BMJ Open

Performance of scoring systems in selecting short stay medical admissions suitable for assessment in Same Day Emergency Care

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064910
Article Type:	Original research
Date Submitted by the Author:	17-May-2022
Complete List of Authors:	Atkin, Catherine; University of Birmingham, Birmingham Acute Care Research Group Gallier, Suzy; University Hospitals Birmingham NHS Foundation Trust, Department of Health Informatics Wallin, Elizabeth; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine Reddy-Kolanu, Vinay; University Hospitals Birmingham NHS Foundation Trust, Acute medicine Sapey, Elizabeth; University of Birmingham, PIONEER HDR-UK Hub; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine
Keywords:	INTERNAL MEDICINE, GENERAL MEDICINE (see Internal Medicine), Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Performance of scoring systems in selecting short stay medical admissions suitable for assessment in Same Day Emergency Care

Catherine Atkin (0000-0003-0596-8515), Suzy Gallier (0000-0003-1026-4125), Elizabeth Wallin, Vinay Reddy-Kolanu, Elizabeth Sapey (0000-0003-3454-5482)

Addresses

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Edgbaston, Birmingham, B15 2GW, UK. Catherine Atkin, NIHR Academic Clinical Lecturer in Acute Medicine.

Department of Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Suzy Gallier, Lead for Research Analytics.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Wallin, Consultant in Nephrology & Acute Medicine.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Vinay Reddy-Kolanu, Consultant in Acute Medicine.

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, and Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Sapey, Professor of Acute and Respiratory Medicine and Honorary Consultant.

Correspondence to: 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 and lasting up to 48 hours, between 1st April 2019 and 9th 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 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

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, and with the facility for patients to return for further investigations on subsequent days as needed, without admission to a hospital bed. 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 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,(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 less than 12 hours or over 48 hours. The study excluded patients who remained in hospital for 12 to 48 hours.

Table 1: Scoring systems to identify medical admissions potentially suitable for discharge from hospital without admission >12 hours. Amb score(5) and Glasgow Admission Prediction Score (GAPS)(7). Amb score of 5 more indicates likely discharge within 12 hours; GAPS of 16 or more suggests patient likely to be admitted to hospital. IV = intravenous, MEWS = Modified early warning score, NEWS = National Early Warning Score, GP = General practitioner

Amb score			**Glasgow Admissions Prediction 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			
IV 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			

The Glasgow Admission Prediction Score (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, assessing in particular the scores' ability to discriminate between acute medical admissions suitable for Same Day Emergency Care and those requiring admission for at least 12 to 48 hours.

Methods

This data study was conducted in collaboration with PIONEER, a Health Data Research Hub in Acute Care, and all study processes were carried out following appropriate ethical approval provided by the East Midlands – Derby REC (reference: 20/EM/0158).

Retrospective data were collected for patients admitted to Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Trust (UHB) between the period of the 1st April 2019 until 9th 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 (CCU) in Europe. The Acute Medical Unit (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 inhouse electronic health record (EHR) called Prescribing Information and Communication System (PICS). 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 focussed 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 and mapped ICD10 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 to 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 08:00-16:59, 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 (IV) treatment was taken as not being anticipated where patients did not receive an IV 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 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 National Early Warning Score 2 (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 15.1. Cell counts containing less than 10 patients were suppressed, due to reporting requirements. For univariate analysis of factors influencing likelihood of discharge within 12 hours, odds ratios for variables included in the original Amb score or GAPS derivation studies were assessed using Chi square. Multivariate analysis of the Amb score and GAPS components was performed using logistic regression, to demonstrate the performance of components within the score in this cohort. Receiver operator characteristic (ROC) curves were calculated for each scoring system, and the area under the ROC (AUROC) calculated. Comparison of proportions was performed using Chi square. A p value of <0.05 is used to signify statistical significance throughout.

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

Results

14314 acute medical inpatient episodes lasting up to 48 hours were identified during the censor period. These episodes were from 12587 patients with 11229 patients having one episode in this time period. Patients 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.

Table 2: Demographics and characteristics of patients with emergency medical admissions lasting up to 48 hours. For whole cohort, and for patients arriving in a normal working day (08:00-16:59, Monday to Friday).P values shown for Chi square comparison of normal working day episodes to episodes starting outside normal working day.

	All episodes N=14314		Normal working day episodes N=7365		Episodes starting outside normal working day N= 6949		P value
	Frequency (%)		Frequency (%)		Frequency (%)		
Age							
16-19	444	(3.1%)	172	(2.3%)	272	(3.9%)	<0.001
20-29	1585	(11.1%)	724	(9.8%)	861	(12.4%)	
30-39	1677	(11.7%)	826	(11.2%)	851	(12.2%)	
40-49	1776	(12.4%)	909	(12.3%)	867	(12.5%)	
50-59	2308	(16.1%)	1255	(17.0%)	1053	(15.2%)	
60-69	2000	(14.0%)	1063	(14.4%)	937	(13.5%)	
70-79	2202	(15.4%)	1205	(16.4%)	997	(14.3%)	
80-89	1749	(12.2%)	941	(12.8%)	808	(11.6%)	
90+	573	(4.0%)	270	(3.7%)	303	(4.4%)	
Under 70	9790	(68.4%)	4949	(67.2%)	4841	(69.7%)	0.001
Over 70	4524	(31.6%)	2416	(32.8%)	2108	(30.3%)	
Gender							
Female	8305	(58.0%)	4246	(57.7%)	4059	(58.4%)	0.357
Ethnicity							
Asian	2259	(15.8%)	1084	(14.7%)	1175	(16.9%)	0.001
Black	655	(4.6%)	332	(4.5%)	323	(4.6%)	
Unknown	1623	(11.3%)	816	(11.1%)	807	(11.6%)	
Mixed	260	(1.8%)	124	(1.7%)	136	(2.0%)	
Other	403	(2.8%)	199	(2.7%)	204	(2.9%)	
White	9114	(63.7%)	4810	(65.3%)	4304	(61.9%)	
Previous attendance in last 30 days	1805	(12.6%)	963	(13.1%)	842	(12.1%)	0.283
Referral source							
ED	9344	(65.3%)	4346	(59.0%)	4998	(71.9%)	<0.001
GP	4970	(34.7%)	3019	(41.0%)	1951	(28.1%)	
Length of stay (hours)							
0-6	3005	(20.1%)	2237	(30.4%)	768	(11.1%)	<0.001
6-12	3389	(23.7%)	1816	(24.7%)	1573	(22.6%)	
12-18	2124	(14.8%)	687	(9.3%)	1437	(20.7%)	
18-24	2072	(14.5%)	903	(12.3%)	1169	(16.8%)	
24-30	1508	(10.5%)	970	(13.2%)	538	(7.7%)	
30-36	740	(5.2%)	301	(4.1%)	439	(6.3%)	
36-42	662	(4.6%)	165	(2.2%)	497	(7.2%)	
42-48	814	(5.7%)	286	(3.9%)	528	(7.6%)	
Death (30 days)	35	(0.2%)	15	(0.2%)	20	(0.3%)	0.308
Readmission							
7 day	1047	(7.3%)	479	(6.5%)	568	(8.2%)	<0.001
14 day	1544	(10.8%)	681	(9.3%)	863	(12.4%)	<0.001
30 day	2268	(15.8%)	1033	(14.0%)	1235	(17.8%)	<0.001

18.4% of episodes occurred on a weekend. Overall, 61.5% of patients arrived between 08:00-16:59; 63.1% of weekday episodes started between these times (Figure 1).

11244 episodes had an associated Emergency Department triage code, with 108 different triage codes used. The commonest triage problem was chest pain (33.5% of episodes), see Supplementary Table 1. 6389 episodes (43.8%) had a length of stay of less than 12 hours.

Normal working day arrivals

There were 7365 episodes in 6848 patients with an arrival time between 08:00-16:59 on a weekday (normal working day, NWD). The triage problem was available for 5272 NWD episodes (71.6%). The commonest triage problem was chest pain (36.8%) (Supplementary Table 1).

4053 episodes (55.0%) had a length of stay of less than 12 hours and 3312 (45.0%) were discharged after 12 to 48 hours. Patients arriving in NWD hours were more likely to be discharged within 12 hours than those arriving outside of these hours (55.0% vs 33.7%, Chi square $p<0.005$).

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

Compared to patients discharged within 12 to 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, Chi square for all).

Factors affecting likelihood of discharge within 12 hours

Univariate comparison of the variables assessed within the original Amb score and GAPS derivation in NWD admissions is shown in Table 3. Age ≥80 and anticipated need for IV therapy were associated with an increased risk of admission lasting more than 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 50-140bpm and systolic blood pressure between 100-200mmHg 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, those with a suspicion of ACS coded into the Emergency Department diagnosis were more likely to be admitted for >12 hours.

*Table 3: Factors considered in derivation of previous scoring systems. Univariate analysis, odds ratio for admission lasting 12-48 hours shown. IV: intravenous; ACS: Acute Coronary Syndrome; RR: respiratory rate; HR: heart rate; MEWS: Modified Early Warning Score; NEWS2: National Early Warning Score 2(9); IHD: ischaemic heart disease; GP: general practice. 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.*

N=7365 unless otherwise stated	Length of stay <12hrs Frequency (%)	Length of stay 12-48 hours Frequency (%)	Odds ratio (OR)	P value	95% CI OR
Age					
16-19	94 (2.3%)	78 (2.4%)	Ref		
20-29	392 (9.7%)	332 (10.0%)	1.02	0.904	0.73 to 1.43
30-39	477 (11.8%)	349 (10.5%)	0.88	0.455	0.63 to 1.23
40-49	548 (13.5%)	361 (10.9%)	0.79	0.168	0.57 to 1.10
50-59	746 (18.4%)	509 (15.4%)	0.82	0.232	0.60 to 1.13
60-69	641 (15.8%)	422 (12.7%)	0.79	0.162	0.57 to 1.10
70-79	634 (15.6%)	571 (17.2%)	1.09	0.617	0.79 to 1.50
80-89	437 (10.8%)	504 (15.2%)	1.39	0.048	1.00 to 1.93
90+	84 (2.1%)	186 (5.6%)	2.69	<0.0005	1.80 to 3.96
≥80	521 (12.9%)	690 (20.8%)	1.78	<0.0005	1.57 to 2.02
Sex (n= 7363)					
Male	1713 (42.3%)	1404 (42.4%)	1.00	0.912	0.92 to 1.10
IV treatment not anticipated	3953 (97.5%)	2704 (81.6%)	0.11	<0.0005	0.09 to 0.14
Not acutely confused (n=6745)	3526 (99.9%)	3197 (99.5%)	0.27	0.005	0.08 to 0.75
If chest pain (1940 pts), ACS not suspected	654 (57.0%)	410 (51.7%)	0.81	0.021	0.67 to 0.97
No neurological deficit*	4024 (99.3%)	3241 (97.9%)	0.33	<0.0005	0.21 to 0.51
Normal temperature (n=6743)	2524 (71.5%)	2242 (69.8%)	0.92	0.140	0.83 to 1.03
Normal RR (n=6735)	3437 (97.5%)	2994 (93.3%)	0.35	<0.0005	0.27 to 0.46
Normal O2 saturations (>95%) (n=6738)	2988 (84.7%)	2525 (78.7%)	0.67	<0.0005	0.59 to 0.76
HR 50-140 (n=6748)	3499 (99.0%)	3144 (97.9%)	0.49	<0.0005	0.32 to 0.74
Systolic blood pressure 100-200 (n=6753)	3430 (96.9%)	3040 (94.6%)	0.56	<0.0005	0.43 to 0.71
MEWS 0 (n=6764)	132 (3.7%)	116 (3.6%)	0.97	0.804	0.74 to 1.26
NEWS2 0 (n=6712)	1381 (39.4%)	1012 (31.6%)	0.71	<0.0005	0.64 to 0.79
NEWS2 0-2 (n=6712)	3213 (91.7%)	2598 (81.0%)	0.39	<0.0005	0.33 to 0.45
NEWS2 (n=6712)					
0	1381 (39.4%)	1012 (31.6%)	ref		
1	1332 (38.0%)	1103 (34.4%)	1.13	0.0352	1.01 to 1.27
2	500 (14.3%)	483 (15.1%)	1.32	<0.0005	1.14 to 1.53
3	188 (5.4%)	272 (8.5%)	1.97	<0.0005	1.61 to 2.42
4	71 (2.0%)	132 (4.1%)	2.54	<0.0005	1.88 to 3.42
5	21 (0.6%)	91 (2.8%)	5.91	<0.0005	3.65 to 9.57
6+	12 (0.3%)	114 (3.6%)	12.96	<0.0005	7.11 to 23.63
Alert (n=6745)	3524 (99.8%)	3170 (98.6%)	0.14	<0.0005	0.05 to 0.32
Not discharged within previous 30 days	3518 (86.8%)	2884 (87.1%)	1.02	0.725	0.89 to 1.18
Admitted within last 1 year	1543 (38.1%)	1499 (45.3%)	1.34	<0.0005	1.22 to 1.48

No history of IHD	3116	(76.9%)	2446	(73.9%)	0.85	0.003	0.76 to 0.95
No history of heart failure	3925	(96.8%)	3113	(94.0%)	0.51	<0.0005	0.40 to 0.64
No history of arrhythmia	3689	(91.0%)	2787	(84.2%)	0.52	<0.0005	0.45 to 0.61
No history of diabetes	3476	(85.8%)	2667	(80.5%)	0.69	<0.0005	0.61 to 0.78
No history of stroke	4033	(99.5%)	3229	(97.5%)	0.19	<0.0005	0.11 to 0.32
No history of renal disease	3866	(95.4%)	3064	(92.5%)	0.60	<0.0005	0.49 to 0.73
No history of chronic lung disease	3264	(80.5%)	2530	(76.4%)	0.78	<0.0005	0.70 to 0.88
Arrival by ambulance	1080	(26.7%)	1384	(41.8%)	1.97	<0.0005	1.79 to 2.18
Referred by GP	2111	(52.1%)	908	(27.4%)	0.35	<0.0005	0.31 to 0.38
Triage category (n=5272)							
Standard	264	(11.2%)	220	(7.6%)	Ref		
Urgent	2072	(87.7%)	2427	(83.4%)	1.41	<0.0005	1.16 to 1.70
Resuscitation	27	(1.1%)	262	(9.0%)	11.6	<0.0005	7.54 to 18.0

Amb score

Multivariate analysis including all components of the Amb score, except access to transportation (which was present for all patients), is shown in Supplementary Table 2. Male sex, absence of acute confusion, normal MEWS and no recent hospital admission did not predict likelihood of discharge within 12 hours in this multivariate analysis. Replacing MEWs with the currently used NEWS2 acuity score, there remained no association of male sex, absence of acute confusion, and no recent hospital admission 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 (Supplementary Table 3). 93.8% (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 less than 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; 12.2% 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%), with a positive predictive value of 55.0% (95% CI 53.8% to 56.2%) and negative predictive value of 87.8% (95% CI 84.3% to 90.8%). 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 to 48 hours).

Table 4: Amb score performance. Performance in normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9)

	Amb score N=6743		Amb score with NEWS2 N=6707	
	Frequency (%)		Frequency (%)	
Score <5 5+	 418 6325	 (6.2%) (93.8%)	 364 6343	 (5.4%) (94.6%)
Score <5 Admission length <12hrs	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				
Sensitivity	98.6% (95% CI 98.1% to 98.9%)		98.8% (95% CI 98.4% to 99.1%)	
Specificity	11.4% (95% CI 10.3% to 12.6%)		10.0% (95% CI 9.0% to 11.1%)	
PPV	55.0% (95% CI 53.8% to 56.2%)		54.5% (95% CI 53.3% to 55.8%)	
NPV	87.8% (95% CI 84.3% to 90.8%)		88.5% (95% CI 84.7% to 91.6%)	
% of patients discharged in <12 hours not identified by score	1.4% (95% CI 1.1% to 2%)		1.2% (95% CI 0.9% to 1.6%)	
Patients identified as suitable by score admitted for >12 hours	45.0% (95% CI 43.8% to 46.2%)		45.5% (95% CI 44.2% to 46.7%)	

Replacing MEWS with NEWS2, the AUROC was 0.612 (95% CI 0.599 to 0.625)(Figure 2b). 94.6% (6343 admissions) had an Amb score of 5 or more; 5.4% (364 admissions) had a score of less than 5. Of those with a raised Amb score incorporating NEWS2, 54.5% were discharged within 12 hours of arrival; 11.5% 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%), with a positive predictive value of 54.5% (95% CI 53.8% to 56.2%) and NPV of 88.5% (95% CI 84.7% to 91.6%). Overall, 56.4% 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%, Chi square p=0.017), in both those discharged within 12 hours (13.7% vs 5.8%, p=0.017) and those discharged in 12 to 48 hours (11.7% vs 7.0%, p=0.001). This was also true for readmission within 30 days (25.6% vs 13.6%, p<0.0005), in those discharged within 12 hours (23.5% vs 12.2%, p=0.015) and those discharged in 12 to 48 hours (25.9 vs 15.3%, p<0.0005). This difference remained when substituting in NEWS2 (7 days: 12.1% vs 6.4%, p<0.0005; 30 days: 25.3% vs 13.8%, p<0.005).

Impact on patient pathway

Patient pathways through acute care incorporating the Amb score were estimated (Figure 3a). Directing short 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 (Supplementary 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, and where comorbidity due to ischaemic heart disease, heart failure, arrhythmia, diabetes, stroke/TIA, renal disease or chronic lung disease was present. A higher proportion of GP referrals with a raised Amb score were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (68.5% vs 44.7%, Chi square $p<0.005$). A higher proportion of patients with a raised Amb score and a NEWS2 of 0-2 were identified correctly compared to those with a raised NEWS2 on arrival.

GAPS

Multivariate analysis including all components of the GAPS is shown in Supplementary 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 more than 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 (Supplementary 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 (57.2%) had a GAPS >15, suggesting need for admission. Of those with a GAPS of 15 or less, 51.4% (1121 episodes) were discharged within 12 hours (Table 5). The sensitivity of the GAPS for identifying patients discharged within 12 hours was 50.4% (95% CI 48.5% to 52.5%), with a PPV of 51.4% (95% CI 49.3% to 53.6%) and 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 to 48 hours).

Table 5: GAPS performance within normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9)

	GAPS N=5091	GAPS with NEWS2 N=4953
	Frequency (%)	Frequency (%)
Score ≤15	2179 (42.8%)	2101 (42.4%)
Score 16+	2912 (57.2%)	2852 (57.6%)
Score ≤15 Admission length <12hrs	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		
Sensitivity	50.4% (95% CI 48.5 to 52.5%)	50.0% (95% CI 47.8% to 52.1%)
Specificity	63.1% (95% CI 61.3% to 64.9%)	63.3% (95% CI 61.5% to 65.0%)
PPV	51.4% (95% CI 49.3% to 53.6%)	50.5% (95% CI 48.4% to 52.7%)
NPV	62.1% (95% CI 60.3% to 63.9%)	62.7% (95% CI 60.9% to 64.5%)
% of patients discharged in <12 hours not identified by score	49.6% (95% CI 47.5% to 51.5%)	50.0% (95% CI 47.9% to 52.2%)
Patients identified as suitable by score admitted for >12 hours	48.6% (95% CI 46.4% to 50.7%)	49.5% (95% CI 47.3% to 51.6%)

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. The sensitivity of the GAPS for identifying patients discharged within 12 hours was 50.0% (95% CI 47.8% to 52.1%), with a PPV of 50.5% (95% CI 48.4% to 52.7%) and 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).

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 to 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%, Chi square p<0.005), both for those discharged within 12 hours (6.0% vs 4.2%, p=0.055), and 12 to 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 to 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).

