

79679_Auto_Edited.docx

Name of Journal: *World Journal of Gastrointestinal Surgery*

Manuscript NO: 79679

Manuscript Type: ORIGINAL ARTICLE

Observational Study

**Development of a warning score for early detection of colorectal anastomotic leakage:
Hype or hope?**

Nuno J G Rama, Óscar Lourenço, Patrícia C Motta Lima, Maria Pedro S Guarino, Diana Parente, Ricardo Castro, Ana Bento, Anabela Rocha, Fernando Castro-Pocas, João Pimentel

Abstract

BACKGROUND

¹Colorectal anastomotic leakage (CAL), a severe postoperative complication, is associated with high morbidity, hospital readmission, and overall healthcare costs. Early detection of CAL remains a challenge in clinical practice. However, some decision models have been developed to increase the diagnostic accuracy of this event.

AIM

To develop a score based on easily accessible variables to detect CAL early.

METHODS

Based on the least absolute shrinkage and selection operator method, a predictive classification system was developed [Early ColoRectAL Leakage (E-CRALL) score] from a prospective observational, single center cohort, carried out in a colorectal division from a non-academic hospital. The score performance and CAL threshold from postoperative day (POD)3 to POD5 were estimated. Based on a precise analytical decision model, the standard clinical practice was compared with the E-CRALL

adoption on POD3, POD4, or POD5. A cost-minimization analysis was conducted, on the assumption that all alternatives delivered similar health-related effects.

RESULTS

In this study, 396 patients who underwent colorectal resection surgery with anastomosis, and 6.3% ($n = 25$) developed CAL. Most of the patients who developed CAL ($n = 23$; 92%) were diagnosed during the first hospital admission, with a median time of diagnosis of 9.0 ± 6.8 d. From POD3 to POD5, the ³ area under the receiver operating characteristic curve of the E-CRALL score was 0.82, 0.84, and 0.95, respectively. On POD5, if a threshold of 8.29 was chosen, 87.4% of anastomotic failures were identified with E-CRALL adoption. Additionally, score usage could anticipate CAL diagnosis in an average of 5.2 d and 4.1 d, if used on POD3 and POD5, respectively. Regardless of score adoption, episode comprehensive costs were markedly greater (up to four times) in patients who developed CAL in comparison with patients who did not develop CAL. Nonetheless, the use of the E-CRALL warning score was associated with cost savings of €421442.20, with most (92.9%) of the savings from patients who did not develop CAL.

CONCLUSION

The E-CRALL score is an accessible tool to predict CAL at an early timepoint. Additionally, E-CRALL can reduce overall healthcare costs, mainly in the reduction of hospital costs, independent of whether a patient developed CAL.

INTRODUCTION

Anastomotic leakage, a severe postoperative complication, remains the Achilles' heel of colorectal surgery, despite the technical advances in this field. Colorectal anastomotic leakage (CAL) ⁶ is associated with high morbidity, mortality, increased length of hospital stay (LOHS), reoperation rate, and healthcare costs^[1-5]. It is worth mentioning that CAL

has a major impact on the patient's quality of life and oncological outcomes, including cancer recurrence^[6-8].

² Nonspecific signs and symptoms often precede the acute and rapid clinical deterioration of a patient with CAL. Late diagnosis and management increase the likelihood of an undesirable outcome. Therefore, timely CAL diagnosis is crucial^[4, 9, 10]. Decision models have been designed to assess the risk of CAL development^[4, 10-13]. These models use regular scores of combined clinical, imaging, and laboratorial parameters, but the relevance of the models in early detection is still uncertain. The limited sensitivity (SS) of computed tomography (CT) in detecting CAL is a particular cause for concern and should be considered to avoid CAL diagnostic and management delays^[14]. Furthermore, it has been reported that an early minimally invasive reoperation should be considered in all patients with CAL suspicion because it is associated with low conversion, mortality, and morbidity rates^[15].

The occurrence of CAL has a significant negative influence on medical resource utilization. Thus, its early identification is critical to generate favorable economic outcomes while avoiding downstream economic impacts of CAL development^[1, 2, 16]. Use of diverting stomas, accurate scores, and attempted reoperation has been demonstrated to decrease LOHS, overall morbidity and readmissions^[16].

