Cureus
Part of SPRINGER NATURE

Open Access Original Article

From Variation to Standardisation: An Electronic Health Record Intervention to Improve Risk Stratification and Admission Decisions in Acute Lower Gastrointestinal Bleeding

Review began 10/03/2026
Review ended 10/20/2026
Published 10/28/2026

© **Copyright** 2026
Bin Rosli. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118748

Muhammad Danial Bin Rosli [1]

1. Department of General Surgery, Altnagelvin Area Hospital, Derry, GBR

Corresponding author: Muhammad Danial Bin Rosli, 97danial1@gmail.com

Abstract

Background

Acute lower gastrointestinal bleeding (LGIB) is a common surgical emergency with wide variation in management despite British Society of Gastroenterology (BSG) guidance. The Oakland score is validated to identify patients suitable for safe discharge, but uptake in routine practice is inconsistent. Where a validated score is not used, disposition reverts to unstructured judgement, risking both premature discharge of high-risk patients and unnecessary admission of low-risk ones. Against a guideline-derived standard of 100%, we aimed to improve documented risk stratification and the appropriateness of admission and discharge decisions.

Methods

We conducted a two-cycle retrospective closed-loop audit of consecutive adult LGIB presentations at a district general hospital, taking the bleeding episode as the unit of analysis (28 episodes in cycle 1, 22 in cycle 2). Between cycles, we introduced a multifaceted intervention comprising clinician education, ward posters, a departmental email, and an electronic health record SmartPhrase that assembles the seven components of the Oakland score at clerking, populating four automatically from the patient record. Primary outcomes were documentation of the Oakland score and concordance of the admission or discharge decision with the patient's Oakland risk category. Proportions were compared by Fisher's exact test using identical denominators in both cycles, with absolute differences and 95% confidence intervals by the Newcombe method.

Results

Oakland score documentation rose from 6/28 (21.4%) to 19/22 (86.4%), an absolute increase of 64.9% (95% CI 38.2 to 79.2; $p < 0.001$). Decisions concordant with risk category rose from 21/27 (77.8%) to 20/20 (100%), an increase of 22.2% (95% CI 2.4 to 40.8; $p = 0.031$); the six cycle 1 discordances comprised three premature discharges (Oakland scores 9, 16 and 22) and three unnecessary admissions. Outpatient investigation was arranged in 9/22 (40.9%) versus 13/18 (72.2%), a difference of 31.3% (95% CI 0.5 to 54.7; $p = 0.062$). Transfusion fell within the restrictive threshold in 3/5 versus 8/8 transfused episodes (difference 40.0%, 95% CI -3.0 to 76.9; $p = 0.128$), with two inappropriate transfusions in cycle 1 and none in cycle 2. Anticoagulant management (12/15 versus 12/13; difference 12.3%, 95% CI -16.4 to 38.3; $p = 0.60$), in which the absolute number managed correctly was unchanged, and CT angiography for shock index >1 (2/4 versus 1/1; $p = 1.00$) showed no significant change.

Conclusions

Embedding risk stratification in the clerking workflow was associated with a substantial increase in documented Oakland score use, and no admission or discharge decision discordant with calculated risk was observed in the second cycle. The intervention was multifaceted and introduced over a short interval in a small single-centre cohort, so the contribution of the SmartPhrase cannot be isolated and causal attribution is not possible. The approach is low-cost, trainee-led and reproducible wherever a configurable electronic health record is used; confirmation that the effect is sustained will require a further cycle including patient-level outcomes.

Categories: General Surgery, Quality Improvement
Keywords: clinical decision support, closed-loop audit, electronic health record, lower gastrointestinal bleeding, oakland score, quality improvement

Introduction

Acute lower gastrointestinal bleeding (LGIB) accounts for a substantial share of emergency general surgical

referrals and affects a predominantly elderly, comorbid population in whom anticoagulant and antiplatelet use is common [1,2]. Most episodes settle spontaneously, and a considerable proportion of patients could be investigated safely as outpatients. The clinical difficulty lies in separating those patients from the minority with major bleeding who require admission, urgent imaging and intervention.

