal respiratory rate, oxygen saturations, heart rate between 50 and 140 bpm and systolic blood pressure between 100 and 200 mm Hg were associated with increased likelihood of discharge within 12 hours; a normal NEWS2 on arrival was associated with increased likelihood of discharge in <12 hours, but MEWS 0 was not. Patients with ischaemic heart disease, heart failure, cardiac arrhythmia, diabetes, previous stroke, chronic kidney disease or chronic lung disease were more likely to be admitted for >12 hours. In those with chest pain as their initial triage problem (1940 patients), those with a suspicion of acute coronary syndrome coded into the emergency department diagnosis were more likely to be admitted for >12 hours (OR 0.80, p=0.025, 95% CI 0.66 to 0.97).

Amb score

Multivariable analysis including all components of the Amb score, except access to transportation (which was present for all patients), is shown in online supplemental table 2. The variables of sex, acute confusion, MEWS and recent hospital admission did not predict likelihood of discharge within 12 hours in this multivariable analysis. Replacing MEWS with the currently used NEWS2 acuity score, there remained no association of sex, acute confusion and recent hospital admission with likelihood of discharge within 12 hours, however NEWS2 of zero was associated with increased likelihood of discharge within 12 hours.

The Amb score could be calculated for 6743 episodes (online supplemental table 3). Ninety-four per cent (6325 admissions) had an Amb score of 5 or more, suggesting they could be discharged within 12 hours; 6.2% (418 admissions) had a score of <5.

The AUROC for the Amb score was 0.601 (95% CI 0.588 to 0.614) (figure 2A). Score performance is shown in table 4. Of those with a raised Amb score suggesting suitability for SDEC, 55% were discharged within 12 hours of arrival (the positive predictive value (PPV), 95% CI 53.8% to 56.2%); 12% of those with an Amb score of <5 were discharged within 12 hours. The sensitivity of the Amb score for identifying patients discharged within 12 hours was 98.6% (95% CI 98.1% to 98.9%). Overall, 57% of patients were correctly identified (Amb score 5+ suggesting suitability for SDEC and length of stay <12 hours, or Amb score <5 and length of stay 12–48 hours).

Replacing MEWS with NEWS2, the AUROC was 0.612 (95% CI 0.599 to 0.625) (figure 2B). Ninety-five per cent (6343 admissions) had an Amb score of 5 or more; 5.4% (364 admissions) had a score of <5. Of those with a raised Amb score incorporating NEWS2, 54.5% were discharged within 12 hours of arrival (PPV, 95% CI 53.8% to 56.2%); 12% of those with a score <5 were discharged within 12 hours. The sensitivity of the Amb score including NEWS2 for identifying patients discharged within 12 hours was 98.8% (95% CI 98.4% to 99.1%). Overall, 56% of patients were correctly identified. There was no significant difference in the performance of the Amb score incorporating MEWS and the Amb score incorporating NEWS2 (table 4).

Those with a low Amb score were more likely to be readmitted within 7 days (13.7% vs 5.8%, χ^2 p=0.017), in both those discharged within 12 hours (13.7% vs 5.8%, p=0.017) and those discharged in 12–48 hours (11.7% vs 7.0%, p=0.001). This was also true for readmission within