The purpose of this study was to develop a classification system capable of assisting clinicians in detecting CAL early and accurately. In addition, we aimed to assess the cost-effectiveness of using this classification system in daily clinical practice.

MATERIALS AND METHODS

PATIENTS AND METHODS

Prospective monocentric study design

A prospective, observational, single center study was conducted in a colorectal division of a non-academic hospital. The study included patients undergoing ¹urgent or elective colorectal resection, regardless of the approach (open or laparoscopic), indication (benign or malignant), and creation of a protective stoma. Data was collected between

March 1, 2017 and August 31, 2019 and recorded in a database according to the study protocol previously published^[17]. CAL, the main endpoint, was defined in accordance with clinical, imaging, and surgical criteria^[5, 17, 18]. Patients were excluded from the study if under 18-years-old, pregnant, unable to give written informed consent, had not received R0 resection with anastomosis, or had inflammatory bowel disease. A 90-d follow-up included data of postoperative complications (including CAL), LOHS and readmissions.

Development of the classification system

We aimed to establish clear and simple rules that can be used in daily clinical practice for recognizing patients at higher risk of CAL early. A predictive classification system was developed from patient-centered data and based on the least absolute shrinkage and selection operator method^[19]. The least absolute shrinkage and selection operator method is a classification technique for variable selection and regularization that results in balanced classifiers in terms of predictive ability and model interpretability^[19]. The classifier was named Early ColoRectAL Leakage (E-CRALL), and logo registration trademark was performed (Figure 1).

The first step to build the classifier included the estimation of conditional probability for developing CAL from the prospective study dataset and sorted into demographic, intraoperative and postoperative classes (Supplementary material Table S1). The postoperative category was grouped into three levels (clinical condition, abdominal pain, and biomarker plasma values) from postoperative day (POD)3 to POD5. The least absolute shrinkage and selection operator *Probit* and *Logit* models, suitable for binary dependent outcomes, were applied.

Further, the risk of overfitting was managed and reduced by splitting the sample. A training sample (70% of total) was used to estimate the models and build alternative classifiers for each POD (3, 4, and 5), and a testing sample (30% of total) was adopted to assess the performance of the classifier and the ability to predict CAL. The classifier with the best predictive performance was selected using cross-validation and

minimizing the deviance and deviance ratio statistics. The performance of alternative classifiers was also evaluated using the area under the receiver operating characteristic curve. Finally, the red flag threshold indicative of CAL was settled, maximizing both the SS and specificity (SP) of the classifier. Three different optimal classifiers were developed, one for each POD (3, 4, and 5).

Cost-minimization analysis

A cost-minimization analysis was conducted to compare the standard clinical practice (no use of E-CRALL) with the adoption of E-CRALL on POD3, POD4, or POD5, assuming that all alternatives delivered similar health-related effects^[20]. The time horizon of the decision problem was the 1st postoperative month, the target population was the prospective study patients, and the analysis perspective was that of the National Health Service. This cost-minimization analysis was based on the analytical decision model (Figure 2) presenting six possible patient pathways after application of E-CRALL^[20, 21]. The patient can be CAL positive or negative (observed ex-post but based on known ex-ante probabilities). In both branches, patients were divided by the optimal classifier, as E-CRALL positive or negative. All E-CRALL positive patients received an abdominal and pelvic CT scan. If the CT scan detected CAL, patients underwent proper management (Figure 2, pathways 1 and 4). Otherwise, if CT scan was negative or doubtful of CAL, patients were re-operated and managed accordingly (Figure 2, pathways 2 and 5). Finally, E-CRALL negative patients maintained appropriate clinical surveillance until CAL diagnosis (Figure 2, pathway 3) or discharge (Figure 2, pathway 6).

The branch probabilities to feed the tree came from several sources. The probabilities of CAL were estimated from the prospective study dataset, and the SS and SP of the E-CRALL score on a specific POD were estimated from the models. The predictive effect of abdominal and pelvic CT scan was drawn from relevant studies^[14, 22-24].