The Oakland score was derived and validated to address this question precisely, identifying patients who may be discharged safely from the emergency department [3]. It has since been externally validated in a separate health system [4], and is recommended in the 2019 British Society of Gastroenterology (BSG) guideline on acute LGIB alongside recommendations for CT angiography in haemodynamically unstable patients, restrictive transfusion thresholds, and timely definitive investigation [1]. Comparable recommendations appear in contemporary American and European guidance [5-7].

Despite this, uptake of formal risk stratification in routine practice remains poor, and the first UK nationwide audit of LGIB documented wide variation in both investigation and disposition [2]. Where a validated score exists but is not used, the admission decision reverts to unstructured clinical judgement, with two failure modes: high-risk patients discharged prematurely, and low-risk patients admitted unnecessarily. The first is a patient safety concern; the second consumes beds without benefit.

Educational interventions alone are known to produce modest and poorly sustained change, whereas decision support delivered automatically within the clinician's existing workflow is among the features most strongly associated with success [8-11]. We therefore undertook a closed-loop audit at a district general hospital to assess adherence to BSG guidance and to test whether embedding the Oakland score directly into the electronic clerking process would improve risk stratification and the appropriateness of the resulting admission decision.

Two primary outcomes were specified: the proportion of episodes in which an Oakland score was documented by the clerking clinician, and the proportion of admission or discharge decisions concordant with the patient's Oakland risk category. Four secondary outcomes were specified: arrangement of outpatient investigation, adherence to the restrictive transfusion threshold, appropriate interruption or reversal of anticoagulant or antiplatelet therapy, and CT angiography in haemodynamically unstable patients. BSG guidance states that all patients presenting with LGIB should be risk stratified and managed accordingly, and does not define a tolerated failure rate; an audit standard of 100% was therefore set for each measure, with the recognition that this represents an aspirational rather than an empirically derived target [1].

Materials And Methods

Setting and design

This was a two-cycle retrospective closed-loop audit conducted in the Department of General Surgery at Altnagelvin Area Hospital, a district general hospital within the Western Health and Social Care Trust, Northern Ireland. The study is reported in accordance with the SQUIRE 2.0 guidelines for quality improvement reporting [12].

Cycle 1 covered consecutive presentations between 1 August and 27 September 2025. The intervention was introduced immediately thereafter, and cycle 2 covered presentations between 1 October and 28 November 2025.

Case identification, participants and unit of analysis

Cases were identified from the general surgical take records for the audit periods, cross-referenced against the EPIC electronic health record (Epic Systems Corporation, Verona, WI), and comprised all patients referred to and accepted by the general surgical service from the emergency department or from inpatient wards with a presenting problem of per-rectal bleeding or suspected LGIB. All consecutive referrals during each period were screened; no sampling was applied.

Adults presenting with visible per-rectal bleeding or clinically suspected LGIB were eligible. Patients with an identified upper gastrointestinal source at presentation, elective admissions for planned procedures, and records with insufficient documentation for core data extraction were excluded.

Thirty referrals were screened in cycle 1 and 25 in cycle 2, of which two and three, respectively, were excluded, leaving 28 and 22 episodes for analysis.

The unit of analysis was the bleeding episode rather than the individual patient. Five patients presented on two separate occasions during the study period (three in cycle 1, two in cycle 2); because each presentation generated an independent risk assessment and disposition decision, these were retained as distinct episodes. Cycle 1 therefore comprised 28 episodes among 25 patients, and cycle 2 comprised 22 episodes among 20 patients.

Data collection

Data were extracted retrospectively from the EPIC electronic health record and from surgical take records, and were managed in Microsoft Excel (Microsoft Corporation, Redmond, WA). Variables collected were age, sex, date of presentation, admission or discharge decision, shock index, whether an Oakland score was documented by the clerking clinician, the Oakland score itself, CT angiography, interventional radiology, haemoglobin concentration below 70 g/L, red cell transfusion, inpatient endoscopy, surgical intervention, anticoagulant or antiplatelet therapy at presentation and whether it was interrupted or reversed, and whether outpatient investigation was arranged.