Table 3 Factors considered in derivation of previous scoring systems

	Length of stay						
	<12 hours		12–48 hours				
N=7365 unless otherwise stated	Frequency (%)		Frequency (%)		OR	P value	95% CI
Age (years)							
16–19	94	(2.3%)	78	(2.4%)	Ref		
20–29	392	(9.7%)	332	(10.%)	1.00	0.99	0.66 to 1.54
30–39	477	(12%)	349	(11%)	0.85	0.45	0.56 to 1.29
40–49	548	(14%)	361	(11%)	0.74	0.17	0.49 to 1.13
50–59	746	(18%)	509	(15%)	0.77	0.21	0.51 to 1.16
60–69	641	(16%)	422	(13%)	0.73	0.14	0.48 to 1.11
70–79	634	(16%)	571	(17%)	1.11	0.62	0.74 to 1.67
80–89	437	(11%)	504	(15%)	1.52	0.049	1.00 to 2.32
90+	84	(2.1%)	186	(5.6%)	2.69	<0.001	2.07 to 5.87
≥80	521	(13%)	690	(21%)	2.11	<0.001	1.76 to 2.52
Sex (*n=7363*)							
Male	1713	(42%)	1404	(42%)	1.00	0.96	0.89 to 1.13
Intravenous treatment not anticipated	3953	(98%)	2704	(82%)	0.08	<0.001	0.06 to 0.11
Not discharged in previous 30 days	3518	(87%)	2884	(87%)	1.02	0.79	0.86 to 1.21
Not admitted within last 1 year	2510	(62%)	1813	(55%)	0.70	<0.001	0.62 to 0.79
No neurological deficit*	4024	(99.3%)	3241	(97.9%)	0.25	<0.001	0.14 to 0.43
Not acutely confused (*n=6745*)	3526	(99.9%)	3197	(99.5%)	0.20	0.007	0.06 to 0.64
Physiological observations							
Normal temperature (*n=6743*)	2524	(72%)	2242	(70%)	0.90	0.12	0.80 to 1.03
Normal RR (*n=6735*)	3437	(98%)	2994	(93%)	0.29	<0.001	0.21 to 0.41
O_2 saturations >95% (*n=6738*)	2988	(85%)	2525	(79%)	0.62	<0.001	0.53 to 0.73
Heart rate 50–140 bpm (*n=6748*)	3499	(99.0%)	3144	(97.9%)	0.42	<0.001	0.25 to 0.69
SBP 100–200 mm Hg (*n=6753*)	3430	(96.9%)	3040	(94.6%)	0.49	<0.001	0.37 to 0.67
Alert (*n=6745*)	3524	(99.8%)	3170	(98.6%)	0.10	<0.001	0.04 to 0.25
MEWS 0 (*n=6764*)	132	(4%)	116	(4%)	0.96	0.80	0.71 to 1.31
NEWS2 0 (*n=6712*)	1381	(39%)	1012	(32%)	0.66	<0.001	0.58 to 0.75
NEWS2 0–2 (*n=6712*)	3213	(92%)	2598	(81%)	0.33	<0.001	0.27 to 0.41
NEWS2 (*n=6712*)							
0	1381	(39%)	1012	(32%)	Ref		
1	1332	(38%)	1103	(34%)	1.15	0.038	1.01 to 1.32
2	500	(14%)	483	(15%)	1.39	<0.001	1.16 to 1.66
3	188	(5.4%)	272	(8.5%)	2.20	<0.001	1.71 to 2.83
4	71	(2.0%)	132	(4.1%)	2.96	<0.001	1.05 to 4.28
5	21	(0.6%)	91	(2.8%)	7.76	<0.001	4.35 to 13.8
≥6	12	(0.3%)	114	(3.6%)	18.5	<0.001	9.15 to 37.5
Previous medical history							
No history of IHD	3116	(77%)	2446	(74%)	0.82	0.004	0.71 to 0.94
No history of heart failure	3925	(97%)	3113	(94%)	0.44	<0.001	0.33 to 0.59
No history of arrhythmia	3689	(91%)	2787	(84%)	0.44	<0.001	0.36 to 0.54
No history of diabetes	3476	(86%)	2667	(81%)	0.62	<0.001	0.53 to 0.73
No history of stroke	4033	(99.5%)	3229	(97.5%)	0.14	<0.001	0.07 to 0.25
No history of renal disease	3866	(95%)	3064	(93%)	0.52	<0.001	0.40 to 0.67

Continued

Table 3 Continued

N=7365 unless otherwise stated	Length of stay <12 hours Frequency (%)		Length of stay 12–48 hours Frequency (%)		OR	P value	95% CI
No history of chronic lung disease	3264	(81%)	2530	(76%)	0.75	<0.001	0.65 to 0.86
Factors on arrival							
Arrival by ambulance	1080	(27%)	1384	(42%)	2.23	<0.001	1.94 to 2.57
Referred by GP	2111	(52%)	908	(27%)	0.28	<0.001	0.24 to 0.34
Triage category (*n=5272*)							
Standard	264	(11%)	220	(7.6%)	Ref		
Urgent	2072	(88%)	2427	(84%)	1.45	0.001	1.17 to 1.80
Resuscitation	27	(1.1%)	262	(9.0%)	14.2	<0.001	8.30 to 24.2

Column percentages shown. Univariate analysis, OR for admission lasting 12–48 hours shown.
Normal ranges for physiological parameters (temperature, heart rate) as defined by the NEWS2 scoring system.[9] Presence of comorbidities assessed from diagnostic codes.
*Neurological deficit recorded as present if neurological deficit was recorded in triage coding of the presenting problem for the admission episode.
GP, general practitioner; IHD, ischaemic heart disease; MEWS, Modified Early Warning Score; NEWS2, National Early Warning Score 2; Ref, reference; RR, respiratory rate; SBP, systolic blood pressure.

30 days (25.6% vs 13.6%, p<0.001), in those discharged within 12 hours (23.5% vs 12.2%, p=0.015) and those discharged in 12–48 hours (25.9% vs 15.3%, p<0.001). This difference remained when substituting in NEWS2 (7 days: 12.1% vs 6.4%, p<0.001; 30 days: 25.3% vs 13.8%, p<0.001), and when assessing episode without another episode in the preceding 30 days (7 days: 11.3% vs 5.6%, χ^2 p<0.001; 30 days: 24.5% vs 12.1%, p<0.001).