The estimation of costs to populate the model (Figure 2) were obtained from the Portuguese National Health Service reimbursement, used as a surrogate indicator for full hospital costs. Costs were based on the Ministerial Order nº 254/2018 of September 7, 2018 (Addendum III). The final costs of each of the six possible pathways were estimated under some assumptions, as presented in Table 1. The expected costs of each alternative were computed by the roll-back method^[20].

The estimation of costs for standard clinical practice were obtained as follow: $iCAL \times \text{Cost_CAL} + (1-iCAL) \times \text{Cost_NoCAL}$, where $iCAL$ was the incidence of CAL in the prospective study dataset and Cost_CAL (Cost_NoCAL) was the cost of treatment of a CAL (No CAL) patient. Costs were based on the Ministerial Order nº 254/2018 of September 7, 2018 (Addendum III). For each patient, the Diagnosis Related Group 221 and 223, respective degree of severity and comprehensive costs, were identified.

All statistical analysis was conducted using Stata Statistical software (Release 16; StataCorp, College Station, TX, United States).

RESULTS

RESULTS

Patients and outcomes

During the study period, we included 396 patients who underwent colorectal resection. Among them, 25 (6.3%) developed CAL. Age, the Charlson Comorbidity Index, and the American Society of Anesthesiologists grade affected the onset of CAL (Table 2). A laparoscopic approach was used in 82% of patients. The surgical approach ($P < 0.001$), the volume of blood loss ($P < 0.001$), the occurrence of intraoperative complications ($P < 0.001$), and the duration of the procedure ($P = 0.011$) were significantly related to the development of CAL (Table 2).

In this study, 92% of patients who developed CAL ($n = 23$) were diagnosed during the first hospital admission. The mean (\pm standard deviation) and median time for CAL diagnosis were 9.0 ± 6.8 d and 8 d (interquartile range = 7), respectively. Anastomotic leakage was significantly associated with a longer hospital stay [median of

21 d (patients who developed CAL) vs 7 d (patients without complications) vs 13 d (patients with other complications); $P < 0.001$]. The 90-d mortality rate was 0.8%, representing 3 patients who developed CAL (Table 2).

E-CRALL Score

Table 3 displays the variables and their respective weight on the score to determine the E-CRALL score for POD3-POD5. Many of the variables were statistically significant with predictive power to detect CAL. The predictive ability of this warning score had an AUROC for POD3 to POD5 of 0.82, 0.84 and 0.95, respectively (Figure 3 and Table 4). The score applied on POD5 had the best predictive power [0.95 (95% confidence interval: 0.90-0.99)].

The cutoff value for applying the E-CRALL score was calculated, defining the threshold for signaling a “patient who developed CAL.”. Setting the optimal cutoff as the one that maximizes both SS and SP of the classifier was established for POD3 and POD5 at 0.0551 and 0.0829, respectively. Considering a discriminant threshold of 5.51 (0.0551×100), the E-CRALL score on POD3 had a SS, SP, positive predictive value, and negative predictive value of 85.7%, 66.1%, 13.8%, and 98.7%, respectively. On POD5, if a threshold of 8.29 (0.0829×100) was chosen, then 87.4% of anastomotic failures were identified (Table 4).

Time to CAL diagnosis

The E-CRALL score adoption from POD3 to POD5 allowed the estimation of different lengths of time to detect CAL and the respective benefits in terms of time saving (Table 5). The E-CRALL score usage could anticipate CAL diagnosis in an average of 5.2 d if used on POD3 and in 4.1 d if used on POD5. CAL diagnosis was possible on the same day of E-CRALL score application on POD4 and POD5.

Cost analysis

Prospective monocentric study: In standard clinical practice, the patients who developed CAL had index admission comprehensive costs markedly greater (286%) than patients who did not develop CAL (€9096.00 *vs* €3177.00, respectively) (Table 6).

E-CRALL score application: In the model setting (Figure 2) after applying the E-CRALL score (on POD5), the adjusted comprehensive costs for each endpoint (pathway 1 to 6) were estimated and summarized in Table 6. In patients who developed CAL, episode comprehensive costs were markedly greater (four times) in comparison with patients who did not develop CAL (€8176.88 *vs* €1946.84, respectively).