Data extraction was performed by the first author, who also independently calculated an Oakland score for every episode from the recorded clinical parameters. No independent duplicate extraction or scoring by a second assessor was undertaken, and inter-rater reliability was therefore not assessed. Missing or unrecorded values were not imputed; episodes in which a required variable was unavailable were excluded from the analysis of that outcome only, and the resulting denominators are reported in full alongside each result.

A single Oakland score was held for each episode and used for all risk categorisation. This score was calculated by the audit team from the recorded clinical parameters for every episode in both cycles, independently of whether, and of what, the clerking clinician had documented. Scores documented by clinicians were used solely to determine the documentation outcome and were not used for risk categorisation. Applying a single, uniformly derived score across both cycles means that the appropriateness of the admission decision was assessed on an identical basis before and after the intervention, and that this outcome is not confounded by the concurrent rise in documentation.

Intervention

Following cycle 1, a multifaceted intervention was implemented. Its components were a departmental teaching session for foundation doctors, senior house officers and registrars; a colour-coded management pathway poster displayed in the emergency department, surgical assessment unit and ward doctors' areas; a targeted email to all surgical and emergency clinicians summarising the pathway; and, as the principal component, a bespoke SmartPhrase built within the EPIC electronic health record.

The SmartPhrase is invoked by typing a full stop followed by the character string "Oakland" at any point in a clinical note, and autocompletes from the first four characters (Figure *1*). On selection, it expands into a structured template listing all seven components of the Oakland score with their published weightings: age, sex, previous admission with LGIB, digital rectal examination findings, heart rate, systolic blood pressure and haemoglobin concentration.



FIGURE 1: Schematic of the SmartPhrase workflow.

The clinician types the phrase name at any point in the clerking note, where it autocompletes after four characters and inserts a structured template listing all seven components of the Oakland score with their published weightings. Four components are populated automatically from the patient record, and three are entered by the clinician, who then sums them and records the total. This figure is an original schematic prepared by the author to illustrate the workflow and the provenance of each component; it is not a reproduction of any vendor software interface. Component weightings are reproduced from Oakland et al. [3].

Four of the seven components are populated automatically from the patient record using EPIC SmartLinks. Age is drawn from the patient header, heart rate and systolic blood pressure from the most recent observation flowsheet entries, and the two most recent haemoglobin results from the laboratory record. The remaining three components, all of which require clinical assessment rather than recorded data, are completed by the clerking clinician. The clinician then sums the component scores and records the total in the field provided.

The intervention therefore removes three distinct barriers to using the score: remembering that a validated score exists, recalling or looking up its component weightings, and retrieving the relevant observations and laboratory values from elsewhere in the record. It does not compute the total, which remains a manual arithmetic step. Automatic provision of decision support within the existing clinical workflow, rather than as a separate task, is among the features most consistently associated with effective computerised decision support [9-11].

The component weightings encoded in the SmartPhrase were checked against the original derivation and validation study and were found to correspond exactly, with haemoglobin thresholds expressed in g/L rather than g/dL in keeping with United Kingdom laboratory reporting conventions [3].

The structure of the expanded template, and the source of each value within it, is shown in Table *1*.

Component	Category	Points	Source of value in the SmartPhrase
Age, years	<40	0	Automatic (patient header)
	40-69	+1	
	≥70	+2	
Heart rate, beats per minute	<70	0	Automatic (observation flowsheet)
	70-89	+1	
	90-109	+2	
	≥110	+3	
Systolic blood pressure, mmHg	50-89	+5	Automatic (observation flowsheet)
	90-119	+4	
	120-129	+3	
	130-159	+2	
	≥160	0	
Haemoglobin, g/L	36-69	+22	Automatic (laboratory record)
	70-89	+17	
	90-109	+13	
	110-129	+8	
	130-159	+4	
	≥160	0	
Sex	Female	0	Entered by clinician
	Male	+1	
Previous admission with lower gastrointestinal bleeding	No	0	Entered by clinician
	Yes	+1	
Digital rectal examination findings	No blood	0	Entered by clinician
	Blood	+1	

TABLE 1: Components of the Oakland score and their implementation within the SmartPhrase.