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, 49.5% 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, 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. 29 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, were female, and where comorbidity due to stroke/TIA was present (Supplementary Table 7). A higher proportion of GP referrals with a low GAPS were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (67.6% vs 50.3%, Chi square p=0.044). A higher proportion of patients with a low GAPS and a NEWS2 of 0-2 were identified correctly compared to 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 (47.1%) where the scoring systems agreed. In 2048 episodes (41.4%) both scores suggested the patient was suitable for SDEC (Amb score 5+ and GAPS ≤15) and in 284 episodes (5.7%) both scores suggested the patient was likely to require admission (Amb score <5 and GAPS 16+). In 2620 episodes (52.9%) the recommendation provided by the score differed. There were 2567 episodes (51.8%) where the Amb score suggested suitability for SDEC while the GAPS suggested admission was likely and 53 episodes (1.1%) where the GAPS suggested likely discharge but the Amb score predicted admission. Those aged over 70, 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 a Amb score suggesting suitability for SDEC with a GAPS suggesting admission (Chi square, p<0.0005 for each subgroup comparison, Figure 4).

Discussion

This paper highlights several important points. Firstly, 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 to 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 negative predictive value than the GAPS, with 88.5% of patients with a low Amb score (suggesting they were unsuitable for SDEC) remaining for more than 12 hours, compared to 62.7% of those with a high GAPS.

Second, some components of both scores included as factors to predict admission or discharge were non-discriminatory in this patient cohort. Multivariate 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 multivariate analysis of GAPS components, and within univariate analysis, referral from GP was associated with decreased likelihood of admission for over 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 significant divergence in the patients identified for SDEC by the Amb score and GAPS. Conflicting recommendations were more likely in those aged over 70, 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 more than 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 focussed on performance of scoring systems to identify patients suitable for SDEC within currently available services; in-depth evaluation of factors necessitating admission over 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 focussed on the ability of the Amb score and GAPS to discriminate between those admitted for <12 hours and 12 to 48 hours. Applying the Amb score or GAPS across all medical admissions will affect the positive and negative predictive value of the score. If some patients with a length of stay >48 hours have a raised Amb score or low GAPS, then the positive predictive value 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 alternative cut-offs was not performed, as multivariate analysis suggested components of the score were not performing as expected within this patient cohort.

This analysis used retrospective data. Amb score calculation presumed IV treatment to be 'anticipated' in patients receiving IV treatment within 6 hours of arrival, as anticipation of IV 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 was 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.

Conclusion

Within this patient cohort, the Amb score and Glasgow Admission Prediction Score 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 Same Day Emergency Care services.

Contributorship

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.

License for publication

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. E.S 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.

Data sharing agreement

Data from this study is 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.

Ethics statement

This research was performed in accordance with the Declaration of Helsinki. All study processes were carried out following appropriate ethical approval provided by the East Midlands – Derby REC (reference: 20/EM/0158). Formal written consent from individual participants was not required.

Transparency declaration

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Funding

No specific funding was available for this project.

References

1. NHS England. A&E Attendances and Emergency Admissions 2021-22 2021 [Available from: 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? Clinical Medicine. 2022:clinmed.2021-0614.
5. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
6. Royal College of Physicians. Acute Care Toolkit 10: Ambulatory Emergency Care. 2014.
7. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.
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, Holland M, Subbe C, Kennedy A, et al. Length of stay in Acute Medical Admissions: Analysis from the Society for Acute Medicine Benchmarking Audit. Acute Med. 2022;21(1):27-33.
11. Cameron A, Jones D, Logan E, O'Keeffe CA, Mason SM, Lowe DJ. Comparison of Glasgow Admission Prediction Score and Amb Score in predicting need for inpatient care. Emerg Med J. 2018;35(4):247-51.
12. Pendlebury ST, Lovett NG, Smith SC, Dutta N, Bendon C, Lloyd-Lavery A, 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(11):e007808.
13. Society for Acute Medicine. Society for Acute Medicine Benchmarking Audit 2021 - SAMBA2021 Report. 2021.
14. Atkin C, Knight T, Subbe C, Holland M, Cooksley T, Lasserson D. Acute care service performance during winter: report from the winter SAMBA 2020 national audit of acute care. Acute Med. 2020;19(4):220-9.

Figure Legends

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

Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions.

Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, 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.

Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

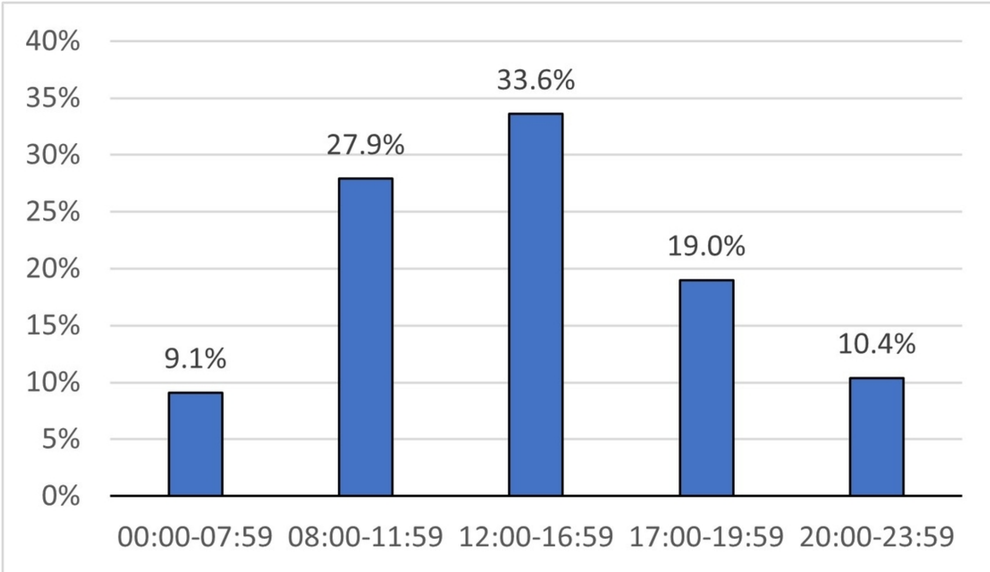


Figure 1: Arrival time for medical attendances lasting up to 48 hours.

99x57mm (300 x 300 DPI)

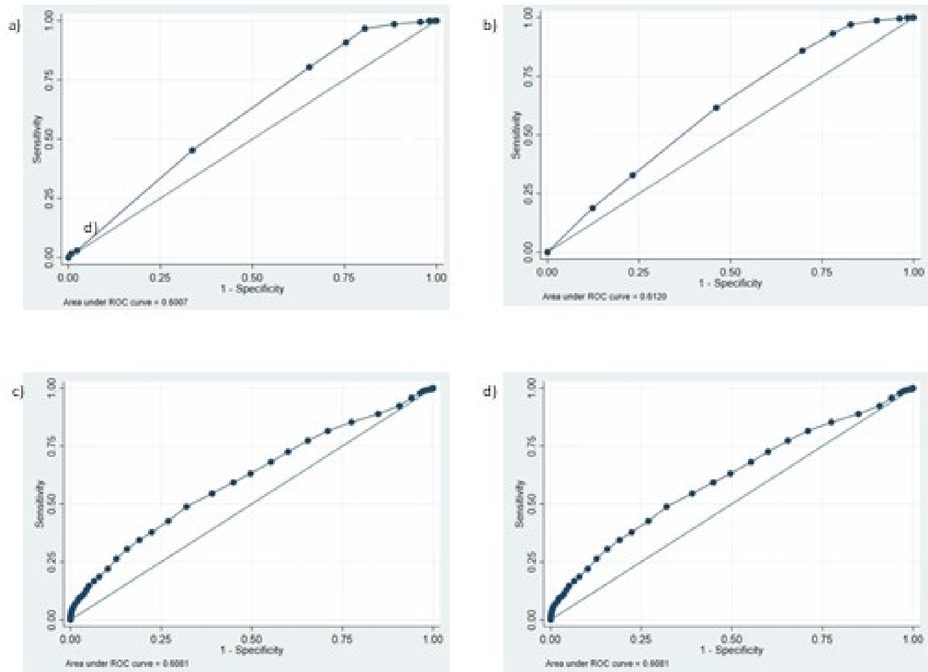


Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions.

143x120mm (96 x 96 DPI)

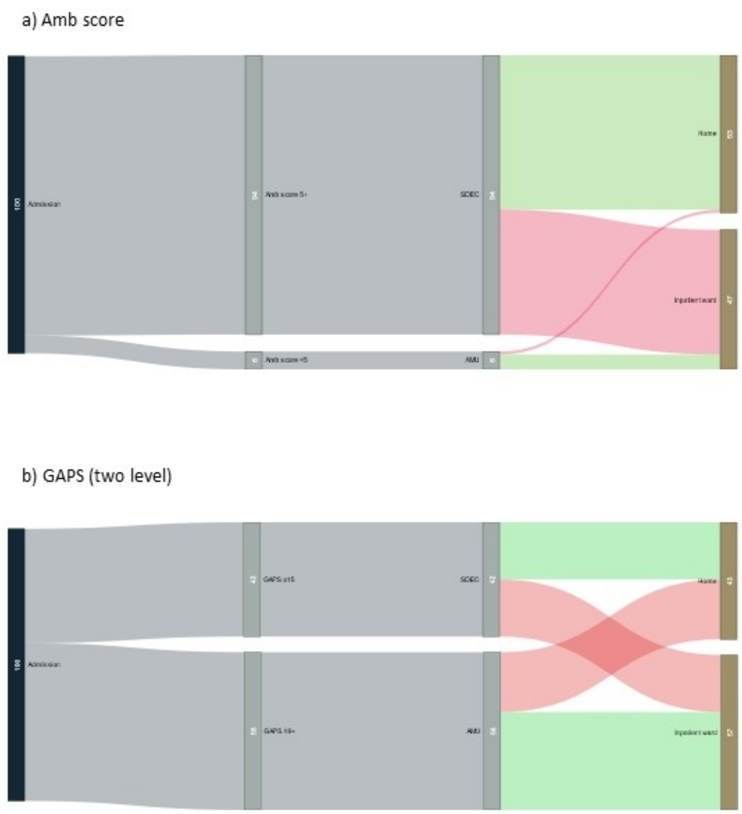


Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, for a) Amb score (5 or more) and b) Glasgow Admission Prediction Score (GAPS)(≤15). Green = currently identified by scoring system, red = incorrectly identified by scoring system.

190x188mm (96 x 96 DPI)

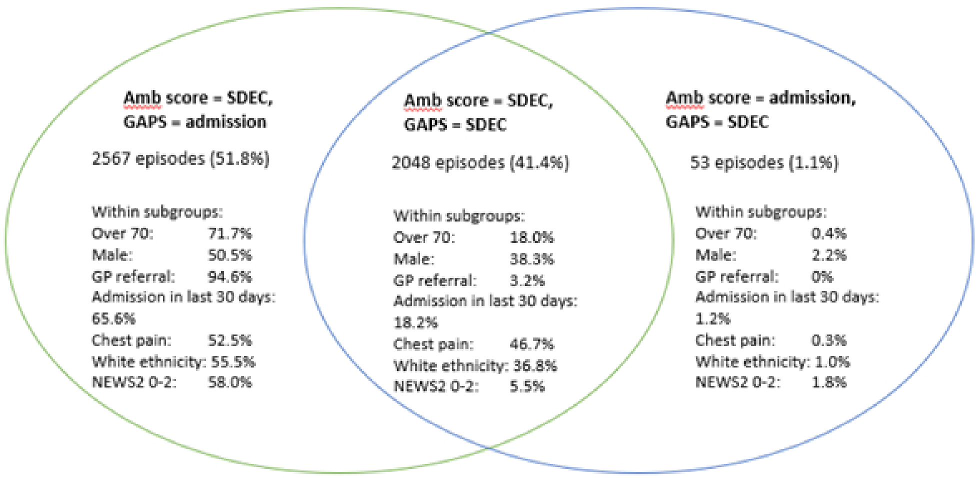


Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

402x212mm (38 x 38 DPI)

Supplementary Table 1: Triage problem. Commonest triage problem recorded on arrival to Emergency Department. Coded presenting problem entered at initial Emergency Department triage. Normal working day admissions defined as episodes starting between 08:00-16:59 Monday-Friday.

All admissions			**Normal working day admissions**		
	Frequency (%)			Frequency (%)	
Chest pain	3762	(33.5%)	Chest pain	1940	(36.8%)
Dyspnoea/difficulty breathing	1586	(14.1%)	Dyspnoea/difficulty breathing	721	(13.7%)
Asthenia	1051	(9.4%)	Asthenia	548	(10.4%)
Headache	609	(5.4%)	Headache	322	(6.1%)
Abdominal pain	408	(3.6%)	Abdominal pain	172	(3.3%)
Near syncope/syncope	282	(2.5%)	Palpitations	145	(2.8%)
Palpitations	256	(2.3%)	Near syncope/syncope	137	(2.6%)
Dizziness	222	(2.0%)	Dizziness	119	(2.3%)
Fever	210	(1.9%)	Pain in lower limb	96	(1.8%)
Substance abuse	210	(1.9%)	Vomiting	82	(1.6%)

Supplementary Table 2: Multivariate analysis of Amb score components. Odds ratio for admission of 12-48 hours, normal working day admissions. IV= intravenous, MEWS= Modified Early Warning Score, NEWS2= National Early Warning Score 2.(2)

Amb score components				Amb score components, substituting NEWS2			
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age >80	1.86	<0.0005	1.63 to 2.13	Age >80	1.85	<0.0005	1.62 to 2.13
Male	1.03	0.568	0.93 to 1.14	Male	1.02	0.733	0.92 to 1.13
IV treatment not anticipated	0.12	<0.0005	0.10 to 0.15	IV treatment not anticipated	0.12	<0.0005	0.10 to 0.15
Not acutely confused	0.38	0.068	0.13 to 1.08	Not acutely confused	0.40	0.09	0.14 to 1.15
MEWS 0	1.05	0.739	0.80 to 1.38	NEWS2 0	0.82	<0.0005	0.74 to 0.92
Not discharged in last 30 days	1.00	0.993	0.86 to 1.16	Not discharged in last 30 days	1.00	0.907	0.87 to 1.17

Supplementary table 3: Amb score for NWD (Normal working day) admission episodes. Normal working day defined as episodes starting between 08:00-16:59 Monday-Friday. Amb score calculated as shown in Table 1.(1) NEWS2: National Early Warning Score 2.(2)

	Amb score		Amb score substituting NEWS2	
Amb score	Number of episodes (%)		Number of episodes (%)	
≤3	12	(0.2%)	12	(0.2%)
3.5	51	(0.8%)	44	(0.7%)
4	98	(1.5%)	81	(1.2%)
4.5	257	(3.8%)	227	(3.4%)
5	327	(4.9%)	287	(4.3%)
5.5	367	(5.4%)	295	(4.4%)
6	690	(10.2%)	522	(7.8%)
6.5	2261	(33.5%)	1605	(23.9%)
7	2502	(37.1%)	1735	(12.6%)
7.5	94	(1.4%)	846	(15.7%)
8	84	(1.3%)	1053	(12.3%)

Supplementary Table 4: Identifying length of admission by Amb score (incorporating NEWS2) within patient subgroups. Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Amb score calculated as per Table 1, with NEWS2 substituted in place of MEWS. NEWS2: National Early Warning Score 2.(2) MEWS: Modified Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for comparisons using Chi square.

	Amb 5+, Admission length <12hrs Correctly identified		Amb 5+, Admission length 12-48 hrs Incorrectly identified		Amb <5, Admission length <12 hours Incorrectly identified		Amb <5, Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by Amb score discharged within 12 hours	P value
Percentage of admissions	51.6%		43.0%		0.6%		4.8%		55.0%	
	N	%	N	%	N	%	N	%		
Age										
16-19	85	52.5%	70	43.2%	<10	<6.2%	<10	<6.2%	54.8%	<0.005
20-29	340	51.1%	291	43.8%	<10	<1.5%	28	4.2%	53.9%	
30-39	404	54.0%	310	41.4%	<10	<1.3%	27	3.6%	56.6%	
40-49	465	56.8%	330	40.3%	<10	<1.2%	20	2.4%	58.5%	
50-59	630	56.0%	445	39.6%	12	1.1%	38	3.4%	58.6%	
60-69	564	57.9%	370	38.0%	<10	<1.0%	38	3.9%	60.4%	
70-79	547	49.5%	506	45.8%	<10	<0.9%	51	4.6%	51.9%	
80-89	357	41.4%	426	49.5%	<10	<1.2%	69	8.0%	45.6%	
90+	67	27.0%	136	54.8%	0	-	45	18.1%	33.0%	
Under 70	3035	54.2%	2322	41.5%	33	0.6%	208	3.7%	56.7%	<0.005
Over 70	424	38.2%	562	50.7%	<10	<0.9%	114	10.3%	43.0%	
Sex										
Female	2022	52.2%	1749	45.1%	12	0.3%	94	2.4%	53.6%	0.077
Male	1437	50.8%	1135	40.1%	30	1.1%	228	8.1%	55.9%	
Ethnicity										
Asian	500	51.4%	440	45.3%	<10	<1.0%	26	2.7%	53.2%	0.191
Black	169	56.0%	122	40.4%	<10	<3.3%	10	3.3%	58.1%	
Unknown	395	54.0%	287	39.3%	11	1.5%	38	5.2%	57.9%	
Mixed	58	51.3%	48	42.5%	<10	<8.8%	<10	<8.8%	54.7%	
Other	103	57.9%	72	40.4%	0		<10	<5.6%	58.9%	
White	2234	50.6%	1915	43.4%	23	1.0%	239	5.4%	53.8%	
Recent admission (30 days)										
Yes	433	50.3%	335	39.0%	11	1.3%	81	9.4%	56.4%	0.273
No	3026	51.8%	2549	43.6%	31	0.5%	241	4.1%	54.3%	
GP referral										
Yes	1792	67.3%	823	30.9%	10	0.4%	39	1.5%	68.5%	<0.005
No	1667	41.2%	2061	51.0%	32	0.8%	283	7.0%	44.7%	
Chest pain as triage problem										
Yes	1032	57.7%	739	41.3%	<10	<0.6%	12	0.7%	58.3%	<0.005
No	2427	49.4%	2145	43.6%	35	0.7%	310	6.3%	53.1%	
History of IHD										
Yes	834	49.7%	766	45.6%	<10	<0.6%	69	4.1%	52.1%	0.025
No	2625	52.2%	2118	42.1%	33	0.7%	253	5.0%	55.3%	
History of HF										
Yes	111	36.0%	167	54.2%	<10	<3.2%	27	8.8%	39.9%	<0.005
No	3348	52.3%	2717	42.5%	39	0.6%	295	4.6%	55.2%	
History of arrhythmia										
Yes	323	37.9%	438	51.3%	<10	<1.2%	83	9.7%	42.4%	<0.005
No	3136	53.6%	2446	41.8%	33	0.6%	239	4.1%	56.2%	
History of diabetes										
Yes	497	44.0%	546	48.4%	<10	<0.9%	79	7.0%	47.7%	<0.005
No	2962	53.1%	2338	41.9%	35	0.6%	243	4.4%	55.9%	
History of stroke										
Yes	18	17.8%	80	79.2%	0	-	<10	<10.0%	18.4%	<0.005
No	3441	52.1%	2804	42.4%	42	0.6%	319	4.8%	55.1%	
History of renal disease										
Yes	167	40.7%	197	48.0%	0	-	46	11.2%	45.9%	<0.005
No	3292	52.3%	2687	42.7%	42	0.7%	276	4.4%	55.1%	
History of chronic lung disease										
Yes	703	47.5%	674	45.5%	12	0.8%	92	6.2%	51.5%	<0.005
No	2756	52.7%	2210	42.3%	32	0.6%	230	4.4%	55.5%	
NEWS2										
0-2	3180	54.8%	2435	41.9%	29	0.5%	162	2.8%	56.6%	<0.005
3-4	252	38.0%	319	48.1%	<10	<1.5%	85	12.8%	44.1%	
5+	27	11.3%	130	54.6%	<10	<4.2%	75	31.5%	17.2%	

Supplementary Table 5: Multivariate analysis of GAPS components. Age – odds ratio (OR) per decade increase in age; NEWS/NEWS2 OR per increase of one point in NEWS/NEWS2. Triage category compared to 'standard' as reference. Odds ratio for admission of 12-48 hours, normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). GP= general practitioner, NEWS= national early warning score

	GAPS				GAPS with NEWS2		
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age	1.06	<0.0005	1.03 to 1.10	Age	1.06	<0.0005	1.03 to 1.10
NEWS	1.24	<0.0005	1.12 to 1.31	NEWS2	1.22	<0.0005	1.16 to 1.28
Triage category*				Triage category*			
Urgent	1.08	0.457	0.88 to 1.32	Urgent	1.04	0.692	0.85 to 1.29
Resuscitation	4.56	<0.0005	2.89 to 7.12	Resuscitation	4.26	<0.0005	2.69 to 6.74
Referred by GP	0.80	<0.0005	0.69 to 0.91	Referred by GP	0.78	0.001	0.68 to 0.90
Arrived in ambulance	1.61	<0.0005	1.41 to 1.83	Arrived in ambulance	1.60	<0.0005	1.40 to 1.82
Admitted <1 year ago	1.41	<0.0005	1.25 to 1.59	Admitted <1 year ago	1.39	<0.0005	1.23 to 1.57

Supplementary Table 6: GAPS for normal working day admissions. GAPS: Glasgow Admission Prediction Score, calculated as described in Table 1.(3) NEWS2: National Early Warning Score 2.(2)

	GAPS score N=5091		GAPS score substituting NEWS2 N=4953	
GAPS score	Number of episodes (%)		Number of episodes (%)	
1-5	93	(1.8%)	88	(1.8%)
6-19	829	(16.3%)	792	(16.0%)
11-15	1257	(24.7%)	1221	(24.7%)
16-20	1329	(26.1%)	1279	(15.8%)
21-25	874	(17.2%)	857	(17.3%)
26-30	354	(7.0%)	360	(7.3%)
31-35	211	(4.1%)	206	(4.2%)
36-40	97	(1.9%)	94	(1.9%)
41-45	41	(0.8%)	45	(0.9%)
46+	<10	(<0.2%)	11	(0.2%)

Supplementary Table 7: Identifying length of admission by GAPS (incorporating NEWS2) within patient subgroups. Analysis of Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Glasgow Admission Prediction Score (GAPS) calculated as per Table 1, with NEWS2 substituted in place of NEWS. NEWS2: National Early Warning Score 2.(2) NEWS: National Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for Chi square comparisons.