Impact on patient pathway

Patient pathways through acute care incorporating the Amb score were estimated (figure 3A). Directing short

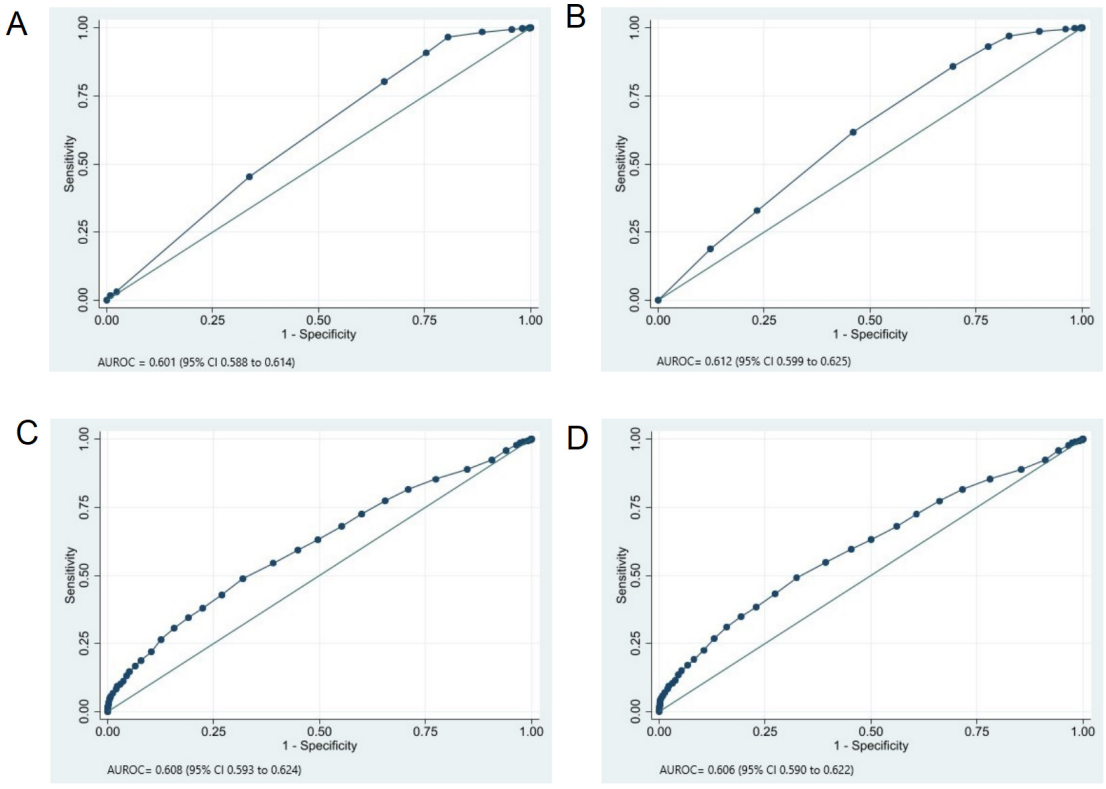

Figure 2 Receiver operator characteristics (ROC) curve for score performance. (A) Amb score; (B) Amb score substituting National Early Warning Score 2 (NEWS2); (C) Glasgow Admission Prediction Score (GAPS); (D) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions.

Table 4 Amb score performance

	Amb score N=6743		Amb score with NEWS2[9] N=6707	
	Frequency (%)		Frequency (%)	
Score				
<5	418	(6.2%)	364	(5.4%)
5+	6325	(93.8%)	6343	(94.6%)
Score <5 Admission length <12 hours	51	(0.8%)	42	(0.6%)
Score <5 Admission length 12–48 hours	367	(5.4%)	322	(4.8%)
Score 5+ Admission length <12 hours	3479	(51.6%)	3459	(51.6%)
Score 5+ Admission length 12–48 hours	2846	(42.2%)	2884	(43.0%)
Score performance	Measures of diagnostic accuracy (95% CI)			
Sensitivity	98.6% (98.1% to 98.9%)		98.8% (98.4% to 99.1%)	
Specificity	11.4% (10.3% to 12.6%)		10.0% (9.0% to 11.1%)	
PPV	55.0% (53.8% to 56.2%)		54.5% (53.3% to 55.8%)	
NPV	87.8% (84.3% to 90.8%)		88.5% (84.7% to 91.6%)	
% of patients discharged in <12 hours not identified by score*	1.4% (1.1% to 2%)		1.2% (0.9% to 1.6%)	
Patients identified as suitable by score admitted for >12 hours†	45.0% (43.8% to 46.2%)		45.5% (44.2% to 46.7%)	

Performance in normal working day admissions.
*1−sensitivity.
†1−PPV.
NEWS2, National Early Warning Score 2; NPV, negative predictive value; PPV, positive predictive value.

stay patients with an Amb score of 5 or more to SDEC, 45% of patients seen in SDEC services would require admission for >12 hours. For an acute medical service assessing 50 potential short stay medical admissions per day, this would mean approximately 47 patients would be seen in SDEC and 22 of these would require admission to an AMU or inpatient ward after review in SDEC. Three patients per day would be streamed directly to AMU, with 1% of those streamed to AMU discharged within 12 hours.