Components of the Oakland score and their implementation within the SmartPhrase, ordered with the four automatically populated components first. Four components (age, heart rate, systolic blood pressure and haemoglobin) are populated automatically by EPIC SmartLinks; three (sex, previous admission with lower gastrointestinal bleeding, and digital rectal examination findings) are entered by the clerking clinician because they require clinical assessment rather than retrieval of a recorded value. Component categories and point weightings are reproduced from the original derivation and validation study by Oakland et al. [3], with haemoglobin expressed in g/L rather than g/dL in keeping with United Kingdom laboratory reporting conventions. Total score ranges from 0 to 35; a score of 8 or below identifies patients suitable for discharge [3,4].

Outcome definitions

All outcomes were defined identically in both cycles. Oakland score documentation was assessed across all episodes. The appropriateness of the admission decision was assessed in every episode with a calculable Oakland score and was classified as appropriate where a high-risk episode (score >8) was admitted or a low-risk episode (score ⩽8) was discharged, using the validated discharge threshold [3,4].

Outpatient investigation was assessed only in episodes that did not undergo inpatient endoscopy or interventional radiology, and in which an outpatient plan was recorded. Cycle 1 data collection did not capture whether a patient had undergone recent endoscopy before presentation, and this field could not be recovered retrospectively; the outcome was therefore defined without reference to prior endoscopy in both cycles. This is deliberately conservative and disadvantages cycle 2, in which three episodes with documented recent endoscopy and no new outpatient booking were counted as non-compliant.

Transfusion compliance was assessed only among transfused episodes and was defined as transfusion at a haemoglobin concentration below 70 g/L in accordance with the restrictive threshold recommended by the BSG and supported by randomised evidence [1,13,14]. No transfused patient in either cycle had cardiovascular disease, so the higher threshold permitted in that group did not apply. Anticoagulant management was assessed among episodes in which the patient was taking anticoagulant or antiplatelet therapy at presentation, and CT angiography among episodes with a shock index greater than 1.

Statistical analysis

Continuous variables are reported as medians with interquartile ranges and categorical variables as counts with percentages. Proportions were compared between cycles using Fisher's exact test. A two-sided p-value below 0.05 was considered statistically significant.

Given the small sample, the principal effect measure is the absolute difference in proportions between cycles, with 95% confidence intervals calculated by the Newcombe hybrid score method, which performs well for small samples and proportions near 0 or 1 [15]. Odds ratios are reported alongside for completeness but are not the basis for interpretation; with denominators as small as one, several are extremely imprecise and should not be read as estimates of effect magnitude.

Fisher's exact test derives the p-value directly from the hypergeometric distribution and does not generate an intermediate test statistic, so no value analogous to t, chi-square or F arises from it. Pearson chi-square statistics are reported for the two comparisons in which all expected cell counts were 5 or greater; for the remaining four, the chi-square approximation was invalid, and no statistic is given. Where a contingency table contained a zero cell, the odds ratio and its confidence interval were estimated using the Haldane-Anscombe correction of 0.5. Analyses were performed in Python 3 (Python Software Foundation, Wilmington, DE) using SciPy (SciPy developers, open-source).

No adjustment was made for multiple comparisons, and the analysis should be regarded as exploratory outside the two prespecified primary outcomes.

Ethical considerations

This was a retrospective analysis of routinely collected clinical records from 45 patients, comprising 50 bleeding episodes. The project was registered with and approved by the Clinical Audit Department of the Western Health and Social Care Trust as an audit of existing practice against national guidance (registration number AUD4590). Because the work used only data already recorded in the course of routine care and involved no change to patient management, no additional procedures and no contact with patients, formal research ethics committee review was not required under United Kingdom Health Research Authority criteria, and the requirement for individual patient consent was waived by the approving body. All data were extracted and analysed by members of the direct clinical care team, records were pseudonymised at the point of extraction, and no identifiable patient information is reported.