GAPS with NEWS2	GAPS≤15 Admission length <12hrs Correctly identified		GAPS ≤15 Admission length 12-48 hrs Incorrectly identified		GAPS 16+, Admission length <12 hours Incorrectly identified		GAPS 16+ Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by GAPS discharged within 12 hours	P value
Percentage of admissions	21.4%		21.0%		21.5%		36.1%		50.5%	
Age (years)										
16-19	32	26.4%	48	39.7%	17	14.0%	24	19.8%	40.0%	<0.005
20-29	140	26.7%	180	34.3%	91	17.4%	113	21.6%	43.8%	
30-39	172	29.6%	185	31.8%	101	17.4%	123	21.2%	48.2%	
40-49	228	35.3%	178	27.6%	104	16.1%	135	20.9%	56.2%	
50-59	237	27.8%	191	22.4%	188	22.1%	235	27.6%	55.4%	
60-69	126	18.1%	102	14.6%	216	31.0%	253	36.3%	55.3%	
70-79	87	11.4%	89	11.7%	201	26.4%	385	50.5%	49.4%	
80-89	33	5.7%	59	10.3%	122	21.2%	361	62.8%	35.9%	
90+	<10	<5.1%	<10	<5.1%	23	11.7%	160	81.2%	50.0%	
Under	1022	24.4%	973	23.3%	918	22.0%	1268	30.3%	51.2%	0.007
Over 70	40	5.2%	66	8.5%	145	18.8%	521	67.5%	37.7%	
Sex										
Female	599	21.3%	633	22.6%	597	21.3%	977	34.8%	48.6%	0.035
Male	463	21.6%	406	18.9%	466	21.7%	811	37.8%	53.3%	
Ethnicity										
Asian	223	27.6%	188	23.2%	157	19.4%	241	29.8%	54.3%	0.25
Black	57	26.1%	48	22.0%	48	22.0%	65	29.8%	54.3%	
Unknown	135	25.7%	127	24.2%	102	19.4%	161	30.7%	51.5%	
Mixed	20	22.0%	27	29.7%	20	22.0%	24	26.4%	42.6%	
Other	35	24.5%	47	32.9%	37	25.9%	24	16.8%	42.7%	
White	592	18.7%	602	19.0%	699	22.1%	1274	40.2%	49.6%	
Recent admission (30 days)										
Yes	45	8.7%	55	10.6%	122	23.6%	295	57.1%	45.0%	0.256
No	1017	22.9%	984	22.2%	941	21.2%	1494	33.7%	50.8%	
GP referral										
Yes	23	2.1%	11	1.0%	533	49.6%	508	47.3%	67.6%	0.044
No	1039	26.8%	1028	26.5%	530	13.7%	1281	33.0%	50.3%	
Chest pain as triage problem										
Yes	523	29.2%	318	17.8%	516	28.8%	433	24.2%	62.2%	<0.005
No	539	17.0%	721	22.8%	547	17.3%	1356	42.9%	42.8%	
History of IHD										
Yes	306	20.3%	231	15.3%	402	26.7%	568	37.7%	57.0%	<0.005
No	756	21.9%	808	23.4%	661	19.2%	1221	35.4%	48.3%	
History of heart failure										
Yes	20	8.0%	17	6.8%	51	20.5%	161	64.7%	54.1%	0.667
No	1042	22.2%	1022	21.7%	1012	21.5%	1628	34.6%	50.5%	
History of arrhythmia										
Yes	72	10.2%	78	11.0%	156	22.1%	401	56.7%	48.0%	0.517
No	990	23.3%	961	22.6%	907	21.4%	1388	32.7%	50.7%	
History of diabetes										
Yes	124	13.9%	146	16.3%	202	22.6%	421	47.1%	45.9%	0.104
No	938	23.1%	893	22.0%	861	21.2%	1368	33.7%	51.2%	
History of stroke										
Yes	<10	<10.4%	29	30.2%	<10	<10.4%	54	56.3%	17.1%	<0.005
No	1056	21.7%	1010	20.8%	1056	21.7%	1735	35.7%	51.1%	
History of renal disease										
Yes	26	8.7%	35	11.7%	61	20.3%	178	59.3%	42.6%	0.209
No	1036	22.3%	1004	21.6%	1002	21.5%	1611	34.6%	50.8%	
History of chronic lung disease										
Yes	191	16.8%	176	15.5%	262	23.0%	510	44.8%	52.0%	0.528
No	871	22.8%	863	22.6%	801	21.0%	1279	33.5%	50.2%	
NEWS2										
0-2	1002	33.0%	954	31.4%	952	31.3%	131	4.3%	51.2%	0.012
3-4	57	11.1%	72	14.0%	93	18.1%	291	56.7%	44.2%	
5+	<10	<4.6%	13	5.9%	18	8.2%	185	84.5%	18.8%	

References

1. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
2. Royal College of Physicians. National Early Warning Score (NEWS) 2. 2017.
3. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	Page 1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	Page 2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	Page 3
Objectives	3	State specific objectives, including any prespecified hypotheses	Page 4
Methods			
Study design	4	Present key elements of study design early in the paper	Page 5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Page 5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	Page 5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	Page 5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Page 5-6
Bias	9	Describe any efforts to address potential sources of bias	Page 5-6, 18-19
Study size	10	Explain how the study size was arrived at	Page 5

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Page 6
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	Page 6-7
		(*b*) Describe any methods used to examine subgroups and interactions	Page 5-7
		(*c*) Explain how missing data were addressed	Page 5-7
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	Page 6
		(*e*) Describe any sensitivity analyses	Page 6
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Page 7, 8, 9
		(b) Give reasons for non-participation at each stage	Page 7, 8, 9
		(c) Consider use of a flow diagram	n/a
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Page 8
		(b) Indicate number of participants with missing data for each variable of interest	Page 8
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	n/a
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	Page 11
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Page 12
		(*b*) Report category boundaries when continuous variables were categorized	n/a
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Page 14, page 16
Discussion			
Key results	18	Summarise key results with reference to study objectives	Page 17
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Page 18
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Page 17
Generalisability	21	Discuss the generalisability (external validity) of the study results	Page 17-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Page 20

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064910.R1
Article Type:	Original research
Date Submitted by the Author:	07-Sep-2022
Complete List of Authors:	Atkin, Catherine; University of Birmingham, Birmingham Acute Care Research Group Gallier, Suzy; University Hospitals Birmingham NHS Foundation Trust, Department of Health Informatics Wallin, Elizabeth; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine Reddy-Kolanu, Vinay; University Hospitals Birmingham NHS Foundation Trust, Acute medicine Sapey, Elizabeth; University of Birmingham, PIONEER HDR-UK Hub; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Evidence based practice
Keywords:	INTERNAL MEDICINE, GENERAL MEDICINE (see Internal Medicine), Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

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 (0000-0003-0596-8515), Suzy Gallier (0000-0003-1026-4125), Elizabeth Wallin, Vinay Reddy-Kolanu, Elizabeth Sapey (0000-0003-3454-5482)

Addresses

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Edgbaston, Birmingham, B15 2GW, UK. Catherine Atkin, NIHR Academic Clinical Lecturer in Acute Medicine.

Department of Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Suzy Gallier, Lead for Research Analytics.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Wallin, Consultant in Nephrology & Acute Medicine.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Vinay Reddy-Kolanu, Consultant in Acute Medicine.

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, and Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Sapey, Professor of Acute and Respiratory Medicine and Honorary Consultant.

Correspondence to: 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 and lasting up to 48 hours, between 1st April 2019 and 9th 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 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

- This study compared performance of the Amb score and GAPS 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 NEWS2, reflecting current clinical practice
- Patients admitted for longer than 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 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 less than 12 hours or over 48 hours. The study excluded patients who remained in hospital for 12 to 48 hours.

Table 1: Scoring systems to identify medical admissions potentially suitable for discharge from hospital without admission >12 hours. Amb score(5) and Glasgow Admission Prediction Score (GAPS)(7). Amb score of 5 more indicates likely discharge within 12 hours; GAPS of 16 or more suggests patient likely to be admitted to hospital. IV = intravenous, MEWS = Modified early warning score, NEWS = National Early Warning Score, GP = General practitioner

Amb score		
Sex	Female	0
	Male	-0.5
Age	<80	0
	≥80	-0.5
Access to personal transport/can take public transport	Agree	2
	Disagree	0
IV treatment not anticipated	Agree	2
	Disagree	0
Not acutely confused	Agree	2
	Disagree	0
MEWS=0	Agree	1
	Disagree	0
Not discharged from hospital within previous 30 days	Agree	1
	Disagree	0

Glasgow Admissions Prediction Score (GAPS)		
NEWS		1 point per point on NEWS score
Age		1 point per decade
Triage category	3 2 (or 2+) 1	5 10 20
Referred by GP		5
Arrived in ambulance		5
Admitted <1 year ago		5

The Glasgow Admission Prediction Score (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 Same Day Emergency Care and those requiring admission for at least 12 to 48 hours.

Methods

This data study was conducted in collaboration with PIONEER, a Health Data Research Hub in Acute Care, and all study processes were carried out following appropriate ethical approval provided by the East Midlands – Derby REC (reference: 20/EM/0158).

Retrospective data were collected for patients admitted to Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Trust (UHB) between the period of the 1st April 2019 until 9th 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 (CCU) in Europe. The Acute Medical Unit (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 inhouse electronic health record (EHR) called Prescribing Information and Communication System (PICS). 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 focussed 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 and mapped ICD10 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 to 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 08:00-16:59, 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 (IV) treatment was taken as not being anticipated where patients did not receive an IV 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 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 National Early Warning Score 2 (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 15.1. Cell counts containing less than 10 patients were suppressed, due to reporting requirements. For univariate analysis of factors influencing likelihood of discharge within 12 hours, odds ratios for variables included in the original Amb score or GAPS derivation studies were assessed using Chi square. Multivariable analysis of the Amb score and GAPS components was performed using logistic regression, to demonstrate the performance of components within the score and allow 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 ROC (AUROC) calculated. Subgroup analysis was performed in prespecified groups based on previous research.(10) Comparison of proportions was performed using Chi square. A p value of <0.05 is 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 less than 48 hours, based on previous research.(10)

Results

14314 acute medical inpatient episodes lasting up to 48 hours were identified during the censor period. These episodes were from 12587 patients with 11229 patients having one episode in this time period. Patients 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.

Table 2: Demographics and characteristics of patients with emergency medical admissions lasting up to 48 hours. For whole cohort, and for patients arriving in a normal working day (08:00-16:59, Monday to Friday).P values shown for Chi square comparison of normal working day episodes to episodes starting outside normal working day.

	All episodes N=14314		Normal working day episodes N=7365		Episodes starting outside normal working day N= 6949		P value
	Frequency (%)		Frequency (%)		Frequency (%)		
Age							
16-19	444	(3%)	172	(2%)	272	(4%)	<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%)	270	(4%)	303	(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	(5%)	332	(5%)	323	(5%)	
Unknown	1623	(11%)	816	(11%)	807	(12%)	
Mixed	260	(2%)	124	(2%)	136	(2%)	
Other	403	(3%)	199	(3%)	204	(3%)	
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 day	1047	(7%)	479	(7%)	568	(8%)	<0.001
14 day	1544	(11%)	681	(9%)	863	(12%)	<0.001
30 day	2268	(16%)	1033	(14%)	1235	(18%)	<0.001

18.4% of episodes occurred on a weekend. Overall, 61.5% of patients arrived between 08:00-16:59 (Figure 1); 63.1% of weekday episodes started between these times.

11244 episodes had an associated Emergency Department triage code, with 108 different triage codes used. The commonest triage problem was chest pain (33.5% of episodes), see Supplementary Table 1. 6394 episodes (43.8%) had a length of stay of less than 12 hours.

Normal working day arrivals

There were 7365 episodes in 6848 patients with an arrival time between 08:00-16:59 on a weekday (normal working day, NWD). The triage problem was available for 5272 NWD episodes (72%). The commonest triage problem was chest pain (37%) (Supplementary Table 1).

4053 episodes (55%) had a length of stay of less than 12 hours and 3312 (45%) were discharged after 12 to 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%, Chi square $p<0.005$).

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

Compared to patients discharged within 12 to 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%, $p<0.005$, Chi square 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 and anticipated need for IV therapy were associated with an increased risk of admission lasting more than 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 50-140bpm and systolic blood pressure between 100-200mmHg 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 ACS coded into the Emergency Department diagnosis were more likely to be admitted for >12 hours (OR 0.81, $p=0.021$, 95% CI 0.67 to 0.97).

*Table 3: Factors considered in derivation of previous scoring systems. Column percentages shown. Univariate analysis, odds ratio for admission lasting 12-48 hours shown. IV: intravenous; RR: respiratory rate; HR: heart rate in beats per minute; SBP: systolic blood pressure in mmHg; MEWS: Modified Early Warning Score; NEWS2: National Early Warning Score 2(9); IHD: ischaemic heart disease; GP: general practice. 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.*

N=7365 unless otherwise stated	Length of stay <12hrs Frequency (%)		Length of stay 12-48 hours Frequency (%)		Odds ratio (OR)	P value	95% CI OR
Age							
16-19	94	(2%)	78	(2%)	Ref		
20-29	392	(10%)	332	(10.%)	1.02	0.90	0.73 to 1.43
30-39	477	(12%)	349	(11%)	0.88	0.46	0.63 to 1.23
40-49	548	(14%)	361	(11%)	0.79	0.17	0.57 to 1.10
50-59	746	(18%)	509	(15%)	0.82	0.23	0.60 to 1.13
60-69	641	(16%)	422	(13%)	0.79	0.16	0.57 to 1.10
70-79	634	(16%)	571	(17%)	1.09	0.62	0.79 to 1.50
80-89	437	(11%)	504	(15%)	1.39	0.05	1.00 to 1.93
90+	84	(2%)	186	(6%)	2.69	<0.001	1.80 to 3.96
≥80	521	(13%)	690	(21%)	1.78	<0.001	1.57 to 2.02
Sex *(n= 7363)*							
Male	1713	(42%)	1404	(42%)	1.00	0.91	0.92 to 1.10
IV treatment not anticipated	3953	(98%)	2704	(82%)	0.11	<0.001	0.09 to 0.14
Not discharged in previous 30 days	3518	(87%)	2884	(87%)	1.02	0.73	0.89 to 1.18
Not admitted within last 1 year	2510	(62%)	1813	(55%)	0.74	<0.001	0.68 to 0.82
No neurological deficit*	4024	(99.3%)	3241	(97.9%)	0.33	<0.001	0.21 to 0.51
Not acutely confused *(n=6745)*	3526	(99.9%)	3197	(99.5%)	0.27	0.005	0.08 to 0.75
Physiological observations							
Normal temperature *(n=6743)*	2524	(72%)	2242	(70%)	0.92	0.14	0.83 to 1.03
Normal RR *(n=6735)*	3437	(98%)	2994	(93%)	0.35	<0.001	0.27 to 0.46
O_2 saturations >95% *(n=6738)*	2988	(85%)	2525	(79%)	0.67	<0.001	0.59 to 0.76
Heart rate 50-140 *(n=6748)*	3499	(99.0%)	3144	(97.9%)	0.49	<0.001	0.32 to 0.74
SBP 100-200 *(n=6753)*	3430	(96.9%)	3040	(94.6%)	0.56	<0.001	0.43 to 0.71
Alert *(n=6745)*	3524	(99.8%)	3170	(98.6%)	0.14	<0.001	0.05 to 0.32
MEWS 0 *(n=6764)*	132	(4%)	116	(4%)	0.97	0.80	0.74 to 1.26
NEWS2 0 *(n=6712)*	1381	(39%)	1012	(32%)	0.71	<0.001	0.64 to 0.79
NEWS2 0-2 *(n=6712)*	3213	(92%)	2598	(81%)	0.39	<0.001	0.33 to 0.45
NEWS2 *(n=6712)*							
0	1381	(39%)	1012	(32%)	Ref		
1	1332	(38%)	1103	(34%)	1.13	0.04	1.01 to 1.27
2	500	(14%)	483	(15%)	1.32	<0.001	1.14 to 1.53
3	188	(5%)	272	(9%)	1.97	<0.001	1.61 to 2.42
4	71	(2%)	132	(4%)	2.54	<0.001	1.88 to 3.42
5	21	(0.6%)	91	(3%)	5.91	<0.001	3.65 to 9.57
≥6	12	(0.3%)	114	(4%)	12.96	<0.001	7.11 to 23.6
Previous medical history							
No history of IHD	3116	(77%)	2446	(74%)	0.85	0.003	0.76 to 0.95
No history of heart failure	3925	(97%)	3113	(94%)	0.51	<0.001	0.40 to 0.64
No history of arrhythmia	3689	(91%)	2787	(84%)	0.52	<0.001	0.45 to 0.61
No history of diabetes	3476	(86%)	2667	(81%)	0.69	<0.001	0.61 to 0.78
No history of stroke	4033	(99.5%)	3229	(97.5%)	0.19	<0.001	0.11 to 0.32
No history of renal disease	3866	(95%)	3064	(93%)	0.60	<0.001	0.49 to 0.73
No history of chronic lung disease	3264	(81%)	2530	(76%)	0.78	<0.001	0.70 to 0.88
Factors on arrival							
Arrival by ambulance	1080	(27%)	1384	(42%)	1.97	<0.001	1.79 to 2.18
Referred by GP	2111	(52%)	908	(27%)	0.35	<0.001	0.31 to 0.38
Triage category *(n=5272)*							
Standard	264	(11%)	220	(8%)	Ref		
Urgent	2072	(88%)	2427	(84%)	1.41	<0.001	1.16 to 1.70
Resuscitation	27	(1%)	262	(9%)	11.6	<0.001	7.54 to 18.0

Amb score

Multivariable analysis including all components of the Amb score, except access to transportation (which was present for all patients), is shown in Supplementary 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 (Supplementary Table 3). 94% (6325 admissions) had an Amb score of 5 or more, suggesting they could be discharged within 12 hours; 6% (418 admissions) had a score of less than 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%); 12.2% 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 to 48 hours).