Cost-minimization analysis

Regardless of CAL status, a cost comparison of the two approaches (standard clinical practice *vs* E-CRALL score application) from POD3 to POD5 was performed (Table 7). Greater cost savings were observed when the E-CRALL score was applied on POD5. Overall, the use of the E-CRALL warning score was associated with a cost savings of €421442.20, with most (92.9%) of the savings from patients who did not develop CAL (Table 8).

DISCUSSION

One strategy to anticipate CAL diagnosis included pooling clinical and laboratory variables in a weighted scoring system to improve the diagnostic accuracy measures of these variable when used separately. Design complexity, the need for external validation, and the difficulties in implementation in daily clinical practice are some of the challenges of score systems. So far, four scores have been developed for early CAL diagnosis; these are the Dutch leakage (DULK) score^[11], its modified version (the modified DULK)^[4], the Diagnostic Leakage (DIACOLE) score^[10], and those based on artificial intelligence methods^[13]. Each score has aimed to identify patients early, with suggestive CAL findings based on a cutoff point (discriminant threshold) to establish a management plan that includes additional exams or reoperation^[4, 10].

The E-CRALL score, proposed and tested in our study, demonstrated a substantial reduction in time to CAL detection (from 3.9 to 5.0 d) and expected time savings (from 4.1 to 5.2 d), depending on the day of its application. The use of the DULK score showed several benefits, namely the decrease in the delay to CAL detection (median 1.5 d compared to 4.0 d) and a reduction in CAL mortality (from 39% to 24%) compared to standard surveillance^[11]. The modified version of the DULK aimed to simplify the original version of the score. It was accomplished through the reduction of the number of parameters necessary to compute the score, becoming user-friendly for clinicians in daily clinical practice^[4]. With an exception for respiratory rate, the other three parameters were included in the E-CRALL warning score. The predictive ability of both the DULK modified version and E-CRALL score was quite similar. However, both score systems were developed based on distinct methodological approaches. Both tools aimed to recognize CAL early and seem to be useful as warning scores for further investigation (for example, CT scan with rectal contrast or reoperation).

The E-CRALL score has the benefits of a high AUROC after POD3, good predictive performance, and the inclusion of variables from the preoperative and intraoperative stages. However, our observations should be confirmed in a different cohort before their full clinical application. After external validation, E-CRALL may be useful for standardizing postoperative monitoring and aiding less experienced clinicians in the early detection of CAL detection, similar to the modified DULK score^[4]. Martin *et al*^[12], concluded that the DULK score was the most reliable instrument for early diagnosis of CAL. They also suggested its integration into risk management health policies to improve the quality of care according to the failure to rescue concept^[12, 18].

Artificial intelligence methods [*i.e.* artificial neural networks (ANNs)] were used by Adams *et al*^[12] to create a tool capable of accurately identifying patients at risk of developing CAL. They developed an ANN-based score and then trained and validated the score on a retrospective cohort. The score included 19 input variables from the three phases of the surgical process, similar to the E-CRALL score. Internal validation produced an AUROC, SS, and SP of 0.89, 85.0%, and 82.1 %, respectively. External

validation was estimated in a small prospective consecutive cohort (12 patients), presenting an SP of 83.3%. These results suggest good generalizability and effective prevention of overfitting by the ANN model. The authors concluded that models based on ANNs can assist in early detection of clinical CAL based on daily clinical data but not measuring this reduction to CAL detection, as E-CRALL score does.

The DIACOLE score was built from the results of a systematic review of the literature. At the onset, the potential laboratorial and clinical postoperative signs and symptoms of CAL were identified and complemented by a binary meta-analysis of those variables previously identified. Based on meta-analysis data, the weight of each identified factor was estimated. The DIACOLE diagnostic index showed an AUROC of 0.91, which was comparable with the E-CRALL score on POD5 (AUROC of 0.95) and was considered a good warning score for CAL diagnosis^[10]. The diagnostic threshold of the DIACOLE score was established using the cutoff point that optimizes SS and SP. This estimation process was identical in both scores, even though the E-CRALL score delivered higher SS and SP (> 90%) than the DIACOLE score (82.9%)^[10]. The authors of the DIACOLE score defined two discriminant thresholds: a lower level (> 3.065) advising daily clinical and laboratorial (with complete blood count) re-evaluation; and a higher level (> 5.436), recommending imaging (CT scan or water-soluble contrast enema)^[10]. On the other hand, the E-CRALL score established just one threshold, dependent on the POD and recommending imaging (CT scan) or early reoperation (if equivocal or negative imaging). Because both score calculations seem to be burdensome due to assessment concerns, the authors developed a user-friendly free software to compute the score value^[10]. Table 9 summarizes the distinctive aspects of the four scores available for CAL diagnosis.