Results

Cohort characteristics

Fifty bleeding episodes were analysed, 28 in cycle 1 and 22 in cycle 2 (Table *2*). The observed baseline characteristics and median Oakland scores were broadly similar between cycles: median Oakland score, calculated for every episode with sufficient data irrespective of documentation, was 17 (IQR 10-23) in cycle 1 and 15 (IQR 10.5-21.2) in cycle 2. With 28 and 22 episodes, however, the absence of an obvious difference does not establish comparability, and residual differences in case mix cannot be excluded.

Characteristic	Cycle 1	Cycle 2
Bleeding episodes	28	22
Individual patients	25	20
Median age, years (range)	72 (20-93)	74 (29-87)
Male	16 (57.1%)	11 (50.0%)
Female	12 (42.9%)	11 (50.0%)
Median Oakland score (IQR)	17 (10-23)	15 (10.5-21.2)
Shock index >1	4 (14.3%)	1 (4.5%)
On anticoagulant or antiplatelet therapy	15 (53.6%)	13 (59.1%)
Received red cell transfusion	5 (17.9%)	8 (36.4%)

TABLE 2: Characteristics of the two audit cycles.

Primary outcomes

Documentation of the Oakland score by the clerking clinician increased from 6/28 episodes (21.4%) in cycle 1 to 19/22 (86.4%) in cycle 2, an absolute increase of 64.9% (95% CI 38.2 to 79.2; $\chi^2 = 20.78$, df = 1; $p < 0.001$). This was the largest change observed and the only one whose confidence interval excludes small effects.

Where a clinician had documented an Oakland score (six episodes in cycle 1 and 19 in cycle 2), the documented value agreed with the audit team's independent recalculation in all 25 cases. No arithmetic or transcription discrepancy was identified, although with 25 documented scores an underlying error rate as high as 11% would not have been excluded (exact one-sided 95% upper bound).

The appropriateness of the admission or discharge decision, assessed against the patient's Oakland risk category, increased from 21/27 episodes (77.8%) to 20/20 (100%), an absolute increase of 22.2% (95% CI 2.4 to 40.8; p = 0.031). The confidence interval excludes zero but is wide and extends close to it, so the magnitude of this change is poorly estimated. The chi-square approximation was invalid because two expected cell counts fell below five. One episode in cycle 1 and two in cycle 2 were excluded because insufficient data were recorded to calculate a score. The six discordant decisions in cycle 1 comprised three high-risk patients who were discharged (Oakland scores 9, 16 and 22) and three low-risk patients who were admitted unnecessarily (scores 7, 8 and 8). No discordant decision was observed in cycle 2.

Secondary outcomes

Outpatient investigation was arranged in 9/22 assessable episodes (40.9%) in cycle 1 and 13/18 (72.2%) in cycle 2, an absolute increase of 31.3% (95% CI 0.5 to 54.7; $\chi^2 = 3.92$, df = 1; p = 0.062). The confidence interval marginally excludes zero while Fisher's exact test does not reach the conventional significance threshold; the more conservative interpretation is adopted here, and the result is treated as non-significant.

Five episodes involved red cell transfusion in cycle 1 and eight in cycle 2. Transfusion occurred within the restrictive threshold in 3/5 episodes (60.0%) and 8/8 episodes (100%), an absolute difference of 40.0% (95% CI -3.0 to 76.9; p = 0.128). In absolute terms, two episodes of transfusion outside the recommended threshold occurred in cycle 1 and none in cycle 2; neither of the two patients concerned had cardiovascular disease that would have justified the higher threshold.