*Table 4: Amb score performance. Performance in normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	Amb score N=6743		Amb score with NEWS2 N=6707	
	Frequency (%)		Frequency (%)	
Score <5 5+	 418 6325	 (6.2%) (93.8%)	 364 6343	 (5.4%) (94.6%)
Score <5 Admission length <12hrs	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				
Sensitivity	98.6% (95% CI 98.1% to 98.9%)		98.8% (95% CI 98.4% to 99.1%)	
Specificity	11.4% (95% CI 10.3% to 12.6%)		10.0% (95% CI 9.0% to 11.1%)	
PPV	55.0% (95% CI 53.8% to 56.2%)		54.5% (95% CI 53.3% to 55.8%)	
NPV	87.8% (95% CI 84.3% to 90.8%)		88.5% (95% CI 84.7% to 91.6%)	
% of patients discharged in <12 hours not identified by score*	1.4% (95% CI 1.1% to 2%)		1.2% (95% CI 0.9% to 1.6%)	
Patients identified as suitable by score admitted for >12 hours^	45.0% (95% CI 43.8% to 46.2%)		45.5% (95% CI 44.2% to 46.7%)	

Replacing MEWS with NEWS2, the AUROC was 0.612 (95% CI 0.599 to 0.625)(Figure 2b). 95% (6343 admissions) had an Amb score of 5 or more; 5% (364 admissions) had a score of less than 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%); 11.5% 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, 56.4% 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%, Chi square p=0.017), in both those discharged within 12 hours (13.7% vs 5.8%, p=0.017) and those discharged in 12 to 48 hours (11.7% vs 7.0%, p=0.001). This was also true for readmission within 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 to 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%, Chi square 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 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 (Supplementary 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, and where comorbidity due to ischaemic heart disease, heart failure, arrhythmia, diabetes, stroke/TIA, renal disease or chronic lung disease was present. A higher proportion of GP referrals with a raised Amb score were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (69% vs 45%, Chi square $p<0.005$). A higher proportion of patients with a raised Amb score and a NEWS2 of 0-2 were identified correctly compared to those with a raised NEWS2 on arrival.

GAPS

Multivariable analysis including all components of the GAPS is shown in Supplementary 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 more than 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 (Supplementary 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 48.5% to 52.5%), with a 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 to 48 hours).

*Table 5: GAPS performance within normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	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 <12hrs	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				
Sensitivity	50.4% (95% CI 48.5 to 52.5%)		50.0% (95% CI 47.8% to 52.1%)	
Specificity	63.1% (95% CI 61.3% to 64.9%)		63.3% (95% CI 61.5% to 65.0%)	
PPV	51.4% (95% CI 49.3% to 53.6%)		50.5% (95% CI 48.4% to 52.7%)	
NPV	62.1% (95% CI 60.3% to 63.9%)		62.7% (95% CI 60.9% to 64.5%)	
% of patients discharged in <12 hours not identified by score*	49.6% (95% CI 47.5% to 51.5%)		50.0% (95% CI 47.9% to 52.2%)	
Patients identified as suitable by score admitted for >12 hours^	48.6% (95% CI 46.4% to 50.7%)		49.5% (95% CI 47.3% to 51.6%)	

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).

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 to 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%, Chi square p<0.005), both for those discharged within 12 hours (6.0% vs 4.2%, p=0.055), and 12 to 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 to 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. 29 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, were female, and where comorbidity due to stroke/TIA was present (Supplementary Table 7). A higher proportion of GP referrals with a low GAPS were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (68% vs 50%, Chi square p=0.044). A higher proportion of patients with a low GAPS and a NEWS2 of 0-2 were identified correctly compared to 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 (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, 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 a Amb score suggesting suitability for SDEC with a GAPS suggesting admission (Chi square, p<0.0005 for each subgroup comparison, Figure 4).

Discussion

This paper highlights several important points. Firstly, 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 to 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 negative predictive value than the GAPS, with 88.5% of patients with a low Amb score (suggesting they were unsuitable for SDEC) remaining for more than 12 hours, compared to 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 (Supplementary 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 over 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, 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 more than 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 focussed on performance of scoring systems to identify patients suitable for SDEC within currently available services; in-depth evaluation of factors necessitating admission over 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 focussed on the ability of the Amb score and GAPS to discriminate between those admitted for <12 hours and 12 to 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 positive and negative

predictive value of the score. Although some aspects of score performance may be appear improved if the scores are able to identify all those admitted for over 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 positive predictive value 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 IV treatment to be 'anticipated' in patients receiving IV treatment within 6 hours of arrival, as anticipation of IV 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 was 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 Glasgow Admission Prediction Score 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 Same Day Emergency Care services.

Contributorship

CA and ES designed the study, CA analysed the data, all authors (CA, SG, EW, VRK, ES) 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.

License for publication

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.

Data sharing agreement
Data from this study is 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.

Ethics statement
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 HDR UK Hub in acute care by the East Midlands – Derby REC (reference: 20/EM/0158). Formal written consent from individual participants was not required.

Transparency declaration
The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Funding
No specific funding was available for this project.

References

1. NHS England. A&E Attendances and Emergency Admissions 2021-22 2021 [Available from: 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? Clinical Medicine. 2022:clinmed.2021-0614.
5. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
6. Royal College of Physicians. Acute Care Toolkit 10: Ambulatory Emergency Care. 2014.
7. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.
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, Holland M, Subbe C, Kennedy A, et al. Length of stay in Acute Medical Admissions: Analysis from the Society for Acute Medicine Benchmarking Audit. Acute Med. 2022;21(1):27-33.
11. Cameron A, Jones D, Logan E, O'Keeffe CA, Mason SM, Lowe DJ. Comparison of Glasgow Admission Prediction Score and Amb Score in predicting need for inpatient care. Emerg Med J. 2018;35(4):247-51.
12. Pendlebury ST, Lovett NG, Smith SC, Dutta N, Bendon C, Lloyd-Lavery A, 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(11):e007808.
13. Society for Acute Medicine. Society for Acute Medicine Benchmarking Audit 2021 - SAMBA2021 Report. 2021.
14. Atkin C, Knight T, Subbe C, Holland M, Cooksley T, Lasserson D. Acute care service performance during winter: report from the winter SAMBA 2020 national audit of acute care. Acute Med. 2020;19(4):220-9.
15. Baier N, Geissler A, Bech M, Bernstein D, Cowling TE, Jackson T, 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):1-10.
16. Canadian Institute for Health Information. NACRS emergency department visits and lengths of stay 2022 [Available from: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-lengths-of-stay.
17. Australian Institute of Health and Welfare. Admitted patients. 2022 [Available from: https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients.

Figure Legends

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

Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions.

Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, 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.

Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

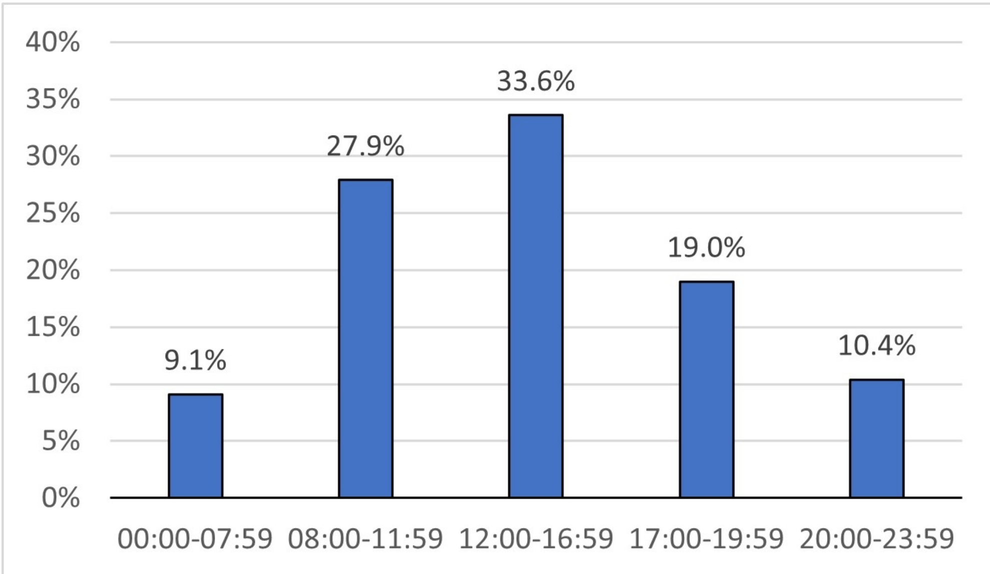


Figure 1: Arrival time for medical attendances lasting up to 48 hours.

99x57mm (330 x 330 DPI)

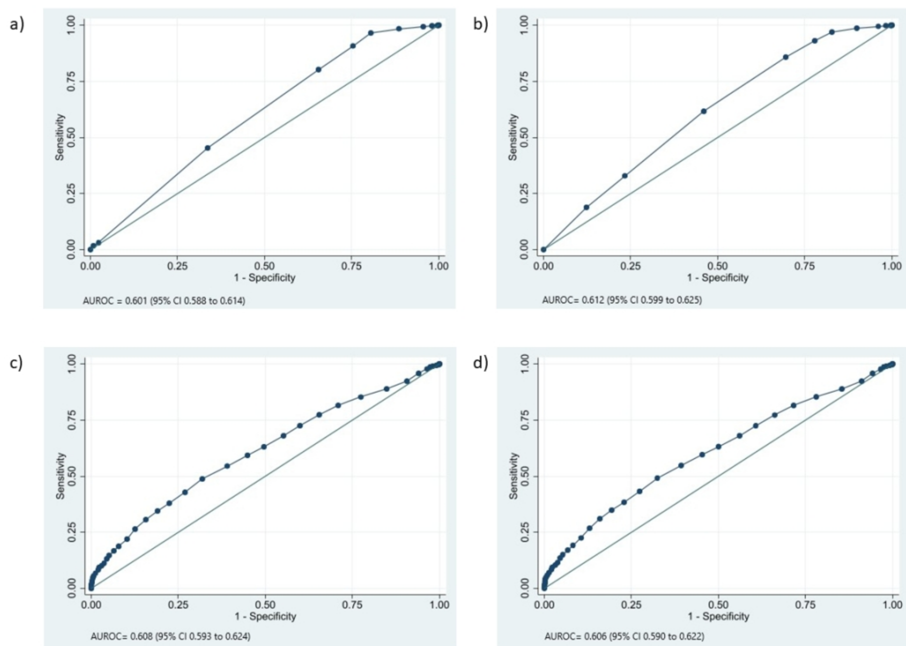
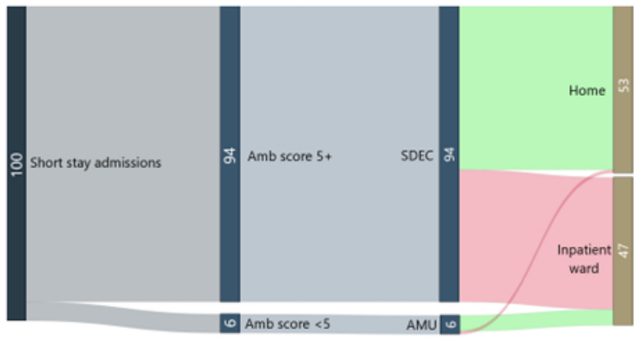


Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions.

855x834mm (38 x 38 DPI)

a) Amb score



b) GAPS (two level)

Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, 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.

481x476mm (38 x 38 DPI)

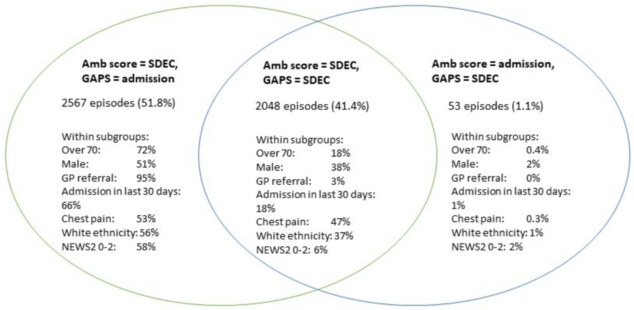


Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

338x190mm (96 x 96 DPI)

Supplementary Table 1: Triage problem. Commonest triage problem recorded on arrival to Emergency Department. Coded presenting problem entered at initial Emergency Department triage. Normal working day admissions defined as episodes starting between 08:00-16:59 Monday-Friday.

All admissions			**Normal working day admissions**		
	Frequency (%)			Frequency (%)	
Chest pain	3762	(33.5%)	Chest pain	1940	(36.8%)
Dyspnoea/difficulty breathing	1586	(14.1%)	Dyspnoea/difficulty breathing	721	(13.7%)
Asthenia	1051	(9.4%)	Asthenia	548	(10.4%)
Headache	609	(5.4%)	Headache	322	(6.1%)
Abdominal pain	408	(3.6%)	Abdominal pain	172	(3.3%)
Near syncope/syncope	282	(2.5%)	Palpitations	145	(2.8%)
Palpitations	256	(2.3%)	Near syncope/syncope	137	(2.6%)
Dizziness	222	(2.0%)	Dizziness	119	(2.3%)
Fever	210	(1.9%)	Pain in lower limb	96	(1.8%)
Substance abuse	210	(1.9%)	Vomiting	82	(1.6%)

Supplementary Table 2: Multivariate analysis of Amb score components. Odds ratio for admission of 12-48 hours, normal working day admissions. IV= intravenous, MEWS= Modified Early Warning Score, NEWS2= National Early Warning Score 2.(2)

Amb score components				Amb score components, substituting NEWS2			
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age >80	1.86	<0.0005	1.63 to 2.13	Age >80	1.85	<0.0005	1.62 to 2.13
Male	1.03	0.568	0.93 to 1.14	Male	1.02	0.733	0.92 to 1.13
IV treatment not anticipated	0.12	<0.0005	0.10 to 0.15	IV treatment not anticipated	0.12	<0.0005	0.10 to 0.15
Not acutely confused	0.38	0.068	0.13 to 1.08	Not acutely confused	0.40	0.09	0.14 to 1.15
MEWS 0	1.05	0.739	0.80 to 1.38	NEWS2 0	0.82	<0.0005	0.74 to 0.92
Not discharged in last 30 days	1.00	0.993	0.86 to 1.16	Not discharged in last 30 days	1.00	0.907	0.87 to 1.17

Supplementary table 3: Amb score for NWD (Normal working day) admission episodes. Normal working day defined as episodes starting between 08:00-16:59 Monday-Friday. Amb score calculated as shown in Table 1.(1) NEWS2: National Early Warning Score 2.(2)

	Amb score		Amb score substituting NEWS2	
Amb score	Number of episodes (%)		Number of episodes (%)	
≤3	12	(0.2%)	12	(0.2%)
3.5	51	(0.8%)	44	(0.7%)
4	98	(1.5%)	81	(1.2%)
4.5	257	(3.8%)	227	(3.4%)
5	327	(4.9%)	287	(4.3%)
5.5	367	(5.4%)	295	(4.4%)
6	690	(10.2%)	522	(7.8%)
6.5	2261	(33.5%)	1605	(23.9%)
7	2502	(37.1%)	1735	(12.6%)
7.5	94	(1.4%)	846	(15.7%)
8	84	(1.3%)	1053	(12.3%)

Supplementary Table 4: Identifying length of admission by Amb score (incorporating NEWS2) within patient subgroups. Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Amb score calculated as per Table 1, with NEWS2 substituted in place of MEWS. NEWS2: National Early Warning Score 2.(2) MEWS: Modified Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for comparisons using Chi square.

	Amb 5+, Admission length <12hrs Correctly identified		Amb 5+, Admission length 12-48 hrs Incorrectly identified		Amb <5, Admission length <12 hours Incorrectly identified		Amb <5, Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by Amb score discharged within 12 hours	P value
Percentage of admissions	51.6%		43.0%		0.6%		4.8%		55.0%	
	N	%	N	%	N	%	N	%		
Age										
16-19	85	52.5%	70	43.2%	<10	<6.2%	<10	<6.2%	54.8%	<0.005
20-29	340	51.1%	291	43.8%	<10	<1.5%	28	4.2%	53.9%	
30-39	404	54.0%	310	41.4%	<10	<1.3%	27	3.6%	56.6%	
40-49	465	56.8%	330	40.3%	<10	<1.2%	20	2.4%	58.5%	
50-59	630	56.0%	445	39.6%	12	1.1%	38	3.4%	58.6%	
60-69	564	57.9%	370	38.0%	<10	<1.0%	38	3.9%	60.4%	
70-79	547	49.5%	506	45.8%	<10	<0.9%	51	4.6%	51.9%	
80-89	357	41.4%	426	49.5%	<10	<1.2%	69	8.0%	45.6%	
90+	67	27.0%	136	54.8%	0	-	45	18.1%	33.0%	
Under 70	3035	54.2%	2322	41.5%	33	0.6%	208	3.7%	56.7%	<0.005
Over 70	424	38.2%	562	50.7%	<10	<0.9%	114	10.3%	43.0%	
Sex										
Female	2022	52.2%	1749	45.1%	12	0.3%	94	2.4%	53.6%	0.077
Male	1437	50.8%	1135	40.1%	30	1.1%	228	8.1%	55.9%	
Ethnicity										
Asian	500	51.4%	440	45.3%	<10	<1.0%	26	2.7%	53.2%	0.191
Black	169	56.0%	122	40.4%	<10	<3.3%	10	3.3%	58.1%	
Unknown	395	54.0%	287	39.3%	11	1.5%	38	5.2%	57.9%	
Mixed	58	51.3%	48	42.5%	<10	<8.8%	<10	<8.8%	54.7%	
Other	103	57.9%	72	40.4%	0		<10	<5.6%	58.9%	
White	2234	50.6%	1915	43.4%	23	1.0%	239	5.4%	53.8%	
Recent admission (30 days)										
Yes	433	50.3%	335	39.0%	11	1.3%	81	9.4%	56.4%	0.273
No	3026	51.8%	2549	43.6%	31	0.5%	241	4.1%	54.3%	
GP referral										
Yes	1792	67.3%	823	30.9%	10	0.4%	39	1.5%	68.5%	<0.005
No	1667	41.2%	2061	51.0%	32	0.8%	283	7.0%	44.7%	
Chest pain as triage problem										
Yes	1032	57.7%	739	41.3%	<10	<0.6%	12	0.7%	58.3%	<0.005
No	2427	49.4%	2145	43.6%	35	0.7%	310	6.3%	53.1%	
History of IHD										
Yes	834	49.7%	766	45.6%	<10	<0.6%	69	4.1%	52.1%	0.025
No	2625	52.2%	2118	42.1%	33	0.7%	253	5.0%	55.3%	
History of HF										
Yes	111	36.0%	167	54.2%	<10	<3.2%	27	8.8%	39.9%	<0.005
No	3348	52.3%	2717	42.5%	39	0.6%	295	4.6%	55.2%	
History of arrhythmia										
Yes	323	37.9%	438	51.3%	<10	<1.2%	83	9.7%	42.4%	<0.005
No	3136	53.6%	2446	41.8%	33	0.6%	239	4.1%	56.2%	
History of diabetes										
Yes	497	44.0%	546	48.4%	<10	<0.9%	79	7.0%	47.7%	<0.005
No	2962	53.1%	2338	41.9%	35	0.6%	243	4.4%	55.9%	
History of stroke										
Yes	18	17.8%	80	79.2%	0	-	<10	<10.0%	18.4%	<0.005
No	3441	52.1%	2804	42.4%	42	0.6%	319	4.8%	55.1%	
History of renal disease										
Yes	167	40.7%	197	48.0%	0	-	46	11.2%	45.9%	<0.005
No	3292	52.3%	2687	42.7%	42	0.7%	276	4.4%	55.1%	
History of chronic lung disease										
Yes	703	47.5%	674	45.5%	12	0.8%	92	6.2%	51.5%	<0.005
No	2756	52.7%	2210	42.3%	32	0.6%	230	4.4%	55.5%	
NEWS2										
0-2	3180	54.8%	2435	41.9%	29	0.5%	162	2.8%	56.6%	<0.005
3-4	252	38.0%	319	48.1%	<10	<1.5%	85	12.8%	44.1%	
5+	27	11.3%	130	54.6%	<10	<4.2%	75	31.5%	17.2%	