Score performance in patient subgroups

The proportion of patients identified correctly varied when comparing patient subgroups (online supplemental table 4). In those with a raised Amb score suggesting suitability for SDEC, a lower proportion of patients were discharged within 12 hours where patients were aged over 70 years, and where comorbidity due to ischaemic heart disease, heart failure, arrhythmia, diabetes, stroke/transient ischaemic attack (TIA), renal disease or chronic lung disease was present. A higher proportion of general practitioner (GP) referrals with a raised Amb score were discharged within 12 hours, compared with those whose first healthcare contact was the emergency department (69% vs 45%, χ^2 p<0.005). A higher proportion of patients with a raised Amb score and a NEWS2 of 0–2 were identified correctly compared with those with a raised NEWS2 on arrival.

Glasgow Admission Prediction Score

Multivariable analysis including all components of the GAPS is shown in online supplemental table 5. Increasing age, increasing NEWS or NEWS2, arrival by ambulance, triage categorisation of requiring resuscitation level care and previous admission within the last 12 months were all associated with increased likelihood of admission for >12 hours. Referral from a GP was associated with increased likelihood of discharge within 12 hours, and not admission.

The GAPS could be calculated for 5091 NWD admissions with scores ranging between 1 and 53 (online supplemental table 6).

The AUROC for the GAPS was 0.608 (95% CI 0.593 to 0.624) (figure 2C). As a binary predictor, 2912 admissions (57%) had a GAPS >15, suggesting need for admission (table 5). Of those with a GAPS of 15 or less, 51.4% were discharged within 12 hours (PPV, 95% CI 49.3% to 53.6%). The sensitivity of the GAPS for identifying patients discharged within 12 hours was 50.4% (95% CI