Among patients taking anticoagulant or antiplatelet therapy, treatment was appropriately interrupted or reversed in 12/15 episodes (80.0%) and 12/13 (92.3%), an absolute difference of 12.3% (95% CI -16.4 to 38.3; p = 0.60). The absolute number of correctly managed episodes was identical in both cycles, and the change in proportion reflects a smaller denominator rather than improved practice. CT angiography was performed in 2/4 episodes with a shock index greater than 1 in cycle 1 and in the single such episode in cycle 2 (absolute difference 50.0%, 95% CI -36.7 to 85.0; p = 1.00); no inference can be drawn from a denominator of one. Full results are shown in Table *3*.

Outcome	Cycle 1	Cycle 2	Absolute difference (95% CI)	Odds ratio (95% CI)	X^2	p
Oakland score documented	6/28 (21.4%)	19/22 (86.4%)	+64.9% (38.2 to 79.2)	23.2 (5.10-105.7)	20.78	<0.001
Admission decision concordant with risk category	21/27 (77.8%)	20/20 (100%)	+22.2% (2.4 to 40.8)	12.4 (0.66-234.3)†	N/A‡	0.031
Outpatient investigation arranged	9/22 (40.9%)	13/18 (72.2%)	+31.3% (0.5 to 54.7)	3.76 (0.99-14.29)	3.92	0.062
Transfusion within restrictive threshold	3/5 (60.0%)	8/8 (100%)	+40.0% (-3.0 to 76.9)	12.1 (0.46-323.3)†	N/A‡	0.128
Anticoagulant interrupted or reversed	12/15 (80.0%)	12/13 (92.3%)	+12.3% (-16.4 to 38.3)	3.00 (0.27-33.09)	N/A‡	0.60
CT angiography when shock index >1	2/4 (50.0%)	1/1 (100%)	+50.0% (-36.7 to 85.0)	3.00 (0.08-115.4)†	N/A‡	1.00

TABLE 3: Guideline adherence before and after intervention.

Absolute differences are cycle 2 minus cycle 1, with 95% confidence intervals by the Newcombe hybrid score method [15]; these are the primary effect measures. Odds ratios are shown for completeness only and are imprecise where denominators are very small. All p-values are from Fisher's exact test, which yields no intermediate test statistic. Denominators: all episodes (Oakland documentation); episodes with a calculable Oakland score (admission decision); episodes not undergoing inpatient endoscopy or interventional radiology and with a recorded plan (outpatient investigation); transfused episodes (transfusion); episodes on anticoagulant or antiplatelet therapy (anticoagulant management); episodes with shock index above 1 (CT angiography). The first two rows, Oakland score documentation and concordance of the admission or discharge decision with risk category, are the two prespecified primary outcomes; the remaining four are secondary. †Haldane-Anscombe correction of 0.5 applied because the contingency table contained a zero cell. ‡Chi-square approximation invalid because one or more expected cell counts were below 5.

Discussion

Embedding the Oakland score into the electronic clerking workflow was followed by a fourfold increase in documented risk stratification and by an increase in the proportion of admission and discharge decisions concordant with the patient's calculated risk. The second finding is the more clinically meaningful of the two. Documentation is a process measure and is valuable chiefly as a means to an end; the end is that patients who need admission are admitted and patients who do not are safely investigated as outpatients. In cycle 1, six of 27 assessable episodes failed that test, including three high-risk patients discharged with Oakland scores of 9, 16 and 22. In cycle 2, there were none, although with only 20 assessable episodes the upper bound of the plausible failure rate in cycle 2 remains appreciable.

The pattern of results is suggestive, though not demonstrative, about which components of the intervention did the work. The domains that changed most are those in which a decision must be made and recorded at clerking, precisely where the SmartPhrase operates. The domains that did not change are either those already close to ceiling performance in cycle 1, such as anticoagulant management, or those involving subgroups too small to permit inference, such as CT angiography in the single haemodynamically unstable episode of cycle 2. This pattern is consistent with the wider literature on computerised decision support, in which automatic provision of guidance within clinician workflow is among the strongest predictors of effect, while education and reminders delivered away from the point of care perform less well [8-11]. It does not, however, establish that the SmartPhrase rather than the concurrent educational components produced the change.