Supplementary Table 5: Multivariate analysis of GAPS components. Age – odds ratio (OR) per decade increase in age; NEWS/NEWS2 OR per increase of one point in NEWS/NEWS2. Triage category compared to 'standard' as reference. Odds ratio for admission of 12-48 hours, normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). GP= general practitioner, NEWS= national early warning score

	GAPS				GAPS with NEWS2		
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age	1.06	<0.0005	1.03 to 1.10	Age	1.06	<0.0005	1.03 to 1.10
NEWS	1.24	<0.0005	1.12 to 1.31	NEWS2	1.22	<0.0005	1.16 to 1.28
Triage category*				Triage category*			
Urgent	1.08	0.457	0.88 to 1.32	Urgent	1.04	0.692	0.85 to 1.29
Resuscitation	4.56	<0.0005	2.89 to 7.12	Resuscitation	4.26	<0.0005	2.69 to 6.74
Referred by GP	0.80	<0.0005	0.69 to 0.91	Referred by GP	0.78	0.001	0.68 to 0.90
Arrived in ambulance	1.61	<0.0005	1.41 to 1.83	Arrived in ambulance	1.60	<0.0005	1.40 to 1.82
Admitted <1 year ago	1.41	<0.0005	1.25 to 1.59	Admitted <1 year ago	1.39	<0.0005	1.23 to 1.57

Supplementary Table 6: GAPS for normal working day admissions. GAPS: Glasgow Admission Prediction Score, calculated as described in Table 1.(3) NEWS2: National Early Warning Score 2.(2)

	GAPS score N=5091		GAPS score substituting NEWS2 N=4953	
GAPS score	Number of episodes (%)		Number of episodes (%)	
1-5	93	(1.8%)	88	(1.8%)
6-19	829	(16.3%)	792	(16.0%)
11-15	1257	(24.7%)	1221	(24.7%)
16-20	1329	(26.1%)	1279	(15.8%)
21-25	874	(17.2%)	857	(17.3%)
26-30	354	(7.0%)	360	(7.3%)
31-35	211	(4.1%)	206	(4.2%)
36-40	97	(1.9%)	94	(1.9%)
41-45	41	(0.8%)	45	(0.9%)
46+	<10	(<0.2%)	11	(0.2%)

Supplementary Table 7: Identifying length of admission by GAPS (incorporating NEWS2) within patient subgroups. Analysis of Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Glasgow Admission Prediction Score (GAPS) calculated as per Table 1, with NEWS2 substituted in place of NEWS. NEWS2: National Early Warning Score 2.(2) NEWS: National Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for Chi square comparisons.

GAPS with NEWS2	GAPS≤15 Admission length <12hrs Correctly identified		GAPS ≤15 Admission length 12-48 hrs Incorrectly identified		GAPS 16+, Admission length <12 hours Incorrectly identified		GAPS 16+ Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by GAPS discharged within 12 hours	P value
Percentage of admissions	21.4%		21.0%		21.5%		36.1%		50.5%	
Age (years)										
16-19	32	26.4%	48	39.7%	17	14.0%	24	19.8%	40.0%	<0.005
20-29	140	26.7%	180	34.3%	91	17.4%	113	21.6%	43.8%	
30-39	172	29.6%	185	31.8%	101	17.4%	123	21.2%	48.2%	
40-49	228	35.3%	178	27.6%	104	16.1%	135	20.9%	56.2%	
50-59	237	27.8%	191	22.4%	188	22.1%	235	27.6%	55.4%	
60-69	126	18.1%	102	14.6%	216	31.0%	253	36.3%	55.3%	
70-79	87	11.4%	89	11.7%	201	26.4%	385	50.5%	49.4%	
80-89	33	5.7%	59	10.3%	122	21.2%	361	62.8%	35.9%	
90+	<10	<5.1%	<10	<5.1%	23	11.7%	160	81.2%	50.0%	
Under	1022	24.4%	973	23.3%	918	22.0%	1268	30.3%	51.2%	0.007
Over 70	40	5.2%	66	8.5%	145	18.8%	521	67.5%	37.7%	
Sex										
Female	599	21.3%	633	22.6%	597	21.3%	977	34.8%	48.6%	0.035
Male	463	21.6%	406	18.9%	466	21.7%	811	37.8%	53.3%	
Ethnicity										
Asian	223	27.6%	188	23.2%	157	19.4%	241	29.8%	54.3%	0.25
Black	57	26.1%	48	22.0%	48	22.0%	65	29.8%	54.3%	
Unknown	135	25.7%	127	24.2%	102	19.4%	161	30.7%	51.5%	
Mixed	20	22.0%	27	29.7%	20	22.0%	24	26.4%	42.6%	
Other	35	24.5%	47	32.9%	37	25.9%	24	16.8%	42.7%	
White	592	18.7%	602	19.0%	699	22.1%	1274	40.2%	49.6%	
Recent admission (30 days)										
Yes	45	8.7%	55	10.6%	122	23.6%	295	57.1%	45.0%	0.256
No	1017	22.9%	984	22.2%	941	21.2%	1494	33.7%	50.8%	
GP referral										
Yes	23	2.1%	11	1.0%	533	49.6%	508	47.3%	67.6%	0.044
No	1039	26.8%	1028	26.5%	530	13.7%	1281	33.0%	50.3%	
Chest pain as triage problem										
Yes	523	29.2%	318	17.8%	516	28.8%	433	24.2%	62.2%	<0.005
No	539	17.0%	721	22.8%	547	17.3%	1356	42.9%	42.8%	
History of IHD										
Yes	306	20.3%	231	15.3%	402	26.7%	568	37.7%	57.0%	<0.005
No	756	21.9%	808	23.4%	661	19.2%	1221	35.4%	48.3%	
History of heart failure										
Yes	20	8.0%	17	6.8%	51	20.5%	161	64.7%	54.1%	0.667
No	1042	22.2%	1022	21.7%	1012	21.5%	1628	34.6%	50.5%	
History of arrhythmia										
Yes	72	10.2%	78	11.0%	156	22.1%	401	56.7%	48.0%	0.517
No	990	23.3%	961	22.6%	907	21.4%	1388	32.7%	50.7%	
History of diabetes										
Yes	124	13.9%	146	16.3%	202	22.6%	421	47.1%	45.9%	0.104
No	938	23.1%	893	22.0%	861	21.2%	1368	33.7%	51.2%	
History of stroke										
Yes	<10	<10.4%	29	30.2%	<10	<10.4%	54	56.3%	17.1%	<0.005
No	1056	21.7%	1010	20.8%	1056	21.7%	1735	35.7%	51.1%	
History of renal disease										
Yes	26	8.7%	35	11.7%	61	20.3%	178	59.3%	42.6%	0.209
No	1036	22.3%	1004	21.6%	1002	21.5%	1611	34.6%	50.8%	
History of chronic lung disease										
Yes	191	16.8%	176	15.5%	262	23.0%	510	44.8%	52.0%	0.528
No	871	22.8%	863	22.6%	801	21.0%	1279	33.5%	50.2%	
NEWS2										
0-2	1002	33.0%	954	31.4%	952	31.3%	131	4.3%	51.2%	0.012
3-4	57	11.1%	72	14.0%	93	18.1%	291	56.7%	44.2%	
5+	<10	<4.6%	13	5.9%	18	8.2%	185	84.5%	18.8%	

Supplementary Table 8: Comparison of key characteristics of this analysis with original derivation of Amb score(1) and Glasgow Admission Prediction Score (GAPS)(3).

	Population	Episode start time	Comparator	Location	Sample size	Study period
This analysis	Unplanned attendances to acute medicine	08:00-16:59, Monday to Friday	Discharged in <12 hours vs admitted for 12-48 hours	Birmingham, UK (single hospital)	7365 episodes	April 2019-March 2020
Amb score – Ala et al, 2012	Unplanned attendances to acute medicine	Unrestricted	Discharged in <12 hours vs admitted for >48 hours	South Wales, UK (single hospital)	625 episodes (derivation: 282, validation: 343)	May-June 2010 (derivation), June-July 2011 (validation)
GAPS score – Cameron et al, 2015	Unplanned attendances to Emergency Department, acute medicine, or minor injuries unit	Unrestricted	Clinical decision to discharge vs clinical decision to admit to hospital	North Glasgow, UK (3 hospitals)	322,846 episodes (derivation: 215,231, validation: 107,615	March 2010-March 2012

References

1. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
2. Royal College of Physicians. National Early Warning Score (NEWS) 2. 2017.
3. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1 & 2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5
	11	Rationale for choosing the reference standard (if alternatives exist)	5
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	5
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5/6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	5/6
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6
	15	How indeterminate index test or reference standard results were handled	6
	16	How missing data on the index test and reference standard were handled	6
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6
	18	Intended sample size and how it was determined	5
RESULTS			
Participants	**19**	Flow of participants, using a diagram	-
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	8
	22	Time interval and any clinical interventions between index test and reference standard	-
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	12,14
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	12,14
	25	Any adverse events from performing the index test or the reference standard	-
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	18
	27	Implications for practice, including the intended use and clinical role of the index test	17-18
OTHER INFORMATION			
	28	Registration number and name of registry	-
	29	Where the full study protocol can be accessed	-
	30	Sources of funding and other support; role of funders	21

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064910.R2
Article Type:	Original research
Date Submitted by the Author:	04-Nov-2022
Complete List of Authors:	Atkin, Catherine; University of Birmingham, Birmingham Acute Care Research Group Gallier, Suzy; University Hospitals Birmingham NHS Foundation Trust, Department of Health Informatics Wallin, Elizabeth; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine Reddy-Kolanu, Vinay; University Hospitals Birmingham NHS Foundation Trust, Acute medicine Sapey, Elizabeth; University of Birmingham, PIONEER HDR-UK Hub; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Evidence based practice
Keywords:	INTERNAL MEDICINE, GENERAL MEDICINE (see Internal Medicine), Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™ Manuscripts

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 (0000-0003-0596-8515), Suzy Gallier (0000-0003-1026-4125), Elizabeth Wallin, Vinay Reddy-Kolanu, Elizabeth Sapey (0000-0003-3454-5482)

Addresses

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Edgbaston, Birmingham, B15 2GW, UK. Catherine Atkin, NIHR Academic Clinical Lecturer in Acute Medicine.

Department of Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Suzy Gallier, Lead for Research Analytics.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Wallin, Consultant in Nephrology & Acute Medicine.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Vinay Reddy-Kolanu, Consultant in Acute Medicine.

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, and Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Sapey, Professor of Acute and Respiratory Medicine and Honorary Consultant.

Correspondence to: 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 and lasting up to 48 hours, between 1st April 2019 and 9th 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 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

- This study compared performance of the Amb score and GAPS 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 NEWS2, reflecting current clinical practice
- Patients admitted for longer than 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 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 less than 12 hours or over 48 hours. That study excluded patients who remained in hospital for 12 to 48 hours.

Table 1: Scoring systems to identify medical admissions potentially suitable for discharge from hospital without admission >12 hours. Amb score(5) and Glasgow Admission Prediction Score (GAPS)(7). Amb score of 5 more indicates likely discharge within 12 hours; GAPS of 16 or more suggests patient likely to be admitted to hospital. IV = intravenous, MEWS = Modified early warning score, NEWS = National Early Warning Score, GP = General practitioner

Amb score		
Sex	Female	0
	Male	-0.5
Age	<80	0
	≥80	-0.5
Access to personal transport/can take public transport	Agree	2
	Disagree	0
IV treatment not anticipated	Agree	2
	Disagree	0
Not acutely confused	Agree	2
	Disagree	0
MEWS=0	Agree	1
	Disagree	0
Not discharged from hospital within previous 30 days	Agree	1
	Disagree	0

Glasgow Admissions Prediction Score (GAPS)		
NEWS		1 point per point on NEWS score
Age		1 point per decade
Triage category	3	5
	2 (or 2+)	10
	1	20
Referred by GP		5
Arrived in ambulance		5
Admitted <1 year ago		5

The Glasgow Admission Prediction Score (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 Same Day Emergency Care and those requiring admission for at least 12 to 48 hours.

Methods

This data study was conducted in collaboration with PIONEER, a Health Data Research Hub in Acute Care, and all study processes were carried out following appropriate ethical approval provided by the East Midlands – Derby REC (reference: 20/EM/0158).

Retrospective data were collected for patients admitted to Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Trust (UHB) between the period of the 1st April 2019 until 9th 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 (CCU) in Europe. The Acute Medical Unit (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 inhouse electronic health record (EHR) called Prescribing Information and Communication System (PICS). 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 focussed 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 and mapped ICD10 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 to 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 08:00-16:59, 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 (IV) treatment was taken as not being anticipated where patients did not receive an IV 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 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 National Early Warning Score 2 (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 15.1. Cell counts containing less than 10 patients were suppressed, due to reporting requirements. For univariate analysis of factors influencing likelihood of discharge within 12 hours, odds ratios 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 was performed using Chi square. A p value of <0.05 is 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 less than 48 hours, based on previous research.(10)

Results

14314 acute medical inpatient episodes lasting up to 48 hours were identified during the censor period. These episodes were from 12587 patients with 11229 patients having one episode in this time period. Patients 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.

Table 2: Demographics and characteristics of patients with emergency medical admissions lasting up to 48 hours. For whole cohort, and for patients arriving in a normal working day (08:00-16:59, Monday to Friday).P values shown for Chi square comparison of normal working day episodes to episodes starting outside normal working day.

	All episodes N=14314		Normal working day episodes N=7365		Episodes starting outside normal working day N= 6949		P value
	Frequency (%)		Frequency (%)		Frequency (%)		
Age							
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 day	1047	(7.3%)	479	(6.5%)	568	(8.2%)	<0.001
14 day	1544	(11%)	681	(9%)	863	(12%)	<0.001
30 day	2268	(16%)	1033	(14%)	1235	(18%)	<0.001

18% of episodes occurred on a weekend. Overall, 62% of patients arrived between 08:00-16:59 (Figure 1); 63% of weekday episodes started between these times.

11244 episodes had an associated Emergency Department triage code, with 108 different triage codes used. The commonest triage problem was chest pain (34% of episodes), see Supplementary Table 1. 6394 episodes (44%) had a length of stay of less than 12 hours.

Normal working day arrivals

There were 7365 episodes in 6848 patients with an arrival time between 08:00-16:59 on a weekday (normal working day, NWD). The triage problem was available for 5272 NWD episodes (72%). The commonest triage problem was chest pain (37%) (Supplementary Table 1).

4053 episodes (55%) had a length of stay of less than 12 hours and 3312 (45%) were discharged after 12 to 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%, Chi square p<0.005).

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

Compared to patients discharged within 12 to 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, Chi square 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 and anticipated need for IV therapy were associated with an increased risk of admission lasting more than 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 50-140bpm and systolic blood pressure between 100-200mmHg 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 ACS 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).

*Table 3: Factors considered in derivation of previous scoring systems. Column percentages shown. Univariate analysis, odds ratio for admission lasting 12-48 hours shown. IV: intravenous; RR: respiratory rate; HR: heart rate in beats per minute; SBP: systolic blood pressure in mmHg; MEWS: Modified Early Warning Score; NEWS2: National Early Warning Score 2(9); IHD: ischaemic heart disease; GP: general practice. 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.*

N=7365 unless otherwise stated	Length of stay				Odds ratio (OR)	P value	95% CI OR
	<12hrs		12-48 hours				
	Frequency (%)		Frequency (%)				
Age							
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
IV 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 *(n=6748)*	3499	(99.0%)	3144	(97.9%)	0.42	<0.001	0.25 to 0.69
SBP 100-200 *(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
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

Amb score

Multivariable analysis including all components of the Amb score, except access to transportation (which was present for all patients), is shown in Supplementary 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 (Supplementary Table 3). 94% (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 less than 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 to 48 hours).

*Table 4: Amb score performance. Performance in normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	Amb score N=6743		Amb score with NEWS2 N=6707	
	Frequency (%)		Frequency (%)	
Score <5 5+	 418 6325	 (6.2%) (93.8%)	 364 6343	 (5.4%) (94.6%)
Score <5 Admission length <12hrs	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%)	

Replacing MEWS with NEWS2, the AUROC was 0.612 (95% CI 0.599 to 0.625)(Figure 2b). 95% (6343 admissions) had an Amb score of 5 or more; 5.4% (364 admissions) had a score of less than 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%, Chi square p=0.017), in both those discharged within 12 hours (13.7% vs 5.8%, p=0.017) and those discharged in 12 to 48 hours (11.7% vs 7.0%, p=0.001). This was also true for readmission within 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 to 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%, Chi square 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 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 (Supplementary 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, and where comorbidity due to ischaemic heart disease, heart failure, arrhythmia, diabetes, stroke/TIA, renal disease or chronic lung disease was present. A higher proportion of GP referrals with a raised Amb score were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (69% vs 45%, Chi square $p<0.005$). A higher proportion of patients with a raised Amb score and a NEWS2 of 0-2 were identified correctly compared to those with a raised NEWS2 on arrival.

GAPS

Multivariable analysis including all components of the GAPS is shown in Supplementary 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 more than 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 (Supplementary 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 48.5% to 52.5%), with a 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 to 48 hours).

*Table 5: GAPS performance within normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	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 <12hrs	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%)	

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).

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 to 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%, Chi square p<0.005), both for those discharged within 12 hours (6.0% vs 4.2%, p=0.055), and 12 to 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 to 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. 29 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, were female, and where comorbidity due to stroke/TIA was present (Supplementary Table 7). A higher proportion of GP referrals with a low GAPS were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (68% vs 50%, Chi square p=0.044). A higher proportion of patients with a low GAPS and a NEWS2 of 0-2 were identified correctly compared to 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 (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, 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 a Amb score suggesting suitability for SDEC with a GAPS suggesting admission (Chi square, p<0.0005 for each subgroup comparison, Figure 4).

Discussion

This paper highlights several important points. Firstly, 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 to 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 negative predictive value than the GAPS, with 88.5% of patients with a low Amb score (suggesting they were unsuitable for SDEC) remaining for more than 12 hours, compared to 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 (Supplementary 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 over 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, 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 more than 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 focussed on performance of scoring systems to identify patients suitable for SDEC within currently available services; in-depth evaluation of factors necessitating admission over 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 focussed on the ability of the Amb score and GAPS to discriminate between those admitted for <12 hours and 12 to 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 positive and negative

predictive value of the score. Although some aspects of score performance may be appear improved if the scores are able to identify all those admitted for over 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 positive predictive value 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 IV treatment to be 'anticipated' in patients receiving IV treatment within 6 hours of arrival, as anticipation of IV 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 was 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 Glasgow Admission Prediction Score 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 Same Day Emergency Care services.

Contributorship

CA and ES designed the study, CA analysed the data, all authors (CA, SG, EW, VRK, ES) 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.

License for publication

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.

Data sharing agreement
Data from this study is 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.

Ethics statement
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 HDR UK Hub in acute care by the East Midlands – Derby REC (reference: 20/EM/0158). Formal written consent from individual participants was not required.

Transparency declaration
The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Funding
No specific funding was available for this project.