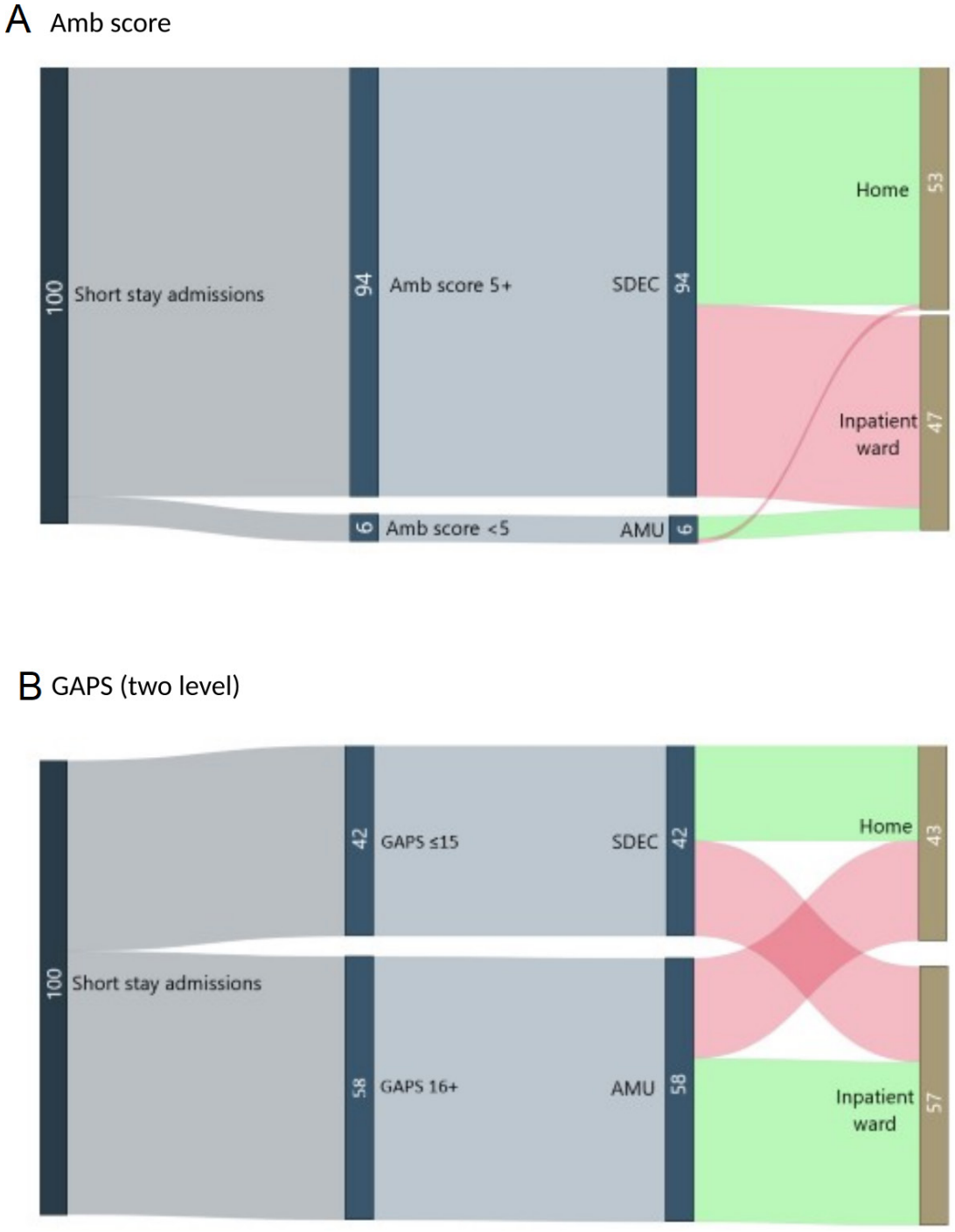

Figure 3 Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when using scoring systems to identify patients for assessment in same day emergency care (SDEC) for (A) Amb score (5 or more) and (B) Glasgow Admission Prediction Score (GAPS) (≤15). Green=currently identified by scoring system, red=incorrectly identified by scoring system. AMU, acute medical unit.

48.5% to 52.5%), with a negative predictive value (NPV) of 62.1% (95% CI 60.3% to 63.9%). Overall, 57.5% of patients were correctly identified (GAPS ≤15 suggesting suitability for SDEC and length of stay <12 hours, or GAPS >15 and length of stay 12–48 hours).

Substituting NEWS2 for NEWS, the AUROC was 0.606 (95% CI 0.590 to 0.622) (figure 2D). As a binary predictor, 2852 admissions (57.6%) had a GAPS (incorporating NEWS2) >15, suggesting need for admission. Of those with a GAPS of 15 or less, 50.5% (1062 episodes) were discharged within 12 hours (PPV, 95% CI 48.4% to 52.7%). The sensitivity of the GAPS for identifying patients discharged within 12 hours was 50.0% (95% CI 47.8% to 52.1%), with a NPV of 62.7% (95% CI 60.9% to 64.5%). Again, 57.5% of patients were correctly identified. Substituting NEWS2 for NEWS within the GAPS did not significantly alter performance of the score (table 5).

Table 5 GAPS performance within normal working day admissions

	GAPS N=5091		GAPS with NEWS2 N=4953	
	Frequency (%)		Frequency (%)	
Score				
≤15	2179	(42.8%)	2101	(42.4%)
16+	2912	(57.2%)	2852	(57.6%)
Score ≤15 Admission length <12 hours	1121	(22.0%)	1062	(21.4%)
Score ≤15 Admission length 12–48 hours	1058	(20.8%)	1039	(21.0%)
Score 16+ Admission length <12 hours	1104	(21.7%)	1063	(21.5%)
Score 16+ Admission length 12–48 hours	1808	(35.5%)	1789	(36.1%)
Score performance	Measures of diagnostic accuracy (95% CI)			
Sensitivity	50.4% (48.5 to 52.5%)		50.0% (47.8% to 52.1%)	
Specificity	63.1% (61.3% to 64.9%)		63.3% (61.5% to 65.0%)	
PPV	51.4% (49.3% to 53.6%)		50.5% (48.4% to 52.7%)	
NPV	62.1% (60.3% to 63.9%)		62.7% (60.9% to 64.5%)	
% of patients discharged in <12 hours not identified by score*	49.6% (47.5% to 51.5%)		50.0% (47.9% to 52.2%)	
Patients identified as suitable by score admitted for >12 hours†	48.6% (46.4% to 50.7%)		49.5% (47.3% to 51.6%)	

*1−sensitivity.
†1−PPV.
GAPS, Glasgow Admission Prediction Score; NEWS2, National Early Warning Score 2; NPV, negative predictive value; PPV, positive predictive value.

Dividing into three risk quantiles, a score of 13 or less (1613 episodes, 32.6%) denotes 'low risk', a score of 14–19 (1536 episodes, 31.0%) denotes medium risk and a score of 20 or more (1804 episodes, 36.4%) denotes high risk. For 'low-risk' patients, 57.8% (835 episodes) were discharged within 12 hours, compared with 46.2% of those with a 'medium-risk' score, and 32.2% of those with a 'high-risk' score.

Those with a GAPS ≥16 were more likely to be readmitted within 7 days (7.4% vs 5.1%, χ^2 p<0.005), both for those discharged within 12 hours (6.0% vs 4.2%, p=0.055), and 12–48 hours (8.3% vs 6.1%, p=0.027). Patients with a GAPS ≥16 were also more likely to be readmitted within 30 days (16.9% vs 10.7%, p<0.005), in those discharged within 12 hours (13.3% vs 9.0%, p=0.001) and those discharged within 12–48 hours (19.0% vs 12.6%, p<0.005). This difference remained when substituting in NEWS2 (7 days: 7.4% vs 5.2%, p<0.005; 30 days: 16.9% vs 11.0%, p<0.005), and when assessing episode without another episode in the preceding 30 days (7 days: 6.1% vs 4.5%, p=0.02; 30 days: 14.4% vs 9.7%, p<0.001).