The change in outpatient investigation is worth noting despite not reaching significance. The definition used here was deliberately conservative: because prior-endoscopy data were unavailable for cycle 1, three cycle 2 episodes in which a patient had undergone recent endoscopy and no further outpatient investigation was booked were counted as failures, although that decision was clinically appropriate. The true improvement may therefore be understated.

Limitations

This audit has several important limitations. It was conducted at a single centre and analysed 50 episodes in total, so the study was underpowered for all outcomes other than the two primary measures; the absence of a statistically significant change in four domains should be read as an absence of evidence rather than evidence of absence. Even for the primary outcomes, the confidence intervals are wide, and the point estimates should not be taken as precise.

The intervention was multifaceted, combining education, visual prompts, email reminders and the SmartPhrase, and all components were introduced simultaneously. The observed changes therefore cannot be attributed to the SmartPhrase alone, and the design does not permit the contribution of individual components to be estimated.

The design was a small uncontrolled before-and-after comparison with no concurrent control group, so secular trends, changes in departmental staffing between cycles, and the Hawthorne effect arising from awareness of an ongoing audit are all plausible alternative explanations for the findings. Causal attribution to the intervention is therefore not possible.

The design was retrospective and depended entirely on the accuracy and completeness of contemporaneous electronic documentation. This is a particular concern for a study whose principal outcome is itself a documentation measure, since an intervention that prompts clinicians to record a score may improve recorded practice more than actual practice.

Risk categorisation used a score calculated uniformly by the audit team in both cycles rather than the score visible to the clinician at the time. This avoids the circularity that would arise from judging decisions against a number the clinician had just documented, and means the second primary outcome is independent of the first. It does, however, mean that the comparator is a score the treating team did not necessarily see, so the outcome measures whether disposition decisions aligned with an objectively calculated risk category rather than whether clinicians correctly applied a score in front of them.

Assessment of appropriateness was performed by the audit team without independent duplicate adjudication or formal assessment of inter-rater reliability, introducing a risk of observer bias. Three episodes in total (one in cycle 1 and two in cycle 2) lacked sufficient data to calculate an Oakland score and were excluded from the primary disposition analysis.

Two limitations arise from the design of the intervention itself. The SmartPhrase displays and populates the components of the score but does not total them, so a documented score depends on manual arithmetic. In this audit, all 25 documented scores matched independent recalculation, but that sample is small enough to be compatible with an error rate of up to 11%, and manual summation remains an avoidable failure point. In addition, the automatically populated observations and haemoglobin values are the most recent entries in the record rather than those at presentation. Where a note is written some hours into the admission, after fluid resuscitation or transfusion, the populated values may differ from the presenting physiology and the resulting score may understate the patient's risk.

The interval between cycles was short. Cycle 2 began within days of the intervention, so the results describe the immediate post-implementation period and provide no evidence of sustained effect. A further cycle at six to 12 months would be required to demonstrate durability, and is planned.

Finally, no patient-level clinical outcomes were captured. We did not measure length of stay, readmission with recurrent bleeding after discharge, transfusion volume, or mortality, so the suggestion that improved risk stratification benefits patients rests on the validity of the Oakland score in external cohorts [3,4] rather than on outcomes observed here.

Conclusions

In this small closed-loop audit at a district general hospital, introducing a multifaceted intervention centred on an electronic health record SmartPhrase was associated with a substantial increase in documented Oakland score use, and no admission or discharge decision discordant with calculated risk was observed in the second cycle. Changes in transfusion practice and outpatient investigation were in the expected direction but did not reach statistical significance in a cohort of this size.

These findings should be interpreted with caution. The audit was small, uncontrolled, and conducted over a short interval, and the intervention combined several components introduced together, so the contribution of the SmartPhrase cannot be isolated and causal attribution is not possible. The results are best regarded as a promising local service improvement rather than as evidence that electronic decision support improves patient outcomes. The intervention nonetheless required no additional staffing or funding, was designed and implemented by trainees, and is reproducible in any institution using a configurable electronic health record. No arithmetic errors were detected in the documented scores, but computing the total and its risk category automatically rather than leaving the final summation to the clinician would remove a residual failure point at no additional cost. Confirmation that the effect is real and sustained will require a further audit cycle with longer follow-up and the inclusion of patient-level clinical outcomes.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the

work.