References

1. NHS England. A&E Attendances and Emergency Admissions 2021-22 2021 [Available from: 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? Clinical Medicine. 2022:clinmed.2021-0614.

5. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.

6. Royal College of Physicians. Acute Care Toolkit 10: Ambulatory Emergency Care. 2014.

7. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.

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, Holland M, Subbe C, Kennedy A, et al. Length of stay in Acute Medical Admissions: Analysis from the Society for Acute Medicine Benchmarking Audit. Acute Med. 2022;21(1):27-33.

11. Cameron A, Jones D, Logan E, O'Keeffe CA, Mason SM, Lowe DJ. Comparison of Glasgow Admission Prediction Score and Amb Score in predicting need for inpatient care. Emerg Med J. 2018;35(4):247-51.

12. Pendlebury ST, Lovett NG, Smith SC, Dutta N, Bendon C, Lloyd-Lavery A, 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(11):e007808.

13. Society for Acute Medicine. Society for Acute Medicine Benchmarking Audit 2021 - SAMBA2021 Report. 2021.

14. Atkin C, Knight T, Subbe C, Holland M, Cooksley T, Lasserson D. Acute care service performance during winter: report from the winter SAMBA 2020 national audit of acute care. Acute Med. 2020;19(4):220-9.

15. Baier N, Geissler A, Bech M, Bernstein D, Cowling TE, Jackson T, 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):1-10.

16. Canadian Institute for Health Information. NACRS emergency department visits and lengths of stay 2022 [Available from: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-lengths-of-stay.

17. Australian Institute of Health and Welfare. Admitted patients. 2022 [Available from: https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients.

Figure Legends

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

Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions. AUROC: area under the receiver operating characteristic curve; 95% CI: 95% confidence interval.

Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, 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.

Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

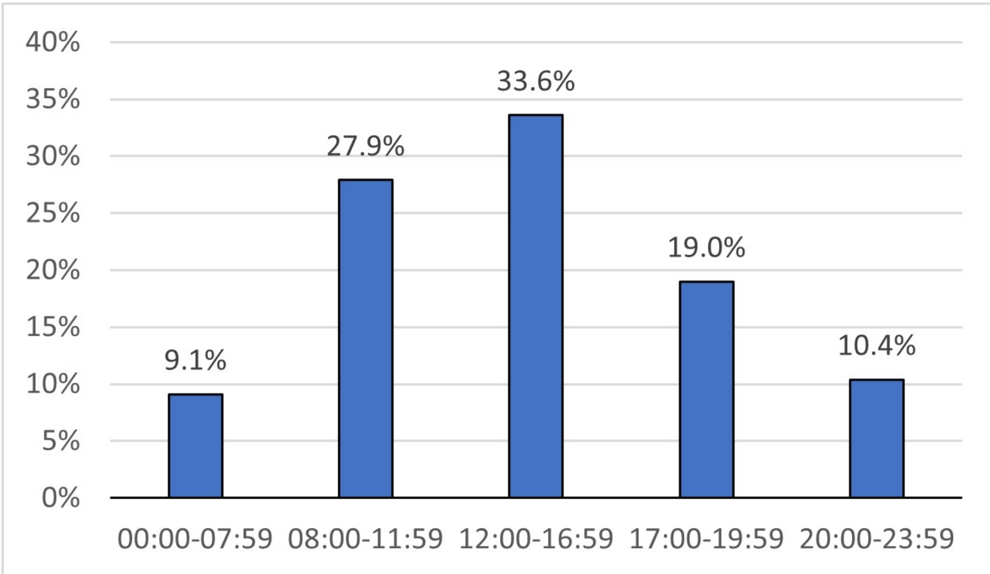


Figure 1: Arrival time for medical attendances lasting up to 48 hours.

99x57mm (330 x 330 DPI)

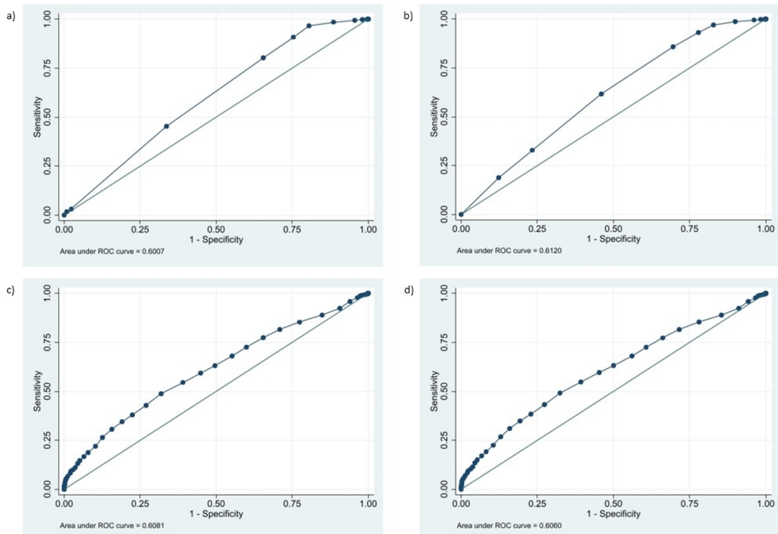


Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions. AUROC: area under the receiver operating characteristic curve; 95% CI: 95% confidence interval.

338x190mm (96 x 96 DPI)

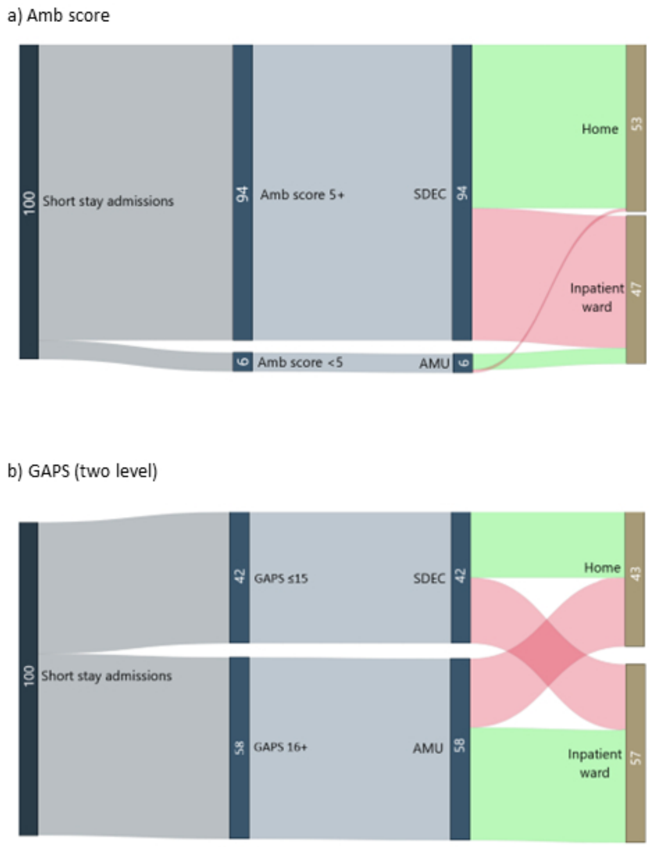


Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, for a) Amb score (5 or more) and b) Glasgow Admission Prediction Score (GAPS)(≤15). Green = currently identified by scoring system, red = incorrectly identified by scoring system.

481x476mm (38 x 38 DPI)

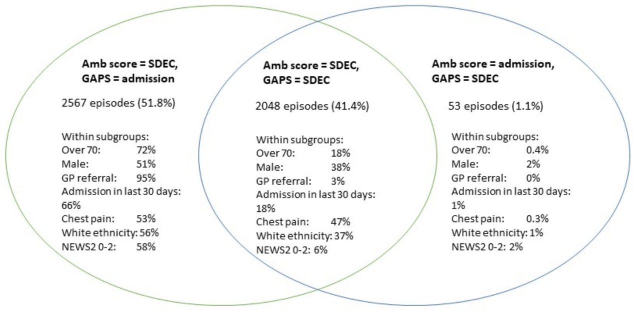


Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

338x190mm (96 x 96 DPI)

Supplementary Table 1: Triage problem. Commonest triage problem recorded on arrival to Emergency Department. Coded presenting problem entered at initial Emergency Department triage. Normal working day admissions defined as episodes starting between 08:00-16:59 Monday-Friday.

All admissions			**Normal working day admissions**		
	Frequency (%)			Frequency (%)	
Chest pain	3762	(34%)	Chest pain	1940	(37%)
Dyspnoea/difficulty breathing	1586	(14%)	Dyspnoea/difficulty breathing	721	(14%)
Asthenia	1051	(9.4%)	Asthenia	548	(10%)
Headache	609	(5.4%)	Headache	322	(6.1%)
Abdominal pain	408	(3.6%)	Abdominal pain	172	(3.3%)
Near syncope/syncope	282	(2.5%)	Palpitations	145	(2.8%)
Palpitations	256	(2.3%)	Near syncope/syncope	137	(2.6%)
Dizziness	222	(2.0%)	Dizziness	119	(2.3%)
Fever	210	(1.9%)	Pain in lower limb	96	(1.8%)
Substance abuse	210	(1.9%)	Vomiting	82	(1.6%)

Supplementary Table 2: Multivariable analysis of Amb score components. Mixed-effects logistic regression, patient as random effect. Odds ratio for admission of 12-48 hours, normal working day admissions. IV= intravenous, MEWS= Modified Early Warning Score, NEWS2= National Early Warning Score 2.(1)

Amb score components				Amb score components, substituting NEWS2			
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age >80	2.03	<0.001	1.71 to 2.41	Age >80	2.01	<0.001	1.69 to 2.38
Male	1.03	0.59	0.92 to 1.16	Male	1.02	0.735	0.91 to 1.14
IV treatment not anticipated	0.10	<0.001	0.7 to 0.13	IV treatment not anticipated	0.12	<0.001	0.07 to 0.14
Not acutely confused	0.32	0.06	0.10 to 1.04	Not acutely confused	0.35	0.08	0.11 to 1.13
MEWS 0	1.06	0.73	0.77 to 1.43	NEWS2 0	0.81	<0.001	0.72 to 0.91
Not discharged in last 30 days	1.00	0.96	0.84 to 1.18	Not discharged in last 30 days	1.01	0.94	0.85 to 1.19

Supplementary table 3: Amb score for NWD (Normal working day) admission episodes. Normal working day defined as episodes starting between 08:00-16:59 Monday-Friday. Amb score calculated as shown in Table 1.(2) NEWS2: National Early Warning Score 2.(1)

	Amb score		Amb score substituting NEWS2	
Amb score	Number of episodes (%)		Number of episodes (%)	
≤3	12	(0.2%)	12	(0.2%)
3.5	51	(0.8%)	44	(0.7%)
4	98	(1.5%)	81	(1.2%)
4.5	257	(3.8%)	227	(3.4%)
5	327	(4.9%)	287	(4.3%)
5.5	367	(5.4%)	295	(4.4%)
6	690	(10.2%)	522	(7.8%)
6.5	2261	(33.5%)	1605	(23.9%)
7	2502	(37.1%)	1735	(12.6%)
7.5	94	(1.4%)	846	(15.7%)
8	84	(1.3%)	1053	(12.3%)

Supplementary Table 4: Identifying length of admission by Amb score (incorporating NEWS2) within patient subgroups. Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Amb score calculated as per Table 1, with NEWS2 substituted in place of MEWS. NEWS2: National Early Warning Score 2.(1) MEWS: Modified Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for comparisons using Chi square.

	Amb 5+, Admission length <12hrs Correctly identified		Amb 5+, Admission length 12-48 hrs Incorrectly identified		Amb <5, Admission length <12 hours Incorrectly identified		Amb <5, Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by Amb score discharged within 12 hours	P value
Percentage of admissions	52%		43%		0.6%		4.8%		55%	
	N	%	N	%	N	%	N	%		
Age										
16-19	85	523%	70	43%	<10	<6.2%	<10	<6.2%	55%	<0.005
20-29	340	51%	291	44%	<10	<1.5%	28	4.2%	54%	
30-39	404	54%	310	41%	<10	<1.3%	27	3.6%	57%	
40-49	465	57%	330	40%	<10	<1.2%	20	2.4%	59%	
50-59	630	56%	445	40%	12	1.1%	38	3.4%	59%	
60-69	564	58%	370	38%	<10	<1.0%	38	3.9%	60%	
70-79	547	50%	506	46%	<10	<0.9%	51	4.6%	52%	
80-89	357	41%	426	50%	<10	<1.2%	69	8.0%	46%	
90+	67	27%	136	558%	0	-	45	18%	33%	
Under 70	3035	54%	2322	42%	33	0.6%	208	3.7%	57%	<0.005
Over 70	424	38%	562	51%	<10	<0.9%	114	10%	43%	
Sex										
Female	2022	52%	1749	45%	12	0.3%	94	2.4%	54%	0.08
Male	1437	50%	1135	40%	30	1.1%	228	8.1%	56%	
Ethnicity										
Asian	500	51%	440	45%	<10	<1.0%	26	2.7%	53%	0.19
Black	169	56%	122	40%	<10	<3.3%	10	3.3%	58%	
Unknown	395	54%	287	39%	11	1.5%	38	5.2%	58%	
Mixed	58	51%	48	43%	<10	<8.8%	<10	<8.8%	55%	
Other	103	58%	72	40%	0		<10	<5.6%	59%	
White	2234	51%	1915	43%	23	1.0%	239	5.4%	54%	
Recent admission (30 days)										
Yes	433	50%	335	39%	11	1.3%	81	9.4%	56%	0.27
No	3026	52%	2549	44%	31	0.5%	241	4.1%	54%	
GP referral										
Yes	1792	67%	823	31%	10	0.4%	39	1.5%	69%	<0.005
No	1667	41%	2061	51%	32	0.8%	283	7.0%	45%	
Chest pain as triage problem										
Yes	1032	58%	739	41%	<10	<0.6%	12	0.7%	58%	<0.005
No	2427	49%	2145	44%	35	0.7%	310	6.3%	53%	
History of IHD										
Yes	834	50%	766	46%	<10	<0.6%	69	4.1%	52%	0.025
No	2625	52%	2118	42%	33	0.7%	253	5.0%	55%	
History of HF										
Yes	111	36%	167	54%	<10	<3.2%	27	8.8%	40%	<0.005
No	3348	52%	2717	43%	39	0.6%	295	4.6%	55%	
History of arrhythmia										
Yes	323	38%	438	51%	<10	<1.2%	83	9.7%	42%	<0.005
No	3136	54%	2446	42%	33	0.6%	239	4.1%	56%	
History of diabetes										
Yes	497	44%	546	48%	<10	<0.9%	79	7.0%	48%	<0.005
No	2962	53%	2338	42%	35	0.6%	243	4.4%	56%	
History of stroke										
Yes	18	18%	80	79%	0	-	<10	<10%	18%	<0.005
No	3441	52%	2804	42%	42	0.6%	319	4.8%	55%	
History of renal disease										
Yes	167	41%	197	48%	0	-	46	11%	46%	<0.005
No	3292	52%	2687	43%	42	0.7%	276	4.4%	55%	
History of chronic lung disease										
Yes	703	48%	674	46%	12	0.8%	92	6.2%	52%	<0.005
No	2756	53%	2210	42%	32	0.6%	230	4.4%	56%	
NEWS2										
0-2	3180	55%	2435	42%	29	0.5%	162	2.8%	57%	<0.005
3-4	252	38%	319	48%	<10	<1.5%	85	13%	44%	
5+	27	11%	130	55%	<10	<4.2%	75	32%	17%	

Supplementary Table 5: Multivariable analysis of GAPS components. Mixed-effects logistic regression, patient as random effect. Age – odds ratio (OR) per decade increase in age; NEWS/NEWS2 OR per increase of one point in NEWS/NEWS2. Triage category compared to 'standard' as reference. Odds ratio for admission of 12-48 hours, normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). GP= general practitioner, NEWS= national early warning score

	GAPS				GAPS with NEWS2		
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age	1.07	<0.001	1.03 to 1.10	Age	1.07	<0.001	1.03 to 1.10
NEWS	1.25	<0.001	1.18 to 1.32	NEWS2	1.22	<0.001	1.16 to 1.29
Triage category*				Triage category*			
Urgent	1.08	0.46	0.88 to 1.33	Urgent	1.04	0.69	0.84 to 1.29
Resuscitation	4.64	<0.001	2.88 to 7.46	Resuscitation	4.32	<0.001	2.68 to 6.95
Referred by GP	0.79	0.002	0.69 to 0.92	Referred by GP	0.78	0.001	0.67 to 0.90
Arrived in ambulance	1.62	<0.001	1.40 to 1.86	Arrived in ambulance	1.61	<0.001	1.40 to 1.86
Admitted <1 year ago	1.42	<0.001	1.24 to 1.61	Admitted <1 year ago	1.40	<0.001	1.22 to 1.60

Supplementary Table 6: GAPS for normal working day admissions. GAPS: Glasgow Admission Prediction Score, calculated as described in Table 1.(3) NEWS2: National Early Warning Score 2.(1)

	GAPS score N=5091		GAPS score substituting NEWS2 N=4953	
GAPS score	Number of episodes (%)		Number of episodes (%)	
1-5	93	(1.8%)	88	(1.8%)
6-19	829	(16.3%)	792	(16.0%)
11-15	1257	(24.7%)	1221	(24.7%)
16-20	1329	(26.1%)	1279	(15.8%)
21-25	874	(17.2%)	857	(17.3%)
26-30	354	(7.0%)	360	(7.3%)
31-35	211	(4.1%)	206	(4.2%)
36-40	97	(1.9%)	94	(1.9%)
41-45	41	(0.8%)	45	(0.9%)
46+	<10	(<0.2%)	11	(0.2%)

Supplementary Table 7: Identifying length of admission by GAPS (incorporating NEWS2) within patient subgroups. Analysis of Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Glasgow Admission Prediction Score (GAPS) calculated as per Table 1, with NEWS2 substituted in place of NEWS. NEWS2: National Early Warning Score 2.(1) NEWS: National Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for Chi square comparisons.