Estimated impact on patient pathway

Patient pathways through acute care incorporating the GAPS were estimated (figure 3B). Directing short stay patients with a GAPS of 15 or less to SDEC, 50% of patients seen in SDEC services would require admission for >12 hours. For an acute medical service assessing 50 short stay medical admissions per day (100 admissions in total), this would mean approximately 21 patients would be seen in SDEC and 10 of these would require admission to an AMU or inpatient ward after review in SDEC. Twenty-nine patients would be streamed directly to AMU, 11 of these patients would be discharged from hospital within 12 hours, and therefore would have been suitable for management via SDEC.

Score performance in patient subgroups

In those with a low GAPS suggesting suitability for SDEC, a lower proportion of patients were discharged within 12 hours where patients were aged over 70 years, were female and where comorbidity due to stroke/TIA was present (online supplemental table 7). A higher proportion of GP referrals with a low GAPS were discharged within 12 hours, compared with those whose first healthcare contact was the emergency department (68% vs 50%, χ^2 p=0.044). A higher proportion of patients with a low GAPS and a NEWS2 of 0–2 were identified correctly compared with those with a raised NEWS2 on arrival.

Differences in patient identification between the two scores

There were 4952 episodes where both the Amb score and GAPS could be calculated. Using both scores (with NEWS2 incorporated), there were 2332 patient episodes

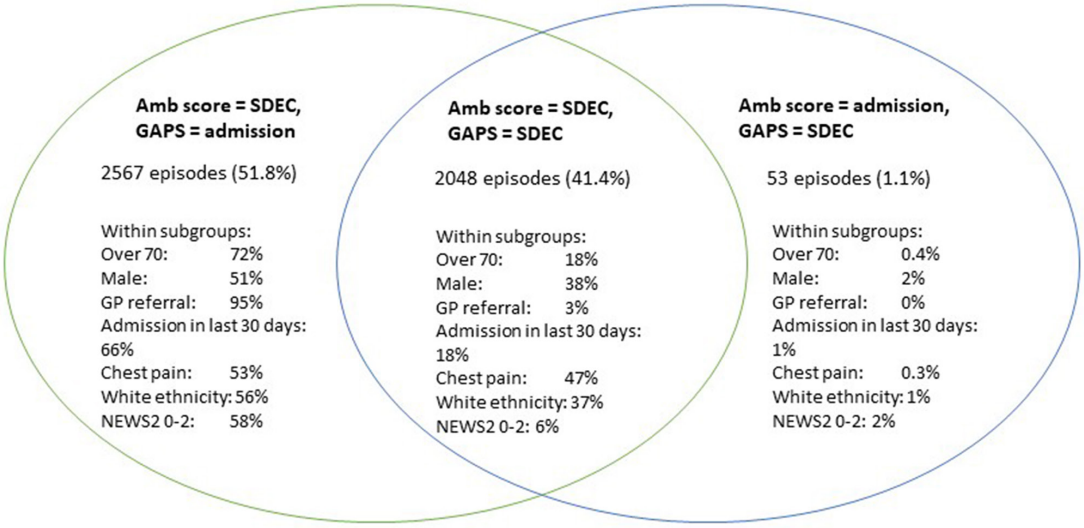

Figure 4 Agreement of Amb score and Glasgow Admission Prediction Score (GAPS) score in identification of patients suitable for same day emergency care (SDEC). Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown. GP, general practitioner; NEWS2, National Early Warning Score 2.

(47%) where the scoring systems agreed. In 2048 episodes (41%), both scores suggested the patient was suitable for SDEC (Amb score 5+ and GAPS ≤15) and in 284 episodes (6%) both scores suggested the patient was likely to require admission (Amb score <5 and GAPS 16+). In 2620 episodes (53%), the recommendation provided by the score differed. There were 2567 episodes (52%), where the Amb score suggested suitability for SDEC while the GAPS suggested admission was likely and 53 episodes (1%) where the GAPS suggested likely discharge but the Amb score predicted admission. Those aged over 70 years, referred by their GP, with a NEWS2 of 0–2 or who had been admitted in the last 30 days were more likely to have an Amb score suggesting suitability for SDEC with a GAPS suggesting admission (χ^2, p<0.0005 for each subgroup comparison, figure 4).