Concept and design: Muhammad Danial Bin Rosli

Acquisition, analysis, or interpretation of data: Muhammad Danial Bin Rosli

Drafting of the manuscript: Muhammad Danial Bin Rosli

Critical review of the manuscript for important intellectual content: Muhammad Danial Bin Rosli

Supervision: Muhammad Danial Bin Rosli

Disclosures

Human subjects: Informed consent for treatment and open access publication was obtained or waived by all participants in this study. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The author thanks Mr Abdalazeez Ahmad and Ms Kathryn Boyce for their contribution to data collection for this audit.

References

1. Oakland K, Chadwick G, East JE, et al.: Diagnosis and management of acute lower gastrointestinal bleeding: guidelines from the British Society of Gastroenterology. Gut. 2019, 68:776-89. 10.1136/gutjnl-2018-317807
2. Oakland K, Guy R, Uberoi R, Hogg R, Mortensen N, Murphy MF, Jairath V: Acute lower GI bleeding in the UK: patient characteristics, interventions and outcomes in the first nationwide audit. Gut. 2018, 67:654-62. 10.1136/gutjnl-2016-313428
3. Oakland K, Jairath V, Uberoi R, et al.: Derivation and validation of a novel risk score for safe discharge after acute lower gastrointestinal bleeding: a modelling study. Lancet Gastroenterol Hepatology. 2017, 2:635-43. 10.1016/S2468-1253(17)30150-4
4. Oakland K, Kothiwale S, Forehand T, et al.: External validation of the Oakland score to assess safe hospital discharge among adult patients with acute lower gastrointestinal bleeding in the US. JAMA Netw Open. 2020, 3:e209630. 10.1001/jamanetworkopen.2020.9630
5. Sengupta N, Feuerstein JD, Jairath V, Shergill AK, Strate LL, Wong RJ, Wan D: Management of patients with acute lower gastrointestinal bleeding: an updated ACG guideline. Am J Gastroenterol. 2023, 118:208-31. 10.14309/ajg.0000000000002130
6. Triantafyllou K, Gkolfakis P, Gralnek IM, et al.: Diagnosis and management of acute lower gastrointestinal bleeding: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy. 2021, 53:850-68. 10.1055/a-1496-8969
7. Strate LL, Gralnek IM: ACG clinical guideline: management of patients with acute lower gastrointestinal bleeding. Am J Gastroenterol. 2016, 111:459-74. 10.1038/ajg.2016.41
8. Ivers N, Jamtvedt G, Flottorp S, et al.: Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012, 2012:CD000259. 10.1002/14651858.CD000259.pub3
9. Kawamoto K, Houlihan CA, Balas EA, Lobach DF: Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005, 330:765. 10.1136/bmj.38398.500764.8F
10. Roshanov PS, Fernandes N, Wilczynski JM, et al.: Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ. 2013, 346:f657. 10.1136/bmj.f657
11. Bates DW, Kuperman GJ, Wang S, et al.: Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc. 2003, 10:523-30. 10.1197/jamia.M1370
12. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D: SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. J Contin Educ Nurs. 2015, 46:501-7. 10.3928/00220124-20151020-02
13. Villanueva C, Colomo A, Bosch A, et al.: Transfusion strategies for acute upper gastrointestinal bleeding . N Engl J Med. 2013, 368:11-21. 10.1056/NEJMoa1211801
14. Hébert PC, Wells G, Blajchman MA, et al.: A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. Transfusion Requirements in Critical Care Investigators, Canadian Critical Care Trials Group. N Engl J Med. 1999, 340:409-17. 10.1056/NEJM199902113400601
15. Newcombe RG: Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med. 1998, 17:873-90. 10.1002/(sici)1097-0258(19980430)17:8<873::aid-sim779>3.0.co;2-i