GAPS with NEWS2	GAPS≤15 Admission length <12hrs Correctly identified		GAPS ≤15 Admission length 12-48 hrs Incorrectly identified		GAPS 16+, Admission length <12 hours Incorrectly identified		GAPS 16+ Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by GAPS discharged within 12 hours	P value
Percentage of admissions	21%		21%		22%		36%		50%	
Age (years)										
16-19	32	26%	48	40%	17	14%	24	20%	40%	<0.005
20-29	140	27%	180	34%	91	17%	113	22%	44%	
30-39	172	30%	185	32%	101	17%	123	21%	48%	
40-49	228	35%	178	28%	104	16%	135	21%	56%	
50-59	237	28%	191	22%	188	22%	235	28%	55%	
60-69	126	18%	102	15%	216	31%	253	36%	55%	
70-79	87	11%	89	12%	201	26%	385	51%	49%	
80-89	33	5.7%	59	10%	122	21%	361	63%	34%	
90+	<10	<5.1%	<10	<5.1%	23	12%	160	81%	50%	
Under	1022	24%	973	23%	918	22%	1268	30%	51%	0.007
Over 70	40	5.2%	66	8.5%	145	19%	521	68%	38%	
Sex										
Female	599	21%	633	23%	597	21.3%	977	35%	48%	0.035
Male	463	22%	406	19%	466	21.7%	811	38%	53%	
Ethnicity										
Asian	223	28%	188	23%	157	19.4%	241	30%	54%	0.25
Black	57	26%	48	22%	48	22.0%	65	30%	54%	
Unknown	135	26%	127	24%	102	19.4%	161	31%	52%	
Mixed	20	22%	27	30%	20	22.0%	24	26%	43%	
Other	35	25%	47	33%	37	25.9%	24	17%	43%	
White	592	19%	602	19%	699	22.1%	1274	40%	50%	
Recent admission (30 days)										
Yes	45	8.7%	55	11%	122	23.6%	295	57%	45%	0.26
No	1017	23%	984	22%	941	21.2%	1494	34%	51%	
GP referral										
Yes	23	2.1%	11	1.0%	533	49.6%	508	47%	68%	0.044
No	1039	27%	1028	27%	530	13.7%	1281	33%	50%	
Chest pain as triage problem										
Yes	523	29%	318	18%	516	28.8%	433	24%	62%	<0.005
No	539	17%	721	23%	547	17.3%	1356	43%	43%	
History of IHD										
Yes	306	20%	231	15%	402	26.7%	568	38%	57%	<0.005
No	756	22%	808	23%	661	19.2%	1221	35%	48%	
History of heart failure										
Yes	20	8.0%	17	6.8%	51	20.5%	161	65%	54%	0.67
No	1042	22%	1022	22%	1012	21.5%	1628	35%	51%	
History of arrhythmia										
Yes	72	10%	78	11%	156	22.1%	401	57%	48%	0.52
No	990	23%	961	23%	907	21.4%	1388	33%	51%	
History of diabetes										
Yes	124	14%	146	16%	202	22.6%	421	47%	46%	0.10
No	938	23%	893	22%	861	21.2%	1368	34%	51%	
History of stroke										
Yes	<10	<10%	29	30%	<10	<10.4%	54	56%	17%	<0.005
No	1056	21%	1010	21%	1056	21.7%	1735	36%	51%	
History of renal disease										
Yes	26	8.7%	35	12%	61	20.3%	178	59%	43%	0.21
No	1036	22%	1004	22%	1002	21.5%	1611	35%	51%	
History of chronic lung disease										
Yes	191	17%	176	16%	262	23.0%	510	45%	52%	0.53
No	871	23%	863	23%	801	21.0%	1279	34%	50%	
NEWS2										
0-2	1002	33%	954	31%	952	31.3%	131	4.3%	51%	0.012
3-4	57	11%	72	14%	93	18.1%	291	57%	44%	
5+	<10	<4.6%	13	5.9%	18	8.2%	185	85%	19%	

Supplementary Table 8: Comparison of key characteristics of this analysis with original derivation of Amb score(1) and Glasgow Admission Prediction Score (GAPS)(3).

	Population	Episode start time	Comparator	Location	Sample size	Study period
This analysis	Unplanned attendances to acute medicine	08:00-16:59, Monday to Friday	Discharged in <12 hours vs admitted for 12-48 hours	Birmingham, UK (single hospital)	7365 episodes	April 2019-March 2020
Amb score – Ala et al, 2012	Unplanned attendances to acute medicine	Unrestricted	Discharged in <12 hours vs admitted for >48 hours	South Wales, UK (single hospital)	625 episodes (derivation: 282, validation: 343)	May-June 2010 (derivation), June-July 2011 (validation)
GAPS score – Cameron et al, 2015	Unplanned attendances to Emergency Department, acute medicine, or minor injuries unit	Unrestricted	Clinical decision to discharge vs clinical decision to admit to hospital	North Glasgow, UK (3 hospitals)	322,846 episodes (derivation: 215,231, validation: 107,615	March 2010-March 2012

References

1. Royal College of Physicians. National Early Warning Score (NEWS) 2. 2017.
2. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
3. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1 & 2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5
	11	Rationale for choosing the reference standard (if alternatives exist)	5
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	5
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5/6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	5/6
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6
	15	How indeterminate index test or reference standard results were handled	6
	16	How missing data on the index test and reference standard were handled	6
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6
	18	Intended sample size and how it was determined	-
RESULTS			
Participants	**19**	Flow of participants, using a diagram	-
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	8
	22	Time interval and any clinical interventions between index test and reference standard	-
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	12,14
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	12,14
	25	Any adverse events from performing the index test or the reference standard	-
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	18
	27	Implications for practice, including the intended use and clinical role of the index test	17-18
OTHER INFORMATION			
	28	Registration number and name of registry	-
	29	Where the full study protocol can be accessed	-
	30	Sources of funding and other support; role of funders	21

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-064910.R3
Article Type:	Original research
Date Submitted by the Author:	01-Dec-2022
Complete List of Authors:	Atkin, Catherine; University of Birmingham, Birmingham Acute Care Research Group Gallier, Suzy; University Hospitals Birmingham NHS Foundation Trust, Department of Health Informatics Wallin, Elizabeth; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine Reddy-Kolanu, Vinay; University Hospitals Birmingham NHS Foundation Trust, Acute medicine Sapey, Elizabeth; University of Birmingham, PIONEER HDR-UK Hub; University Hospitals Birmingham NHS Foundation Trust, Acute Medicine
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Evidence based practice
Keywords:	INTERNAL MEDICINE, GENERAL MEDICINE (see Internal Medicine), Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

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 (0000-0003-0596-8515), Suzy Gallier (0000-0003-1026-4125), Elizabeth Wallin, Vinay Reddy-Kolanu, Elizabeth Sapey (0000-0003-3454-5482)

Addresses

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, Edgbaston, Birmingham, B15 2GW, UK. Catherine Atkin, NIHR Academic Clinical Lecturer in Acute Medicine.

Department of Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Suzy Gallier, Lead for Research Analytics.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Wallin, Consultant in Nephrology & Acute Medicine.

Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Vinay Reddy-Kolanu, Consultant in Acute Medicine.

Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, University of Birmingham, and Department of Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham B15 2GW, UK. Elizabeth Sapey, Professor of Acute and Respiratory Medicine and Honorary Consultant.

Correspondence to: Catherine Atkin c.atkin@nhs.net

Number of references: 17

Abstract word count: 271

Word count: 4998

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 and lasting up to 48 hours, between 1st April 2019 and 9th 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 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

- This study compared performance of the Amb score and GAPS 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 NEWS2, reflecting current clinical practice
- Patients admitted for longer than 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 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 less than 12 hours or over 48 hours. That study excluded patients who remained in hospital for 12 to 48 hours.

Table 1: Scoring systems to identify medical admissions potentially suitable for discharge from hospital without admission >12 hours. Amb score(5) and Glasgow Admission Prediction Score (GAPS)(7). Amb score of 5 more indicates likely discharge within 12 hours; GAPS of 16 or more suggests patient likely to be admitted to hospital. IV = intravenous, MEWS = Modified early warning score, NEWS = National Early Warning Score, GP = General practitioner

Amb score		
Sex	Female	0
	Male	-0.5
Age	<80	0
	≥80	-0.5
Access to personal transport/can take public transport	Agree	2
	Disagree	0
IV treatment not anticipated	Agree	2
	Disagree	0
Not acutely confused	Agree	2
	Disagree	0
MEWS=0	Agree	1
	Disagree	0
Not discharged from hospital within previous 30 days	Agree	1
	Disagree	0

Glasgow Admissions Prediction Score (GAPS)		
NEWS		1 point per point on NEWS score
Age		1 point per decade
Triage category	3	5
	2 (or 2+)	10
	1	20
Referred by GP		5
Arrived in ambulance		5
Admitted <1 year ago		5

The Glasgow Admission Prediction Score (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 Same Day Emergency Care and those requiring admission for at least 12 to 48 hours.

Methods

This data study was conducted in collaboration with PIONEER, a Health Data Research Hub in Acute Care, and all study processes were carried out following appropriate ethical approval provided by the East Midlands – Derby REC (reference: 20/EM/0158).

Retrospective data were collected for patients admitted to Queen Elizabeth Hospital Birmingham, University Hospitals Birmingham NHS Trust (UHB) between the period of the 1st April 2019 until 9th 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 (CCU) in Europe. The Acute Medical Unit (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 inhouse electronic health record (EHR) called Prescribing Information and Communication System (PICS). 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 focussed 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 and mapped ICD10 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 to 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 08:00-16:59, 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 (IV) treatment was taken as not being anticipated where patients did not receive an IV 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 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 National Early Warning Score 2 (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 15.1. Cell counts containing less than 10 patients were suppressed, due to reporting requirements. For univariate analysis of factors influencing likelihood of discharge within 12 hours, odds ratios 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 Chi square. A p value of <0.05 is 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 less than 48 hours, based on previous research.(10)

Results

14314 acute medical inpatient episodes lasting up to 48 hours were identified during the censor period. These episodes were from 12587 patients with 11229 patients having one episode in this time period. Patients 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.

Table 2: Demographics and characteristics of patients with emergency medical admissions lasting up to 48 hours. For whole cohort, and for patients arriving in a normal working day (08:00-16:59, Monday to Friday).P values shown for Chi square comparison of normal working day episodes to episodes starting outside normal working day.

	All episodes N=14314		Normal working day episodes N=7365		Episodes starting outside normal working day N= 6949		P value
	Frequency (%)		Frequency (%)		Frequency (%)		
Age							
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 day	1047	(7.3%)	479	(6.5%)	568	(8.2%)	<0.001
14 day	1544	(11%)	681	(9%)	863	(12%)	<0.001
30 day	2268	(16%)	1033	(14%)	1235	(18%)	<0.001

18% of episodes occurred on a weekend. Overall, 62% of patients arrived between 08:00-16:59 (Figure 1); 63% of weekday episodes started between these times.

11244 episodes had an associated Emergency Department triage code, with 108 different triage codes used. The commonest triage problem was chest pain (34% of episodes), see Supplementary Table 1. 6394 episodes (44%) had a length of stay of less than 12 hours.

Normal working day arrivals

There were 7365 episodes in 6848 patients with an arrival time between 08:00-16:59 on a weekday (normal working day, NWD). The triage problem was available for 5272 NWD episodes (72%). The commonest triage problem was chest pain (37%) (Supplementary Table 1).

4053 episodes (55%) had a length of stay of less than 12 hours and 3312 (45%) were discharged after 12 to 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%, Chi square p<0.005).

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

Compared to patients discharged within 12 to 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, Chi square 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, with comparison of . Age ≥80 and anticipated need for IV therapy were associated with an increased risk of admission lasting more than 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 50-140bpm and systolic blood pressure between 100-200mmHg 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 ACS 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).

*Table 3: Factors considered in derivation of previous scoring systems. Column percentages shown. Univariate analysis, odds ratio for admission lasting 12-48 hours shown. IV: intravenous; RR: respiratory rate; HR: heart rate in beats per minute; SBP: systolic blood pressure in mmHg; MEWS: Modified Early Warning Score; NEWS2: National Early Warning Score 2(9); IHD: ischaemic heart disease; GP: general practice. 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.*

N=7365 unless otherwise stated	Length of stay				Odds ratio (OR)	P value	95% CI OR
	<12hrs		12-48 hours				
	Frequency (%)		Frequency (%)				
Age							
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
IV 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 *(n=6748)*	3499	(99.0%)	3144	(97.9%)	0.42	<0.001	0.25 to 0.69
SBP 100-200 *(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
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

Amb score

Multivariable analysis including all components of the Amb score, except access to transportation (which was present for all patients), is shown in Supplementary 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 (Supplementary Table 3). 94% (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 less than 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 to 48 hours).

*Table 4: Amb score performance. Performance in normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	Amb score N=6743		Amb score with NEWS2 N=6707	
	Frequency (%)		Frequency (%)	
Score <5	418	(6.2%)	364	(5.4%)
Score 5+	6325	(93.8%)	6343	(94.6%)
Score <5 Admission length <12hrs	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%)	

Replacing MEWS with NEWS2, the AUROC was 0.612 (95% CI 0.599 to 0.625)(Figure 2b). 95% (6343 admissions) had an Amb score of 5 or more; 5.4% (364 admissions) had a score of less than 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%, Chi square p=0.017), in both those discharged within 12 hours (13.7% vs 5.8%, p=0.017) and those discharged in 12 to 48 hours (11.7% vs 7.0%, p=0.001). This was also true for readmission within 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 to 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%, Chi square 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 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 (Supplementary 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, and where comorbidity due to ischaemic heart disease, heart failure, arrhythmia, diabetes, stroke/TIA, renal disease or chronic lung disease was present. A higher proportion of GP referrals with a raised Amb score were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (69% vs 45%, Chi square $p<0.005$). A higher proportion of patients with a raised Amb score and a NEWS2 of 0-2 were identified correctly compared to those with a raised NEWS2 on arrival.

GAPS

Multivariable analysis including all components of the GAPS is shown in Supplementary 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 more than 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 (Supplementary 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 48.5% to 52.5%), with a 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 to 48 hours).

*Table 5: GAPS performance within normal working day admissions. PPV: positive predictive value; NPV: negative predictive value. NEWS2: National Early Warning Score 2.(9) *(1-sensitivity); ^(1-PPV)*

	GAPS N=5091	GAPS with NEWS2 N=4953
	Frequency (%)	Frequency (%)
Score ≤15 16+	 2179 (42.8%) 2912 (57.2%)	 2101 (42.4%) 2852 (57.6%)
Score ≤15 Admission length <12hrs	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%)

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).

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 to 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%, Chi square p<0.005), both for those discharged within 12 hours (6.0% vs 4.2%, p=0.055), and 12 to 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 to 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. 29 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, were female, and where comorbidity due to stroke/TIA was present (Supplementary Table 7). A higher proportion of GP referrals with a low GAPS were discharged within 12 hours, compared to those whose first healthcare contact was the emergency department (68% vs 50%, Chi square p=0.044). A higher proportion of patients with a low GAPS and a NEWS2 of 0-2 were identified correctly compared to 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 (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, 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 a Amb score suggesting suitability for SDEC with a GAPS suggesting admission (Chi square, p<0.0005 for each subgroup comparison, Figure 4).

Discussion

This paper highlights several important points. Firstly, 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 to 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 negative predictive value than the GAPS, with 88.5% of patients with a low Amb score (suggesting they were unsuitable for SDEC) remaining for more than 12 hours, compared to 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 (Supplementary 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 over 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, 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 more than 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 focussed on performance of scoring systems to identify patients suitable for SDEC within currently available services; in-depth evaluation of factors necessitating admission over 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 focussed on the ability of the Amb score and GAPS to discriminate between those admitted for <12 hours and 12 to 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 positive and negative

predictive value of the score. Although some aspects of score performance may be appear improved if the scores are able to identify all those admitted for over 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 positive predictive value 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 IV treatment to be 'anticipated' in patients receiving IV treatment within 6 hours of arrival, as anticipation of IV 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 was 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 Glasgow Admission Prediction Score 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 Same Day Emergency Care services.

Contributorship

CA and ES designed the study, CA analysed the data, all authors (CA, SG, EW, VRK, ES) 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.

License for publication

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.

Data sharing agreement
Data from this study is 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.

Ethics statement
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 HDR UK Hub in acute care by the East Midlands – Derby REC (reference: 20/EM/0158). Formal written consent from individual participants was not required.

Transparency declaration
The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Funding
No specific funding was available for this project.

References

1. NHS England. A&E Attendances and Emergency Admissions 2021-22 2021 [Available from: 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? Clinical Medicine. 2022:clinmed.2021-0614.
5. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
6. Royal College of Physicians. Acute Care Toolkit 10: Ambulatory Emergency Care. 2014.
7. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.
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, Holland M, Subbe C, Kennedy A, et al. Length of stay in Acute Medical Admissions: Analysis from the Society for Acute Medicine Benchmarking Audit. Acute Med. 2022;21(1):27-33.
11. Cameron A, Jones D, Logan E, O'Keeffe CA, Mason SM, Lowe DJ. Comparison of Glasgow Admission Prediction Score and Amb Score in predicting need for inpatient care. Emerg Med J. 2018;35(4):247-51.
12. Pendlebury ST, Lovett NG, Smith SC, Dutta N, Bendon C, Lloyd-Lavery A, 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(11):e007808.
13. Society for Acute Medicine. Society for Acute Medicine Benchmarking Audit 2021 - SAMBA2021 Report. 2021.
14. Atkin C, Knight T, Subbe C, Holland M, Cooksley T, Lasserson D. Acute care service performance during winter: report from the winter SAMBA 2020 national audit of acute care. Acute Med. 2020;19(4):220-9.
15. Baier N, Geissler A, Bech M, Bernstein D, Cowling TE, Jackson T, 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):1-10.
16. Canadian Institute for Health Information. NACRS emergency department visits and lengths of stay 2022 [Available from: https://www.cihi.ca/en/nacrs-emergency-department-visits-and-lengths-of-stay.
17. Australian Institute of Health and Welfare. Admitted patients. 2022 [Available from: https://www.aihw.gov.au/reports-data/myhospitals/sectors/admitted-patients.

Figure Legends

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

Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions. AUROC: area under the receiver operating characteristic curve; 95% CI: 95% confidence interval.

Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, 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.

Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

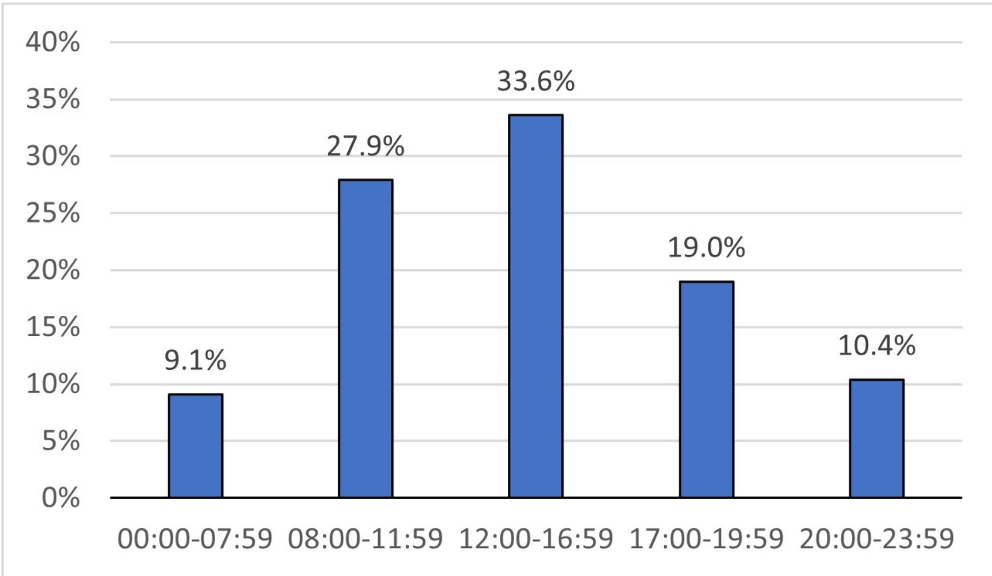


Figure 1: Arrival time for medical attendances lasting up to 48 hours.

99x57mm (330 x 330 DPI)

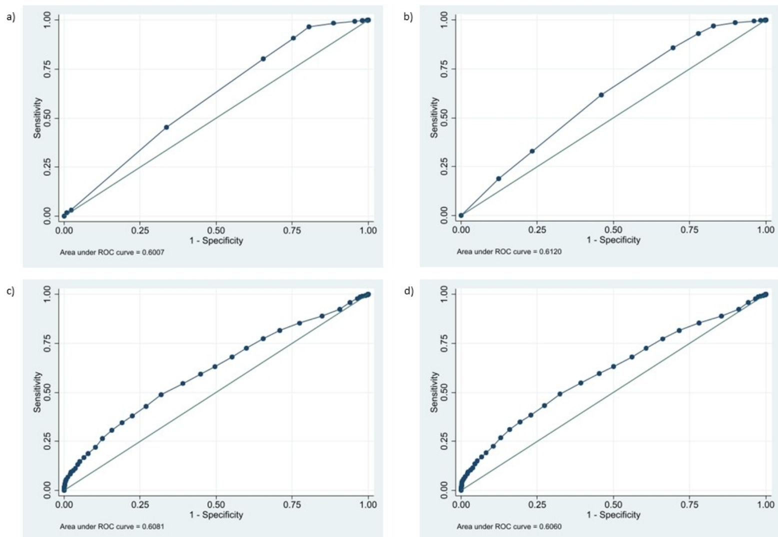


Figure 2: Receiver operator characteristics (ROC) curve for score performance. A) Amb score; b) Amb score substituting NEWS2; c) GAPS; d) GAPS substituting NEWS2. Performance in identifying patients with length of stay <12 hours in normal working day admissions. AUROC: area under the receiver operating characteristic curve; 95% CI: 95% confidence interval.

338x190mm (96 x 96 DPI)

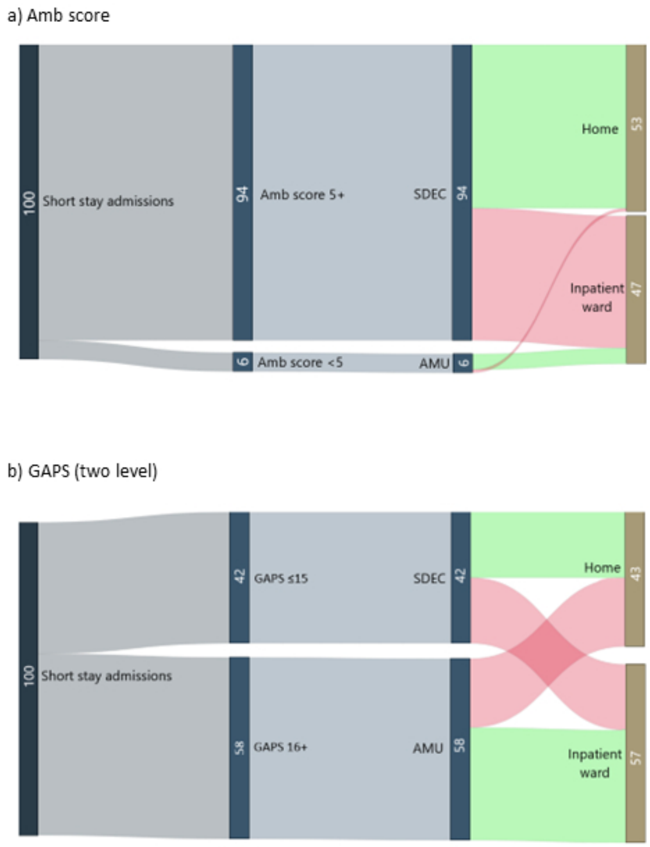


Figure 3: Sankey diagram estimating patient pathways through acute medical services for short stay medical admissions when utilising scoring systems to identify patients for assessment in Same Day Emergency Care, for a) Amb score (5 or more) and b) Glasgow Admission Prediction Score (GAPS)(≤15). Green = currently identified by scoring system, red = incorrectly identified by scoring system.

481x476mm (38 x 38 DPI)

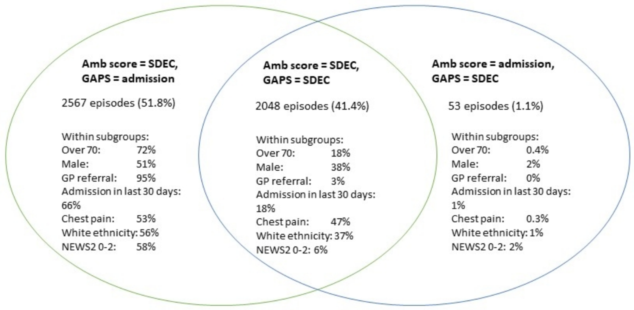


Figure 4: Agreement of Amb score and GAPS score in identification of patients suitable for SDEC. Within each patient subgroup, the percentage of patients where the Amb score and GAPS suggested suitability for SDEC is shown.

338x190mm (96 x 96 DPI)

Supplementary Table 1: Triage problem. Commonest triage problem recorded on arrival to Emergency Department. Coded presenting problem entered at initial Emergency Department triage. Normal working day admissions defined as episodes starting between 08:00-16:59 Monday-Friday.

All admissions			**Normal working day admissions**		
	Frequency (%)			Frequency (%)	
Chest pain	3762	(34%)	Chest pain	1940	(37%)
Dyspnoea/difficulty breathing	1586	(14%)	Dyspnoea/difficulty breathing	721	(14%)
Asthenia	1051	(9.4%)	Asthenia	548	(10%)
Headache	609	(5.4%)	Headache	322	(6.1%)
Abdominal pain	408	(3.6%)	Abdominal pain	172	(3.3%)
Near syncope/syncope	282	(2.5%)	Palpitations	145	(2.8%)
Palpitations	256	(2.3%)	Near syncope/syncope	137	(2.6%)
Dizziness	222	(2.0%)	Dizziness	119	(2.3%)
Fever	210	(1.9%)	Pain in lower limb	96	(1.8%)
Substance abuse	210	(1.9%)	Vomiting	82	(1.6%)

Supplementary Table 2: Multivariable analysis of Amb score components. Mixed-effects logistic regression, patient as random effect. Odds ratio for admission of 12-48 hours, normal working day admissions. IV= intravenous, MEWS= Modified Early Warning Score, NEWS2= National Early Warning Score 2.(1)

Amb score components				Amb score components, substituting NEWS2			
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age >80	2.03	<0.001	1.71 to 2.41	Age >80	2.01	<0.001	1.69 to 2.38
Male	1.03	0.59	0.92 to 1.16	Male	1.02	0.735	0.91 to 1.14
IV treatment not anticipated	0.10	<0.001	0.7 to 0.13	IV treatment not anticipated	0.12	<0.001	0.07 to 0.14
Not acutely confused	0.32	0.06	0.10 to 1.04	Not acutely confused	0.35	0.08	0.11 to 1.13
MEWS 0	1.06	0.73	0.77 to 1.43	NEWS2 0	0.81	<0.001	0.72 to 0.91
Not discharged in last 30 days	1.00	0.96	0.84 to 1.18	Not discharged in last 30 days	1.01	0.94	0.85 to 1.19

Supplementary table 3: Amb score for NWD (Normal working day) admission episodes. Normal working day defined as episodes starting between 08:00-16:59 Monday-Friday. Amb score calculated as shown in Table 1.(2) NEWS2: National Early Warning Score 2.(1)

	Amb score		Amb score substituting NEWS2	
Amb score	Number of episodes (%)		Number of episodes (%)	
≤3	12	(0.2%)	12	(0.2%)
3.5	51	(0.8%)	44	(0.7%)
4	98	(1.5%)	81	(1.2%)
4.5	257	(3.8%)	227	(3.4%)
5	327	(4.9%)	287	(4.3%)
5.5	367	(5.4%)	295	(4.4%)
6	690	(10.2%)	522	(7.8%)
6.5	2261	(33.5%)	1605	(23.9%)
7	2502	(37.1%)	1735	(12.6%)
7.5	94	(1.4%)	846	(15.7%)
8	84	(1.3%)	1053	(12.3%)

Supplementary Table 4: Identifying length of admission by Amb score (incorporating NEWS2) within patient subgroups. Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Amb score calculated as per Table 1, with NEWS2 substituted in place of MEWS. NEWS2: National Early Warning Score 2.(1) MEWS: Modified Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for comparisons using Chi square.

	Amb 5+, Admission length <12hrs Correctly identified		Amb 5+, Admission length 12-48 hrs Incorrectly identified		Amb <5, Admission length <12 hours Incorrectly identified		Amb <5, Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by Amb score discharged within 12 hours	P value
Percentage of admissions	52%		43%		0.6%		4.8%		55%	
	N	%	N	%	N	%	N	%		
Age										
16-19	85	523%	70	43%	<10	<6.2%	<10	<6.2%	55%	<0.005
20-29	340	51%	291	44%	<10	<1.5%	28	4.2%	54%	
30-39	404	54%	310	41%	<10	<1.3%	27	3.6%	57%	
40-49	465	57%	330	40%	<10	<1.2%	20	2.4%	59%	
50-59	630	56%	445	40%	12	1.1%	38	3.4%	59%	
60-69	564	58%	370	38%	<10	<1.0%	38	3.9%	60%	
70-79	547	50%	506	46%	<10	<0.9%	51	4.6%	52%	
80-89	357	41%	426	50%	<10	<1.2%	69	8.0%	46%	
90+	67	27%	136	558%	0	-	45	18%	33%	
Under 70	3035	54%	2322	42%	33	0.6%	208	3.7%	57%	<0.005
Over 70	424	38%	562	51%	<10	<0.9%	114	10%	43%	
Sex										
Female	2022	52%	1749	45%	12	0.3%	94	2.4%	54%	0.08
Male	1437	50%	1135	40%	30	1.1%	228	8.1%	56%	
Ethnicity										
Asian	500	51%	440	45%	<10	<1.0%	26	2.7%	53%	0.19
Black	169	56%	122	40%	<10	<3.3%	10	3.3%	58%	
Unknown	395	54%	287	39%	11	1.5%	38	5.2%	58%	
Mixed	58	51%	48	43%	<10	<8.8%	<10	<8.8%	55%	
Other	103	58%	72	40%	0		<10	<5.6%	59%	
White	2234	51%	1915	43%	23	1.0%	239	5.4%	54%	
Recent admission (30 days)										
Yes	433	50%	335	39%	11	1.3%	81	9.4%	56%	0.27
No	3026	52%	2549	44%	31	0.5%	241	4.1%	54%	
GP referral										
Yes	1792	67%	823	31%	10	0.4%	39	1.5%	69%	<0.005
No	1667	41%	2061	51%	32	0.8%	283	7.0%	45%	
Chest pain as triage problem										
Yes	1032	58%	739	41%	<10	<0.6%	12	0.7%	58%	<0.005
No	2427	49%	2145	44%	35	0.7%	310	6.3%	53%	
History of IHD										
Yes	834	50%	766	46%	<10	<0.6%	69	4.1%	52%	0.025
No	2625	52%	2118	42%	33	0.7%	253	5.0%	55%	
History of HF										
Yes	111	36%	167	54%	<10	<3.2%	27	8.8%	40%	<0.005
No	3348	52%	2717	43%	39	0.6%	295	4.6%	55%	
History of arrhythmia										
Yes	323	38%	438	51%	<10	<1.2%	83	9.7%	42%	<0.005
No	3136	54%	2446	42%	33	0.6%	239	4.1%	56%	
History of diabetes										
Yes	497	44%	546	48%	<10	<0.9%	79	7.0%	48%	<0.005
No	2962	53%	2338	42%	35	0.6%	243	4.4%	56%	
History of stroke										
Yes	18	18%	80	79%	0	-	<10	<10%	18%	<0.005
No	3441	52%	2804	42%	42	0.6%	319	4.8%	55%	
History of renal disease										
Yes	167	41%	197	48%	0	-	46	11%	46%	<0.005
No	3292	52%	2687	43%	42	0.7%	276	4.4%	55%	
History of chronic lung disease										
Yes	703	48%	674	46%	12	0.8%	92	6.2%	52%	<0.005
No	2756	53%	2210	42%	32	0.6%	230	4.4%	56%	
NEWS2										
0-2	3180	55%	2435	42%	29	0.5%	162	2.8%	57%	<0.005
3-4	252	38%	319	48%	<10	<1.5%	85	13%	44%	
5+	27	11%	130	55%	<10	<4.2%	75	32%	17%	

Supplementary Table 5: Multivariable analysis of GAPS components. Mixed-effects logistic regression, patient as random effect. Age – odds ratio (OR) per decade increase in age; NEWS/NEWS2 OR per increase of one point in NEWS/NEWS2. Triage category compared to 'standard' as reference. Odds ratio for admission of 12-48 hours, normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). GP= general practitioner, NEWS= national early warning score

	GAPS				GAPS with NEWS2		
	Adjusted OR	P value	95% CI		Adjusted OR	P value	95% CI
Age	1.07	<0.001	1.03 to 1.10	Age	1.07	<0.001	1.03 to 1.10
NEWS	1.25	<0.001	1.18 to 1.32	NEWS2	1.22	<0.001	1.16 to 1.29
Triage category* Urgent Resuscitation	 1.08 4.64	 0.46 <0.001	 0.88 to 1.33 2.88 to 7.46	Triage category* Urgent Resuscitation	 1.04 4.32	 0.69 <0.001	 0.84 to 1.29 2.68 to 6.95
Referred by GP	0.79	0.002	0.69 to 0.92	Referred by GP	0.78	0.001	0.67 to 0.90
Arrived in ambulance	1.62	<0.001	1.40 to 1.86	Arrived in ambulance	1.61	<0.001	1.40 to 1.86
Admitted <1 year ago	1.42	<0.001	1.24 to 1.61	Admitted <1 year ago	1.40	<0.001	1.22 to 1.60

Supplementary Table 6: GAPS for normal working day admissions. GAPS: Glasgow Admission Prediction Score, calculated as described in Table 1.(3) NEWS2: National Early Warning Score 2.(1)

	GAPS score N=5091		GAPS score substituting NEWS2 N=4953	
GAPS score	Number of episodes (%)		Number of episodes (%)	
1-5	93	(1.8%)	88	(1.8%)
6-19	829	(16.3%)	792	(16.0%)
11-15	1257	(24.7%)	1221	(24.7%)
16-20	1329	(26.1%)	1279	(15.8%)
21-25	874	(17.2%)	857	(17.3%)
26-30	354	(7.0%)	360	(7.3%)
31-35	211	(4.1%)	206	(4.2%)
36-40	97	(1.9%)	94	(1.9%)
41-45	41	(0.8%)	45	(0.9%)
46+	<10	(<0.2%)	11	(0.2%)

Supplementary Table 7: Identifying length of admission by GAPS (incorporating NEWS2) within patient subgroups. Analysis of Normal working day admissions (episodes starting 08:00-16:59, Monday-Friday). Glasgow Admission Prediction Score (GAPS) calculated as per Table 1, with NEWS2 substituted in place of NEWS. NEWS2: National Early Warning Score 2.(1) NEWS: National Early Warning Score. SDEC: Same Day Emergency Care. GP: general practice; IHD: Ischaemic heart disease; HF: heart failure. Presence of chest pain as recorded on initial Emergency Department triage. P values shown for Chi square comparisons.

GAPS with NEWS2	GAPS≤15 Admission length <12hrs Correctly identified		GAPS ≤15 Admission length 12-48 hrs Incorrectly identified		GAPS 16+, Admission length <12 hours Incorrectly identified		GAPS 16+ Admission length 12-48hrs Correctly identified		Proportion 'SDEC suitable' by GAPS discharged within 12 hours	P value
Percentage of admissions	21%		21%		22%		36%		50%	
Age (years)										
16-19	32	26%	48	40%	17	14%	24	20%	40%	<0.005
20-29	140	27%	180	34%	91	17%	113	22%	44%	
30-39	172	30%	185	32%	101	17%	123	21%	48%	
40-49	228	35%	178	28%	104	16%	135	21%	56%	
50-59	237	28%	191	22%	188	22%	235	28%	55%	
60-69	126	18%	102	15%	216	31%	253	36%	55%	
70-79	87	11%	89	12%	201	26%	385	51%	49%	
80-89	33	5.7%	59	10%	122	21%	361	63%	34%	
90+	<10	<5.1%	<10	<5.1%	23	12%	160	81%	50%	
Under	1022	24%	973	23%	918	22%	1268	30%	51%	0.007
Over 70	40	5.2%	66	8.5%	145	19%	521	68%	38%	
Sex										
Female	599	21%	633	23%	597	21.3%	977	35%	48%	0.035
Male	463	22%	406	19%	466	21.7%	811	38%	53%	
Ethnicity										
Asian	223	28%	188	23%	157	19.4%	241	30%	54%	0.25
Black	57	26%	48	22%	48	22.0%	65	30%	54%	
Unknown	135	26%	127	24%	102	19.4%	161	31%	52%	
Mixed	20	22%	27	30%	20	22.0%	24	26%	43%	
Other	35	25%	47	33%	37	25.9%	24	17%	43%	
White	592	19%	602	19%	699	22.1%	1274	40%	50%	
Recent admission (30 days)										
Yes	45	8.7%	55	11%	122	23.6%	295	57%	45%	0.26
No	1017	23%	984	22%	941	21.2%	1494	34%	51%	
GP referral										
Yes	23	2.1%	11	1.0%	533	49.6%	508	47%	68%	0.044
No	1039	27%	1028	27%	530	13.7%	1281	33%	50%	
Chest pain as triage problem										
Yes	523	29%	318	18%	516	28.8%	433	24%	62%	<0.005
No	539	17%	721	23%	547	17.3%	1356	43%	43%	
History of IHD										
Yes	306	20%	231	15%	402	26.7%	568	38%	57%	<0.005
No	756	22%	808	23%	661	19.2%	1221	35%	48%	
History of heart failure										
Yes	20	8.0%	17	6.8%	51	20.5%	161	65%	54%	0.67
No	1042	22%	1022	22%	1012	21.5%	1628	35%	51%	
History of arrhythmia										
Yes	72	10%	78	11%	156	22.1%	401	57%	48%	0.52
No	990	23%	961	23%	907	21.4%	1388	33%	51%	
History of diabetes										
Yes	124	14%	146	16%	202	22.6%	421	47%	46%	0.10
No	938	23%	893	22%	861	21.2%	1368	34%	51%	
History of stroke										
Yes	<10	<10%	29	30%	<10	<10.4%	54	56%	17%	<0.005
No	1056	21%	1010	21%	1056	21.7%	1735	36%	51%	
History of renal disease										
Yes	26	8.7%	35	12%	61	20.3%	178	59%	43%	0.21
No	1036	22%	1004	22%	1002	21.5%	1611	35%	51%	
History of chronic lung disease										
Yes	191	17%	176	16%	262	23.0%	510	45%	52%	0.53
No	871	23%	863	23%	801	21.0%	1279	34%	50%	
NEWS2										
0-2	1002	33%	954	31%	952	31.3%	131	4.3%	51%	0.012
3-4	57	11%	72	14%	93	18.1%	291	57%	44%	
5+	<10	<4.6%	13	5.9%	18	8.2%	185	85%	19%	

Supplementary Table 8: Comparison of key characteristics of this analysis with original derivation of Amb score(1) and Glasgow Admission Prediction Score (GAPS)(3).

	Population	Episode start time	Comparator	Location	Sample size	Study period
This analysis	Unplanned attendances to acute medicine	08:00-16:59, Monday to Friday	Discharged in <12 hours vs admitted for 12-48 hours	Birmingham, UK (single hospital)	7365 episodes	April 2019-March 2020
Amb score – Ala et al, 2012	Unplanned attendances to acute medicine	Unrestricted	Discharged in <12 hours vs admitted for >48 hours	South Wales, UK (single hospital)	625 episodes (derivation: 282, validation: 343)	May-June 2010 (derivation), June-July 2011 (validation)
GAPS score – Cameron et al, 2015	Unplanned attendances to Emergency Department, acute medicine, or minor injuries unit	Unrestricted	Clinical decision to discharge vs clinical decision to admit to hospital	North Glasgow, UK (3 hospitals)	322,846 episodes (derivation: 215,231, validation: 107,615	March 2010-March 2012

References

1. Royal College of Physicians. National Early Warning Score (NEWS) 2. 2017.
2. Ala L, Mack J, Shaw R, Gasson A, Cogbill E, Marion 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 (Lond). 2012;12(5):420-6.
3. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emergency Medicine Journal. 2015;32(3):174.

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1 & 2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	5
	11	Rationale for choosing the reference standard (if alternatives exist)	5
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	5
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5/6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	5/6
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6
	15	How indeterminate index test or reference standard results were handled	6
	16	How missing data on the index test and reference standard were handled	6
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6
	18	Intended sample size and how it was determined	-
RESULTS			
Participants	**19**	Flow of participants, using a diagram	-
	20	Baseline demographic and clinical characteristics of participants	8
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	8
	22	Time interval and any clinical interventions between index test and reference standard	-
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	12,14
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	12,14
	25	Any adverse events from performing the index test or the reference standard	-
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	18
	27	Implications for practice, including the intended use and clinical role of the index test	17-18
OTHER INFORMATION			
	28	Registration number and name of registry	-
	29	Where the full study protocol can be accessed	-
	30	Sources of funding and other support; role of funders	21

STARD 2015

AIM

STARD stands for “Standards for Reporting Diagnostic accuracy studies”. This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or “2x2” table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.