DISCUSSION

This paper highlights several important points. First, this analysis suggests that both the Amb score and the GAPS have limited ability to discriminate between patients discharged within 12 hours and those discharged in 12–48 hours in this diverse and urban health setting. Both scores had an AUROC suggesting they could not identify those discharged within 12 hours to an acceptable level, with the Amb score having an AUROC of 0.612 and GAPS an AUROC of 0.606. Score performance was worse than in previously published research, with the Amb score suggested to have an AUROC of 0.91 (95% CI 0.88 to 0.94) in the original derivation study,[5] and 0.743 (95% CI 0.717 to 0.769) in a subsequent evaluation,[11] and the GAPS having an AUROC of 0.877 (95% CI 0.875 to 0.880) during its original derivation[7] and 0.807 (95% CI 0.785 to 0.830) on subsequent assessment.[11] In our analysis, the Amb score has a higher NPV than the GAPS, with 88.5% of patients with a low Amb score (suggesting they were unsuitable for SDEC) remaining for >12 hours, compared with 62.7% of those with a high GAPS. Although differences in performance may relate to utilisation in a setting that differs from the original studies (online supplemental table 8), this reflects potential performance when implemented in clinical practice in our setting.

Second, some components of both scores included as factors to predict admission or discharge were non-discriminatory in this patient cohort. Multivariable analysis suggested that sex and confusion did significantly affect admission length when considered with other Amb score components, and sex was not associated with longer length of stay in univariate analysis. This may reduce overall performance of the Amb score within our population. Previous research suggests confusion is associated with increased length of hospital stay[12]; differences in admission length in our analysis may have been masked as only a small number of patients had new confusion recorded. Within multivariable analysis of GAPS components, and within univariate analysis, referral from GP was associated with decreased likelihood of admission for >12 hours. This contradicts the original GAPS derivation study, where referral from GP was associated with increased likelihood of admission.[7] This will affect performance of the GAPS in our cohort, and highlights the importance of evaluating the influence of each score component in local patient cohorts. Underlying reasons for this difference, such as availability of local referral pathways or additional community services, cannot be assessed within this analysis.

Third, there was a marked difference in the proportion of patients that would be directed through SDEC services when implementing the two scores, with the Amb score directing 94% of this short stay cohort and GAPS only 42%. This suggests that score choice may have considerable impact on patient pathway and subsequent

service demand. There was also significant divergence in the patients identified for SDEC by the Amb score and GAPS. Conflicting recommendations were more likely in those aged over 70 years, referred by their GP, or with a normal NEWS2 score. This highlights specific subgroups of patients within our cohort where implementation of either scoring system into clinical practice may impact access to SDEC services.

Fourth, updating both the Amb score and GAPS with NEWS2 did not noticeably improve performance. NEWS2 was incorporated into both scores within this analysis to reflect current practice.[9] Within the Amb score, and in univariate analysis, NEWS2 appeared to be a more significant predictor than MEWS. This may reflect the low number of patients with a MEWS of zero on arrival; a higher proportion of patients had a NEWS2 of zero due to the amended normal ranges of the early warning score components.

Implementing the Amb score or GAPS to select patients for review in SDEC within our cohort would result in >45% of patients assessed in SDEC requiring subsequent admission to an inpatient bed. This is likely to be higher than is acceptable for both patient experience and flow through acute services. As SDEC services have a fixed capacity, with limited space and staffing, each patient awaiting admission within SDEC services reduces the capacity to deliver SDEC to subsequent patients that day and may expose patients to additional delays due to multiple location changes and waits for inpatient beds.

Limitations

This analysis was restricted admissions during 'normal working' hours to reflect operation of SDEC services. Most SDEC services in the UK operate during daytime hours with associated increased availability of investigations and specialty input.[13] Scoring system performance outside these hours may differ, due to differences in access to services and in the patient cohort admitted outside daytime hours.[14]

This analysis focused on performance of scoring systems to identify patients suitable for SDEC within currently available services; in-depth evaluation of factors necessitating admission >12 hours, for example, ongoing therapy input or delays in diagnostic imaging, were outside the scope of this analysis. Pathway changes facilitating discharge within 12 hours, such as ambulatory pathways, may alter performance of any patient selection scoring system, and should therefore prompt reassessment of score performance.

This analysis focused on the ability of the Amb score and GAPS to discriminate between those admitted for <12 hours and 12–48 hours. Applying the Amb score or GAPS across all medical admissions, including those with a length of stay over 48 hours, will affect the PPV and NPV of the score. Although some aspects of score performance may be appear improved if the scores are able to identify all those admitted for >48 hours correctly, the proportion of patients incorrectly directed through SDEC will not improve. If some patients with a length of stay >48 hours have a raised Amb score or low GAPS, then the PPV will be lower than suggested within this analysis, resulting in a higher proportion of patients deemed 'suitable for SDEC' being admitted to inpatient wards.

GAPS was assessed as a binary outcome using a cut-off of 15 to indicate higher likelihood of discharge within 12 hours, although adjusting the cut-off to maximise performance within each centre is advised.[7] Full analysis of the potential impact of using alternative cut-offs on patient selection and pathway use was not performed, as multivariable analysis suggested components of the score were not performing as expected within this patient cohort.