This study has validated that the overall cost increases markedly for patients who develop CAL, being significantly greater (286.3%) than for patients who did not develop CAL. This result is in line with other reports. Ashraf *et al*^[16] found an increase of 154% in the mean in-patient hospital cost for 20 patients with anastomotic leakage

after anterior resection (£6233 ± £965 *vs* £9605 ± £6908 for non-CAL and CAL patients, respectively). Similar results were observed by other studies^[2, 25, 26].

One of the aims of this study was to assess the economic value of the use of the E-CRALL score. When comparing expected costs of E-CRALL application with those of standard practice, the results clearly pointed to the economic advantage of E-CRALL. We assumed that the health outcomes with and without the E-CRALL score were similar. Overall costs decreased after E-CRALL use, revealing a reduction of 32.0% and 13.6% in non-CAL and CAL patients, respectively, compared with standard clinical practice. These overall savings were first and foremost explained by the reduction in LOHS, as evidenced by the high proportion of savings that were seen in the non-CAL group (92.9%). Decision support systems based on inaccurate data are a source of false positive and false negative results, with possible adverse impacts on health and financial outcomes. Both potential false positives (*i.e.* excessive investigations) and false negatives (*i.e.* missed diagnoses) were incorporated in this analysis. However, in this study, costs related to false positive and false negative results had a lower impact than the benefits of the reduction in the LOHS. Moreover, reducing the time to CAL diagnosis had a smaller positive economic effect, accounting for 7.1% of cost savings (£32074.00). So far, a cost minimization analysis has not been performed in any of the similar scores mentioned above, but these tools may provide useful real-world information for improving financial outcomes.

A strength of the E-CRALL score is the combination of preoperative, intraoperative, and postoperative variables, emphasizing the clinical method because it incorporates technology (three biomarkers: calprotectin, C-reactive protein, and eosinophil cell count) and information from clinical data and physical examination (preoperative and intraoperative aspects, abdominal pain, and clinical condition).

Another strength of the E-CRALL score is defined as a single warning threshold, depending on the POD, and then recommending imaging (CT scan) or early reoperation (if equivocal or negative imaging). This simplifies the CAL detection approach. Additionally, an early operation in cases of dubious or negative imaging,

helps reduce the time to CAL detection and consequently starts CAL treatment promptly. Other authors concluded that early reoperation, namely re-laparoscopy, for managing complications following colorectal surgery appears to be safe and effective in highly selected patients^[27-29]. The key approach for this selection can involve the adoption of the E-CRALL score. In addition, a policy of early reoperation in patients with suspected complications enables timely management with expedient resolution, saving time to CAL diagnosis and to discharge^[29].

This study has several limitations. First, it is noteworthy that the E-CRALL score was developed and tested on only one dataset. Therefore, these findings should be considered with caution and should be validated externally, which is planned for a future multicentric, prospective study. Another limitation is related to the E-CRALL complexity for daily clinical implementation. It includes 13 diverse variables, which may increase the workload for healthcare staff.

Furthermore, this study addressed the economic burden of CAL in routine practice if all alternatives deliver equivalent health outcomes. This assumption is based on a conservative estimation since health outcomes improve with the early diagnosis^[29, 30]. In addition, there was a large divergence in the cost estimation of CAL, depending on the method of its calculation. This prospective study adopted comprehensive costs as there is the usual practice of public (National Health Service) reimbursement paid to the hospital. These methods may inadvertently underestimate costs due to under-coding or in contrast raise the practice of 'gaming' to receive more revenue. The estimation of personalized cost (tailored approach) by the aggregate of the index costs would be a more appropriate method^[16, 31].

Finally, it is crucial to estimate costs related to a delayed diagnosis as well as costs related to a high rate of false positive cases, unjustified reoperations, or frequent readmissions. Consequences of false negative cases on LOHS are difficult to accurately assess. A conservative policy was applied with the adoption of a cutoff with a SS around 100% to minimize the impact of false negatives on LOHS and the consequences of inappropriate early discharge.

CONCLUSION

CONCLUSION

The E-CRALL score demonstrated a high predictive ability, with SS and a negative predictive value of 100% after POD4 and a significant SP (86.6%) on POD5. This study internally validated the E-CRALL score for the early diagnosis of CAL and will integrate the local risk management policy, improving the quality of colorectal surgical healthcare. The routine adoption of the E-CRALL score may help prioritize CAL detection, supporting the policy of early reoperation in patients with suspected anastomotic failure. Even though the reduced time to CAL diagnosis had a smaller positive economic effect, overall costs decreased after E-CRALL use, revealing a noteworthy reduction of in-hospital costs, independent of CAL status, which was primarily due to the reduction in the LOHS in patients who did not develop CAL.

ARTICLE HIGHLIGHTS

Research background

Colorectal anastomotic leakage (CAL) is a surgical complication with a huge impact on morbidity and mortality. Early diagnosis of CAL can reduce these complications as well as hospital readmission and overall healthcare costs.

Research motivation

Decision models have been developed to increase the diagnostic accuracy of CAL. A user-friendly score applied in routine clinical practice can have a positive impact on the timely diagnosis of CAL and minimize healthcare costs.

Research objectives

To develop a score capable of assisting clinicians in early and accurate detection of CAL. In addition, we aimed to assess the cost-effectiveness of using this classification system in daily clinical practice.

Research methods

From March 1, 2017 to August 31, 2019, 396 patients who underwent colorectal resection with anastomosis were enrolled in a prospective, observational, single center study. A score based on the least absolute shrinkage and selection operator method developed and named the Early ColoRectAL Leakage (E-CRALL) score. The score performance and CAL threshold from postoperative day (POD)3 to POD5 were estimated. A cost-minimization analysis was also conducted.

Research results

This study included 396 patients who underwent colorectal resection with anastomosis. Among them, 6.3% ($n = 25$) developed CAL. The median time to CAL diagnosis was 9.0 ± 6.8 d. From POD3 to POD5, the ³area under the receiver operating characteristic curve of the E-CRALL score was 0.82, 0.84, and 0.95, respectively. The score anticipated CAL diagnosis in an average of 5.2 d and 4.1 d if used on POD3 and POD5, respectively. Overall costs in patients who developed CAL were markedly higher in comparison with patients who did not develop CAL. The E-CRALL warning score was associated with a cost savings of €421442.20.

Research conclusions

The E-CRALL score demonstrated a high predictive ability, with sensitivity and a negative predictive value of 100% on POD4 and a significant specificity (86.6%) on POD5. The routine adoption of the E-CRALL score may help prioritize CAL detection. Overall costs decreased after E-CRALL use, revealing a noteworthy reduction of in-hospital costs, independent of CAL status, which was primarily from the reduction in the LOHS for patients who did not develop CAL.

Research perspectives

A prospective, multicentric study will be conducted to test the warning score and promote external validation of our research.

6%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|---|-----------------|
| 1 | www.wjgnet.com
Internet | 166 words — 4% |
| <hr/> | | |
| 2 | Nuno Rama, Diana Parente, Cândida Silva, Miguel Neves et al. "Anastomotic Leak in Colorectal Cancer Surgery: From Diagnosis to Management or Failure - A Retrospective Cohort Study.", Research Square, 2021
Crossref Posted Content | 31 words — 1% |
| <hr/> | | |
| 3 | www.ncbi.nlm.nih.gov
Internet | 24 words — 1% |
| <hr/> | | |
| 4 | G. Martin, A. Dupré, A. Mulliez, F. Prunel, K. Slim, D. Pezet. "Validation of a score for the early diagnosis of anastomotic leakage following elective colorectal surgery", Journal of Visceral Surgery, 2015
Crossref | 18 words — < 1% |
| <hr/> | | |
| 5 | www.labome.org
Internet | 18 words — < 1% |
| <hr/> | | |
| 6 | pure.uva.nl
Internet | 13 words — < 1% |

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES

EXCLUDE MATCHES

OFF

< 12 WORDS