This analysis used retrospective data. Amb score calculation presumed intravenous treatment to be 'anticipated' in patients receiving intravenous treatment within 6 hours of arrival, as anticipation of intravenous therapy is not routinely collected with EHR. This may have altered the patients receiving points for this component. Both scores were calculated only for patients where data were available for all components. For the GAPS score, this restricted included episodes to those where patients arrived through the emergency department, as direct arrivals to AMU do not receive categorisation of triage urgency. This may affect score performance when assessing the overall cohort, particularly in patients referred from their GP. The missing scores highlight potential issues when considering implementation; in routinely collected EHR data, score components may be incompletely documented. This should be considered when evaluating proposed scoring systems, as performance in real-world healthcare settings will be influenced by data availability.

These scores were suggested to be used at triage on initial arrival. Implementing these scores prospectively in clinical practice may alter the length of patients' pathways through acute services, and therefore length of stay. This may have some impact on the number of patients discharged within 12 hours, therefore any scoring system to be implemented would require prospective evaluation.

This study took place within a UK setting, and there is considerable variability in the structure of acute care services internationally, including in the delivery of ambulatory services for patients with acute medical emergencies.[15] However, increased demand for acute services is noted in other healthcare systems,[16 17] and so methods for identifying patients suitable to be managed without inpatient admission may be beneficial in these settings.

CONCLUSION

Within this patient cohort, the Amb score and GAPS could not accurately identify acute medical admissions that were likely to be discharged within 12 hours of admission, limiting their utility in selecting patients suitable for SDEC services.

Author affiliations
[1]Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, UK
[2]Department of Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
[3]Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

Contributors CA and ES designed the study, CA analysed the data, all authors contributed to interpretation of the data and approved the final manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. CA is acting as guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; CA is funded by an NIHR clinical lectureship. ES reports grant funding from HDR UK, Wellcome Trust, MRC, BLF, NIHR, EPSRC and Alpha 1 Foundation; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the 'Methods' section for further details.

Patient consent for publication Not applicable.

Ethics approval This research was performed in accordance with the Declaration of Helsinki. All study processes were carried out following appropriate ethical approval provided for PIONEER, the Health Data Research UK Hub in acute care by the East Midlands—Derby REC (reference: 20/EM/0158). Formal written consent from individual participants was not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Data from this study are available from PIONEER, the Health Data Hub in Acute care, in accordance with Hub processes. See www.pioneerdatahub.co.uk and contact PIONEER@uhb.nhs.uk for more details.

ORCID iDs
Catherine Atkin http://orcid.org/0000-0003-0596-8515
Elizabeth Sapey http://orcid.org/0000-0003-3454-5482

REFERENCES

1 NHS England. A&E Attendances and Emergency Admissions 2021-22, 2021. Available: https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/ae-attendances-and-emergency-admissions-2021-22/
2 NHS England, NHS Improvement. Same-day emergency care: clinical definition, patient selection and metrics; 2019.
3 National Health Service. The NHS long term plan 2019.
4 Atkin C, Riley B, Sapey E. How do we identify acute medical admissions that are suitable for same day emergency care? *Clin Med* 2022;22:131–9.
5 Ala L, Mack J, Shaw R, *et al*. Selecting ambulatory emergency care (AEC) patients from the medical emergency in-take: the derivation and validation of the Amb score. *Clin Med* 2012;12:420–6.
6 Royal College of Physicians. Acute care toolkit 10: ambulatory emergency care; 2014.
7 Cameron A, Rodgers K, Ireland A, *et al*. A simple tool to predict admission at the time of triage. *Emerg Med J* 2015;32:174–9.
8 Royal College of Physicians. National early warning score (NEWS): standardising the assessment of acute-illness severity in the NHS; 2012.
9 Royal College of Physicians. National early warning score (NEWS) 2; 2017.
10 Atkin C, Knight T, Cooksley T, *et al*. Length of stay in acute medical admissions: analysis from the Society for Acute Medicine Benchmarking Audit. *Acute Med* 2022;21:27–33.
11 Cameron A, Jones D, Logan E, *et al*. Comparison of Glasgow admission prediction score and Amb score in predicting need for inpatient care. *Emerg Med J* 2018;35:247–51.
12 Pendlebury ST, Lovett NG, Smith SC, *et al*. Observational, longitudinal study of delirium in consecutive unselected acute medical admissions: age-specific rates and associated factors, mortality and re-admission. *BMJ Open* 2015;5:e007808.
13 Society for Acute Medicine. Society for Acute Medicine Benchmarking Audit 2021 - SAMBA2021 Report; 2021.
14 Atkin C, Knight T, Subbe C, *et al*. Acute care service performance during winter: report from the winter SAMBA 2020 national audit of acute care. *Acute Med* 2020;19:220–9.
15 Baier N, Geissler A, Bech M, *et al*. Emergency and urgent care systems in Australia, Denmark, England, France, Germany and the Netherlands - Analyzing organization, payment and reforms. *Health Policy* 2019;123:1–10.
16 Canadian Institute for Health Information. NACRS emergency department visits and lengths of stay, 2022. Available: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-lengths-of-stay
17 Australian Institute of Health and Welfare. Admitted patients, 2022. Available: https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients