Journal of Cancer Policy 47 (2026) 100702

Contents lists available at ScienceDirect

Journal of Cancer Policy

journal homepage: www.elsevier.com/locate/jcpo

Population-Based Cancer Screening analysis in Northern Portugal Using Process Mining

Hugo Monteiro [a,*,1], Mariana Oliveira [a], Ricardo Martinho [b], João Reis [c], Fernando Tavares [c], Óscar Felgueiras [d], Carlos Martins [e]

[a] Faculdade de Medicina da Universidade do Porto, Faculty of Medicine, University of Porto, Al. Professor Hernâni Monteiro, Porto 4200-319, Portugal
[b] Centro de Investigação em Tecnologias e Serviços de Saúde, CINTESIS, ESTG, Polytechnic University of Leiria Portugal, Portugal
[c] Direção Executiva do SNS, Departamento de Estudos e Planeamento , Rua Da Constituição 195, Porto 4200-198, Portugal
[d] Faculdade de Ciências do Porto, Faculty of Sciences, University of Porto, Portugal
[e] Centro de Investigação em Tecnologias e Serviços de Saúde University of Porto, Centre for Health Technology and Services Research (CINTESIS@RISE), Portugal

ARTICLE INFO

Keywords:
Cancer screening
Process mining
Healthcare resilience
Colorectal cancer

ABSTRACT

Background: This study focuses on the Colorectal Cancer Screening Program in Northern Portugal, aiming to evaluate the disruption effects on its performance and efficiency.
Methods: We conducted an observational analyses of 271 637 administrative records from 2020 to 2022. Administrative timestamps were converted into a step-by-step dataset of screening activities (an "event log") and analysed using process mining and comparative performance analysis across time periods and ACeS (primary care administrative clusters).
Results: Consultation-to-colonoscopy time lengthened by 53 %, rising from a median 58 days (IQR 29–92) in early 2020 to 89 days (IQR 53–127) in 2021, before improving to 73 days in 2022. Conversely, referral-to-consultation time fell from 110 days to 26 days (−76 %), reflecting targeted backlog clearance. Screening volumes declined in 2020 but recovered above baseline levels by 2022. Performance differences across primary care administrative clusters were significant ($p < 0.001$), with some units outperforming regional median transition times. Early adoption of automated electronic referrals and flexible consultation scheduling may have contributed to improved programme performance during the recovery period following pandemic-related disruptions. Substantial heterogeneity across units was observed for key transitions, indicating uneven disruption and recovery patterns across administrative units.
Conclusion: Process Mining techniques revealed critical vulnerabilities in the screening program during the initial stages of the period in analysis (matching the pandemic). These findings support targeted monitoring and prioritisation of operational improvements to reduce avoidable delays and strengthen continuity of population-based screening.
Policy summary: Policies aimed at strengthening healthcare service continuity and operational capacity benefit from analytical methods like process mining. Key recommendations include standardizing workflows, enhancing coordination between primary care and hospital services, and investing in digital monitoring systems to mitigate disruptions and ensure continuity in cancer screening programs during periods of system stress.

1. Introduction

Healthcare services rely on complex workflows that are vulnerable to disruption due to resource shortages, skill gaps, demand surges, and public health emergencies [1]. For instance, in 2020, the COVID-19 pandemic disrupted healthcare services worldwide, impacting many activities, including preventive care such as screening programs [2–8]. Such disruptions can recur, and screening programs are particularly

* Corresponding author.
E-mail addresses: up200204272@edu.med.up.pt, hugo.monteiro@ulsge.min-saude.pt (H. Monteiro), mariana_ramos_oliveira@hotmail.com (M. Oliveira), ricardo.martinho@ipleiria.pt (R. Martinho), joao.reis@sns.min-saude.pt (J. Reis), fernando.tavares@sns.min-saude.pt (F. Tavares), olfelgue@fc.up.pt (Ó. Felgueiras), carlosmartins20@gmail.com (C. Martins).
[1] Present address: Departamento de Estudos e Planeamento, Direção Executiva do SNS, Rua Da Constituição 195, Porto 4200–198, Portugal.

https://doi.org/10.1016/j.jcpo.2026.100702
Received 28 March 2025; Received in revised form 22 December 2025; Accepted 3 January 2026
Available online 5 January 2026

susceptible to being deprioritized during periods of stress. Understanding vulnerabilities and variation in operational performance across settings can help policymakers and providers strengthen screening services during disruptions and support recovery planning.

One such program focuses on the screening of colorectal cancer (CRC), which remains a leading cause of cancer mortality in Portugal. Its incidence, mortality and stage distribution vary widely across Europe. Countries with long-standing colorectal cancer screening programs (CCSP) have seen different epidemiological trajectories. Other studies observed that countries without broad coverage continue to see rising incident rates; early postimplementation spikes in age-standardized decrease [7]. These findings support the relevance of effective screening, both in managing individual health and influencing national cancer trends, to improve health outcomes [9,10].

Process mining has gained prominence as a method for reconstructing real-world clinical workflows from digital traces, allowing for detailed analysis of performance, bottlenecks, and deviations. Rinner et al. [1] applied process mining techniques to melanoma surveillance, revealing hidden delays and variations. More recently, Aversano et al. [11] systematically reviewed process mining applications in healthcare, including colorectal cancer screening, highlighting its role in characterizing pathway changes, bottlenecks, and performance variation during disruptive periods.

The Colorectal Cancer Screening Program (CCSP), launched in 2018 in the North Region of Portugal, targets individuals aged 50–74 through an organized and staged workflow across multiple care points. Patients are enrolled through ACeS (*Agrupamentos de Centros de Saúde*), administrative clusters of primary care centres (PCC) responsible for defined catchment populations and invited in planned cycles to manage resource use. This process entails that those patients do not overwhelm different resources and helps maintain program manageability. As illustrated in Fig. 1, the process begins with an invitation letter, followed by a faecal immunochemical test (FIT) kit – a type of faecal occult blood test (FOBT). Returned samples are analysed at a central laboratory, and patients are then informed of results. FIT-positive individuals are referred to their family doctor for a pre-colonoscopy consultation and, when appropriate, scheduled for colonoscopy at the designated hospital. The colonoscopy findings, including polyp histology, guide subsequent care [12]. Cases may be excluded from a screening cycle due to administrative or clinical reasons. Those not returning the FOBT or excluded but still eligible are invited again in future cycles. This creates a complex operational system with interdependent steps and feedback loops. The CCSP's design makes it a strong candidate for Process Mining analysis, enabling the reconstruction of real workflows, detection of inefficiencies, and performance benchmarking. Operational monitoring focuses on three stages: 1) FIT kit distribution and return, 2) result handling and consultations, and 3) colonoscopy coordination. These insights support targeted improvements and are applicable to similar preventive programs globally.

The CCSP is an example where healthcare services need to align with a range of operations, making it susceptible to disruptions [13]. Such interruptions can affect the program's overall performance, and the efficiency of the resources involved [14]. Even though most performance indicators cover clinical output measures [15], there are other strategies like process mining capable of assessing performance and providing insights of system behaviour [16].

Although the CCSP began in 2018, recording 318,963 screening episodes through December 2022, for the main analysis, we focused on 271,637 episodes that occurred between January 1st, 2020, and December 31st, 2022. The North Region accounted for more than 75 % of all the screening program execution at the national level during the study period. This study addresses the following research questions:

1. How did the COVID-19 pandemic affect key transition times within the colorectal cancer screening workflow (e.g., invitation to FIT return, consultation to colonoscopy) across the study periods?
2. Which ACeS (administrative units) showed comparatively better performance (shorter delays and/or smaller deterioration) during the pandemic period?
3. Which workflow patterns characterised higher-performing ACeS, and how might these inform targeted future improvements in national screening programs?

2. Methods

To address these objectives, we applied a multi-step analytical approach combining process mining with statistical performance evaluation, as detailed below.

2.1. Data source and event log creation

We first developed an event log file in tabular format for use with process mining algorithms [17]. In this study, an event log refers to a structured table where each row represents a time-stamped step recorded for a screening episode (case). This format enables reconstruction of the observed workflow from routine administrative records. A screening episode corresponds to a single program cycle for an eligible individual; the same individual may re-enter the program in later cycles. Based on CCSP administrative records, this dataset included timestamps for key activities, such as invitation mailings, FIT kit distribution and return, laboratory results, and colonoscopy scheduling. The generated event log incorporated metadata to track the specific healthcare resource associated with each activity, enabling analysis of regional differences in performance.

In the Portuguese NHS, *ACeS* are geographically defined primary care clusters that group multiple Primary Care Centres and serve a resident population. In our dataset, ACeS were used as the administrative unit ('resource') to compare operational performance across territories. The constructed event log included the following variables:

- case_id: unique identifier for each screening episode,
- activity: name of the recorded step in the workflow (e.g., Invitation_mail, FIT_mail, FIT_return),
- timestamp: the date when the activity occurred,
- resource: *ACeS* identifier (primary care administrative cluster responsible for the episode).

Laboratory outcomes were encoded explicitly (FIT-positive vs FIT-negative); FIT-negative episodes terminate after result notification (returning for invitation after 2 years), while FIT-positive episodes may continue to primary care referral and onward coordination for colonoscopy.

2.2. Process mining and workflow analysis

We applied process mining using Disco (Fluxicon)[2] software to reconstruct the observed colorectal screening workflow from the event log. We first generated a directly-follows representation of the process and inspected dominant variants (most frequent activity sequences). To improve interpretability and avoid over-emphasizing noise from rare or incomplete traces, we filtered infrequent variants and administrative anomalies (e.g., implausible event orderings), while ensuring that the retained set still represented the full set of recorded activity types [18]. The resulting models were then compared across predefined time periods and across ACeS units.

2.3. Statistical analysis and comparative performance

We extended the comparison to analyse performance across different

[2] https://fluxicon.com/disco

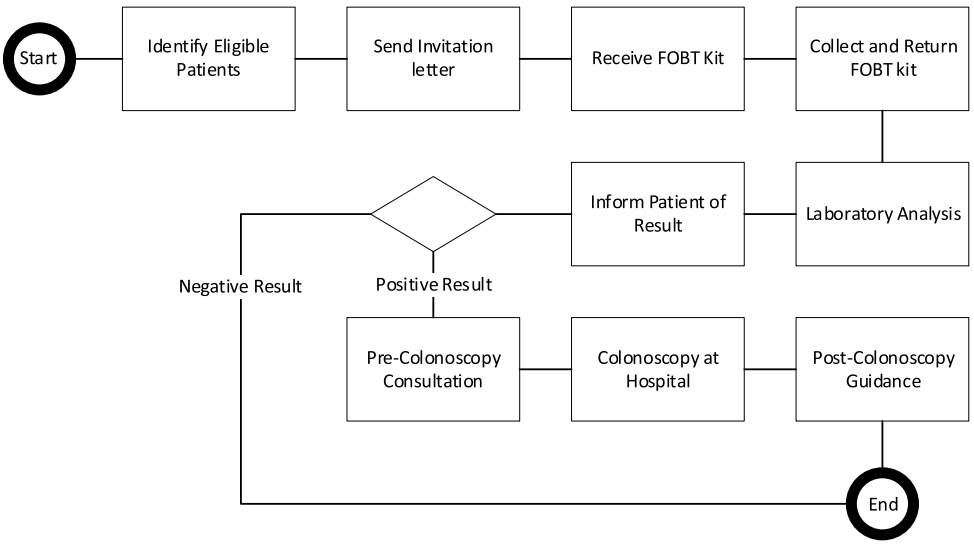

Fig. 1. Colorectal cancer screening process.

resources, particularly Primary Care Centre administrative units (ACeS), which groups cases geographically, aiming to pinpoint those that might be influencing overall performance changes [14]. The resources were present in the dataset and numbered sequentially.

After identifying deviations and filtering traces to focus on areas of interest, we performed comparative performance analysis. Timestamps were computed between activities of the main traces: T1 = Invitation_mail to FIT_mail, T2A = FIT_mail to FIT_return, T3 = FIT_return to Lab_receive, T4 = Lab_receive to Lab_result, T5P = Lab_result (positive) to PCC_fwd, T6 = PCC_fwd to PCC_observation, and T7 = PCC_observation to Colonoscopy center. All transition times were calculated in days as differences between consecutive timestamps. Statistical tests included Shapiro-Wilk for normality, ANOVA for comparing periods: 1st semester of 2020 (significantly impacted by the onset of the COVID-19 pandemic in Portugal), the 2nd semester of 2020 (where adjustments to pandemic conditions continued), and the subsequent years of 2021 and 2022. For pairwise comparisons we applied Tukey's post-hoc test. Based on initial results, we then conducted ANOVA by resource to determine units with significant performance differences. Given the large sample size, we used ANOVA as a robustness-oriented comparison of period means and confirmed consistency of findings using non-parametric Kruskal–Wallis tests in parallel. To reduce bias from unbalanced activity across periods, we restricted comparisons to ACeS units active in all evaluated periods.

2.4. Validation through stakeholder engagement

Key findings and anomalies identified through statistical analysis were reviewed with the CCSP regional coordination team in structured feedback sessions. The aim was to validate interpretation of administrative timestamps and clarify operational mechanisms underlying observed patterns (e.g., negative intervals attributable to event ordering/data-entry artefacts; referral steps registered automatically in the information system). This input was used exclusively to support methodological interpretation and data-quality decisions; no formal qualitative analysis was conducted, and no interview-derived results are reported.

2.5. Ethics and compliance

All data were fully anonymized prior to analysis. No individual patient consent was required, as all records were de-identified in compliance with GDPR and with the approval of the Health Ethics Committee of the Regional Health Administration of Northern Portugal (CE/2023/96).

3. Results

The results are organised to present the discovered process models, followed by an analysis of performance deviations across time periods and resources.

3.1. Process model discovery and workflow variants

The initial ad-hoc analysis revealed some unexpected findings, including long transition periods or atypical sequences (e.g. lab results preceding kit delivery). Many traces had only one case (high process variability) or other bottleneck traces that provided little value for performance analysis.

We performed data-quality filtering to remove temporal anomalies and sequences incompatible with the programme logic (e.g., Lab_result recorded before FIT_return). Episodes were retained even when they terminated early (e.g., after a negative FIT result), because such terminations represent most screening outcomes. These filters removed approximately 5 % of the original 285,390 episodes. After cleaning, the final dataset comprised 271,637 episodes (95 % of the total), preserving all recorded activity types while improving interpretability of the discovered workflow.

3.2. Key workflow variants and frequencies

Four variants were discovered and are depicted in Fig. 2. Importantly, a negative FIT result terminates the screening episode and therefore constitutes the expected most frequent pathway in population-based screening:

- Variant 1 (85 % of cases): Standard workflow - invitation, receive FIT kit, lab return and lab result, primary care consultation of positive cases, referral and execution of colonoscopy
- Variant 2 (~3 %): FIT kits are rejected at the primary care centre before reaching the lab, requiring re-invitation
- Variant 3 (~2 %): A negative FIT result leads to immediate exclusion without referral or follow-up – explicit post-laboratory exclusion variant.
- Variant 4 (~10 %): Incomplete traces due to dropouts or missing follow-up.

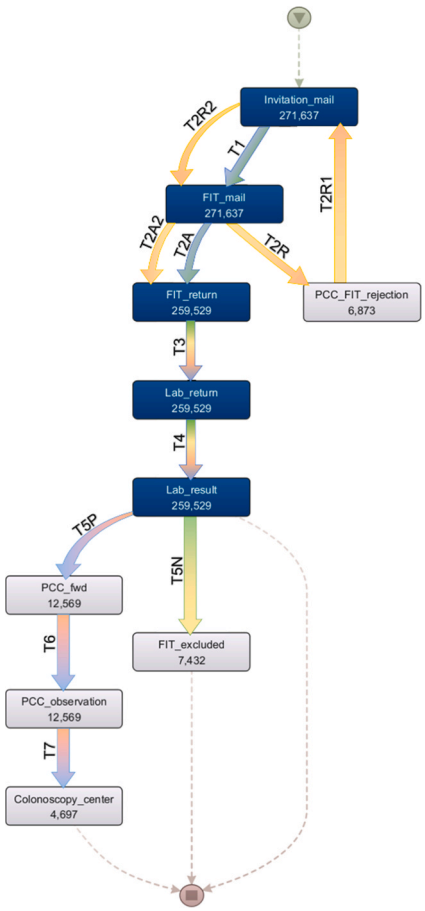

Fig. 2. Process model (Direct Flow Graph) of the four variants obtained after cleaning infrequent traces. T1: time between invitation mail and KIT receiving mail; T2R: time between receiving the KIT and having it rejected; T2R1: time between having the KIT rejected and receiving new invitation mail; T2R2: time between receiving new invitation mail and new KIT receiving mail; T2A: time between receiving the KIT and returning it; T2A2: time between receiving the second KIT and returning it; T3: time between returning the KIT and the lab receiving it; T4: time between the lab receiving the KIT and having the result; T5N: time between the result being negative and the patient being excluded from the study; T5P: time between the result being positive and the patient being forwarded to the family doctor; T6: time between the patient being forwarded to the family doctor and the actual appointment; T7: time between the actual appointment and having the colonoscopy.

These variants are depicted in Fig. 2, with arrows labelled according to timestamp intervals (T1–T7). Most cases follow Variant 1, while the remaining patterns reflect procedural exceptions or data terminations.

Fig. 3 presents a dotted chart illustrating process dispersion over time. Each dot represents a recorded activity, and horizontal alignments correspond to distinct screening episodes. The x-axis represents calendar time, and colours distinguish activity types to support visual separation of steps (e.g., invitation, kit return, laboratory, referral). The chart highlights a screening programme interruption during the first half of 2020 (first COVID-19 wave) and a second disruption around March 2022, where many episodes appear to halt and restart from early steps, consistent with a temporary pause or capacity constraint affecting upstream activities.

3.3. Temporal and regional performance analysis

3.3.1. Data preparation

For the performance analysis, transition times between two consecutive activities were calculated in days by subtracting the timestamps of each event from its subsequent event. We then compared performance metrics across different periods, reflecting the disruption and recovery phases. The data were then segmented into four distinct periods: the first semester of 2020 (8 % of the episodes/cases), the second semester of 2020 (25 %), 2021 (41 %) and 2022 (26 %). The Shapiro-Wilk showed non-normal distribution. Despite this, we applied ANOVA based on the large sample size ($n > 200{,}000$), under the Central Limit Theorem and confirmed results using Kruskal–Wallis tests in parallel. Consequently, median and interquartile ranges were computed using data from all ACeS available.

3.3.2. Period Performance Conformance analysis

T4 (Lab_receive to Lab_result), T5P (Lab_result positive to PCC_fwd), and T5N (Lab_result negative to exclusion) consistently registered a median of 0 days (IQR: 0–0), indicating automated and/or uniform system entries for these steps (Fig. 2). Only a small fraction of episodes had non-zero T4 values; after validation with programme coordination, these were treated as anomalies and excluded from detailed interpretation. Because these timestamps convey limited operational delay, subsequent comparisons focused on transitions reflecting organisational capacity and scheduling constraints (notably T6 and T7).

As for the timestamp T5P, a similar analysis showed that most patients (over 90 %) had no delay. This timeframe represents the time spent between the registry of the test result and the patient's referral registry to the Family Doctor consultation, and since this is an automatic process, as clarified with the program's coordination, it is expected to have zero duration when test results are positive. Therefore, this timestamp variable was removed from further analysis.

To assess the differences across various timestamps, an ANOVA test was performed, with a subsequent Tukey test applied to explore these differences over different time periods. To ensure comparability and minimize bias from unbalanced data, we filtered data to only consider primary care centres that were active both in the first semester of 2020 and in the remaining time periods.

The ANOVA results illustrate significant differences across most timestamps in our healthcare process analysis. T7, the period from the primary care centre consultation to colonoscopy, showed the greatest variability, suggesting potential inefficiencies or delays at this stage. Conversely, T1, which tracks the period from the invitation mail to kit mailing, exhibited the significant differences across periods, indicating relative stability in this early stage of the process.

The results from the Tukey post-hoc tests, compare differences between specific semesters. Most notably, the first semester of 2020 shows consistent and considerable differences when compared with subsequent periods.

3.3.3. Resource Performance Comparative Analysis

Observation of the ACeS resources used to execute each case of the process showed that some had considerably fewer traces, leading to the exclusion of ACeS 7, 10 and 11 from this analysis. The remaining resources for analysis represented 88 % of all episodes.

To further assess the inter-resource variation, we applied a Kruskal-Wallis's test, comparing each ACeS against all others, including those active in every period (Table 1). The T2R transition (FITMail to PCC-FITRejection) showed significant differences, suggesting variability in how different centres manage potential rejections of the FIT mail at the primary care level; the observation of T4 (LabReturn to LabResult) implies variability in the time it takes for the laboratory to process the biological samples and to provide its results to the management platform; the T7 (PCCFwd to Colonoscopy) is the final step up to the realization of the colonoscopy and it presented significant differences in most of the resources. We can infer from observations from ACeS such as 24 and 19 that there are units showing statistically significant differences across several timestamps, suggesting a broad range of areas that require a more detailed analysis.

To visualize inter-resource variation during disruption and recovery,

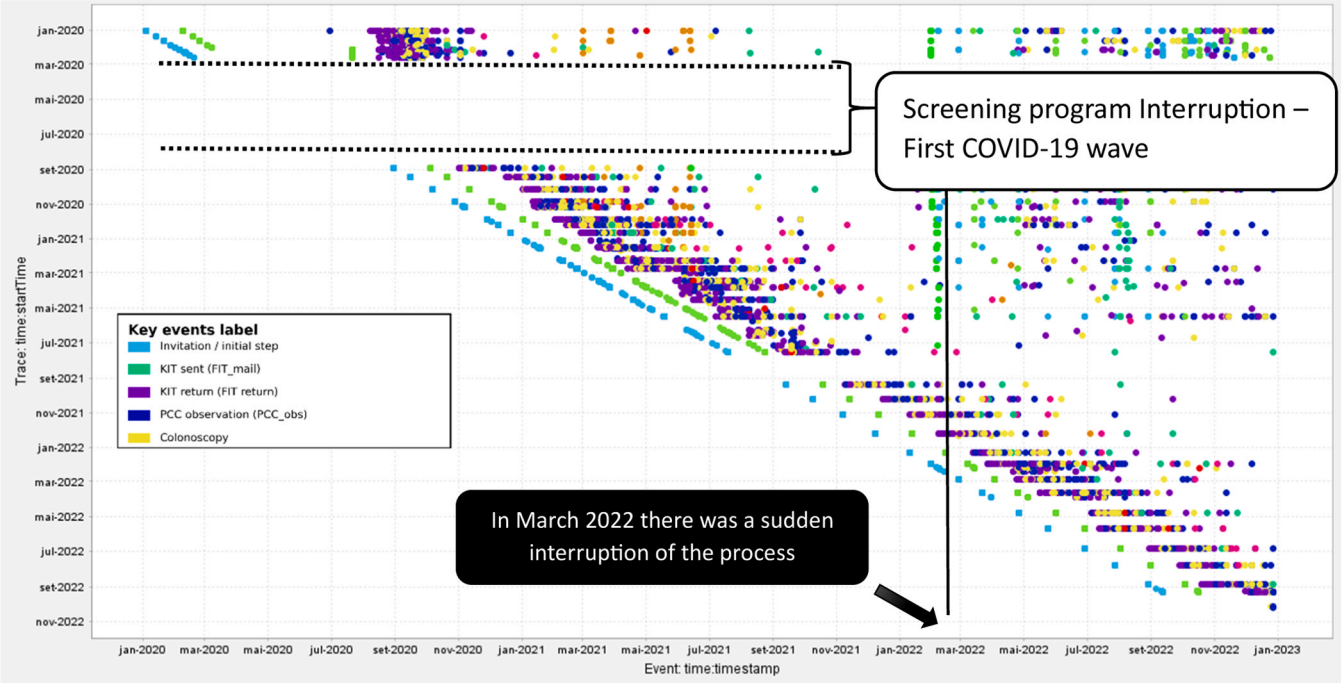

Fig. 3. Dotted chart illustrating the temporal dispersion of screening process activities across individual cases. Each dot represents a recorded event (activity) in the screening pathway, and horizontal alignment of dots corresponds to a single screening episode (case). Colours indicate distinct activities identified through process mining, reflecting different sequences of activities observed in the data. The figure highlights two major disruptions: (i) a complete interruption of screening activities during the first COVID-19 wave in early 2020, and (ii) a sudden disruption in March 2022, characterised by widespread case interruptions followed by re-initiation from the initial invitation step.

Table 1
Inter-ACeS differences in key screening timestamps (waiting time in days). Global variability across timestamps was assessed using ANOVA, and Kruskal–Wallis tests were used to identify inter-ACeS heterogeneity. Timestamps represent intervals between key activities in the screening process. (*- p-value ≤0,05, **- p-value≤ 0.01). The "# ACeS with heterogeneity" column indicate the number of ACeS (out of 12) showing significant differences when compared against all others (KW $p \leq 0.01$). Timestamp definitions are provided in Methods (Section 2.3) and Fig. 2 caption; full outputs are available in Supplementary Table S1.

Timestamp	Transition	# *ACeS* with heterogeneity (KW $p \leq 0.01$)	Interpretation
T2R	Kit received → Rejection	8/12	Variation in primary care admin handling of kit rejection
T3	Kit returned → Lab received	12/12	Territorial differences in transport / lab intake timing
T6	Forwarded → Consultation	9/12	Primary care consultation scheduling capacity differences
T7	Consultation → Colonoscopy	11/12	Hospital capacity / scheduling differences for colonoscopy

Fig. 4 shows boxplots of waiting times (days) comparing COVID vs post-COVID periods for T6 (referral to consultation) and T7 (consultation to colonoscopy). T6 reached 110 days in early 2020 and decreased to 26 days by 2022 (−76 %). ACeS10 showed the highest T6 values, while ACeS13 and ACeS19 remained consistently lower. In contrast, T7, increased median times from 58 days in 2020–89 days in 2021 (+53 %) and then decreased to 73 days in 2022. Several ACeS showed significant variation: some units presenting extended delays while others maintaining shorter transitions, highlighting higher resilience. This discrepancy between T6 and T7 suggests that referral-to-consultation workflows (T6) were more vulnerable to disruption, highlighting the need for focused improvement efforts in this earlier phase of the screening pathway.

4. Conclusion

The implementation of advanced data and process mining techniques facilitated the analysis of over one million screening records from the Colorectal Cancer Screening Program. Tools such as Disco software enabled the transformation of complex datasets into structured event logs, allowing efficient data processing and analysis within a "low code" environment. This efficient handling of big datasets underscores the significance of robust data analytics tools in contemporary healthcare research.

The COVID-19 pandemic caused measurable delays across several screening stages. Median T7 (consultation to colonoscopy) increased from 58 days [IQR: 29–92] in early 2020–89 days [IQR: 53–127] in 2021, with partial recovery 73 days [IQR: 40–120] in 2022. T6 (referral to consultation) decreased from 110 days [IQR: 23–400] in early 2020–26 days [IQR: 13–52] in 2022. Some ACeS showed prolonged disruptions, while others recovered earlier. Post-lockdown surges reflect backlog clearance. These delays, also reported in other European programs, expose a shared vulnerability in early-stage pathways and highlight the need for robust service continuity planning during crises [8,19].

Limitations include: First, administrative records may omit unlogged events, such as patient cancellations or system-level overrides, which can bias the discovery process of complete patient traces. Second, data quality issues (e.g., negative T4 values or duplicated transitions) required data cleaning and stakeholder validation, potentially excluding rare behaviours. Third, lack of sociodemographic or clinical data limits interpretation of pathway differences. Fourth, although the data reflect the Northern Region, which accounted for over 75 % of national screenings, generalization to other systems requires caution.

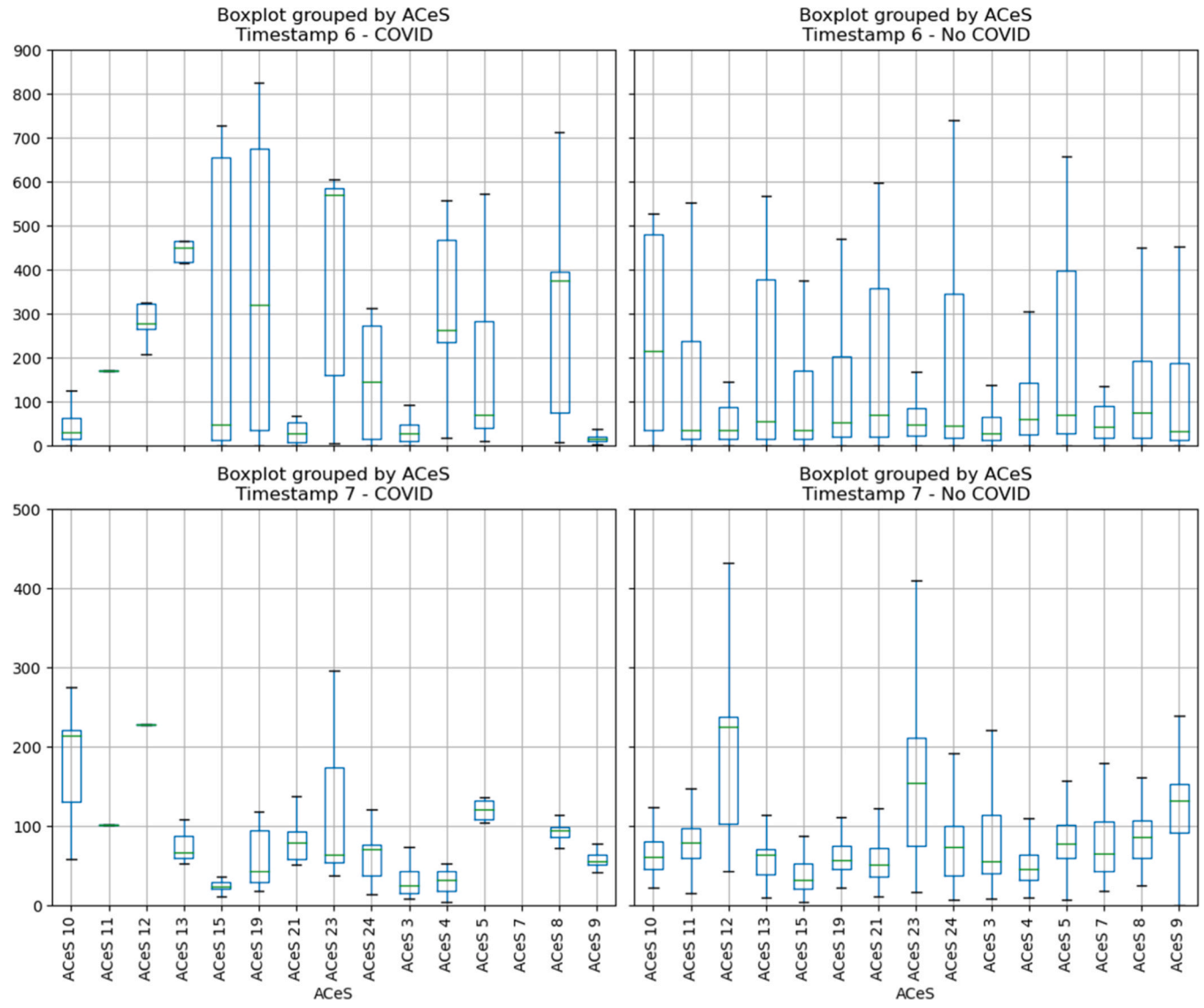

Fig. 4. Boxplots of waiting times (in days) during the COVID-19 period (left) and post-COVID period (right) for key screening transitions: T6 (top), referral to primary care consultation, and T7 (bottom), consultation to colonoscopy. Boxes represent medians and interquartile ranges; whiskers indicate variability across ACeS.

This study presents performance nuances within the CCSP during COVID-19 and demonstrates the merit of process mining in revealing vital workflow deviations. Although the pandemic served as the initial disruption context, these results also illuminate how preventive and non-acute healthcare services can be deprioritized under systemic stress, making it important to put in place proactive measures for protecting such programs in future crises.

Future research should address key opportunities identified in this study and broader evaluations of process mining in healthcare. Mixed-methods audits of high-performing ACeS could reveal contextual factors critical to maintaining continuity during disruptions. Integrating longitudinal patient-level data would enable outcome modelling, explicitly linking delays in transitions to clinical consequences such as diagnostic stage or survival. Extending per-case conformance checking could support near-real-time detection of workflow deviations, enabling proactive interventions. Finally, evaluating the cost-effectiveness of digital monitoring under stress scenarios would provide evidence to support the scaling of resilient healthcare systems. Together, these proposals constitute a practical research agenda aimed at operationalizing process mining insights and enhancing the resilience of cancer screening programs [11].

5. Policy summary

Following this study's findings, the following policy actions are proposed:

- Implement a unified digital protocol for FIT kit distribution and tracking, to reduce early-stage variability;
- Establish joint screening teams to meet biweekly, monitor case backlogs, and adjust colonoscopy schedules to reduce delays;
- Deploy real-time dashboards to monitor event transitions with thresholds (e.g., T7 $>$ 30 days) that trigger operational reviews;
- Expand performance frameworks to include equity, follow-up, and pathway consistency, with sociodemographic and outcome tracking;
- Embed resilience goals into national legislative and budgetary instruments, aligning with ongoing public health reforms in Portugal.

These measures translate data into actionable governance, strengthen system responsiveness, and help ensure continuity of cancer

screening.

CRediT authorship contribution statement

Hugo Monteiro, M.D., M.P.H.: Conceptualization, Data Curation, Formal analysis, methodology, resources, visualization, writing – original draft, review and editing. **Mariana Oliveira, PhD**: Formal analysis, visualization, writing – original draft, review and editing. **Ricardo Martinho, PhD**: Writing – review and editing, supervision. **João Reis, M.D.**: Data curation, resources. **Fernando Tavares, M.D.**: Data curation, resources, validation. **Óscar Felgueiras, PhD**: Writing – review and editing, formal analysis. **Carlos Martins, M.D., PhD**: Writing- review and editing, supervision.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT from OpenAI to review text. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hugo Monteiro, João Reis and Fernando Tavares were employed in the organization responsible for the implementation of the screening programs at regional level. Mariana Oliveira, Carlos Martins, Óscar Felgueiras e Ricardo Martinho have nothing to declare.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jcpo.2026.100702.

References

[1] C. Rinner, E. Helm, R. Dunkl, H. Kittler, S. Rinderle-Ma, Process mining and conformance checking of long running processes in the context of melanoma surveillance, Int. J. Environ. Res. Public Health 15 (12) (2018) 2809.

[2] T. Li, B. Nickel, P. Ngo, K. McFadden, M. Brennan, M.L. Marinovich, et al., A systematic review of the impact of the COVID-19 pandemic on breast cancer screening and diagnosis, Breast J. 67 (2023) 78–88.

[3] L.M. Nogueira, B. Palis, D. Boffa, S. Lum, K.R. Yabroff, H. Nelson, Evaluation of the impact of the COVID-19 pandemic on reliability of cancer surveillance data in the National Cancer Database, Ann. Surg. Oncol. 30 (4) (2023) 2087–2093.

[4] E. Heer, Y. Ruan, D.J. Boyne, T.N. Jarada, D. Heng, J.-W. Henning, et al., Impact of the COVID-19 pandemic on cancer diagnoses, stage and survival in Alberta, Can. Med. Assoc. J. 195 (23) (2023) E804–E812.

[5] M. Angelini, F. Teglia, L. Astolfi, G. Casolari, P. Boffetta, Decrease of cancer diagnosis during COVID-19 pandemic: a systematic review and meta-analysis, Eur. J. Epidemiol. 38 (1) (2023) 31–38.

[6] J. Worthington, F. van Wifferen, Z. Sun, L. de Jonge, J.-B. Lew, M.J. Greuter, et al., Potential global loss of life expected due to COVID-19 disruptions to organised colorectal cancer screening, EClinicalMedicine (2023) **62**.

[7] R. Fujisawa, Impact of the COVID-19 pandemic on cancer care in OECD countries, OECD Health Work. Pap. (141) (2022), 0_1-43.

[8] D. Mendes, D. Figueiredo, C. Alves, A. Penedones, B. Costa, F. Batel-Marques, Impact of the COVID-19 pandemic on cancer screenings in Portugal, Cancer Epidemiol. 88 (2024) 102496.

[9] R. Cardoso, F. Guo, T. Heisser, M. Hackl, P. Ihle, H. De Schutter, et al., Colorectal cancer incidence, mortality, and stage distribution in European countries in the colorectal cancer screening era: an international population-based study, Lancet Oncol. 22 (7) (2021) 1002–1013.

[10] H. Brenner, T. Heisser, R. Cardoso, M. Hoffmeister, Reduction in colorectal cancer incidence by screening endoscopy, Nat. Rev. Gastroenterol. Hepatol. 21 (2) (2024) 125–133.

[11] L. Aversano, M. Iammarino, A. Madau, G. Pirlo, G. Semeraro, Process mining applications in healthcare: a systematic literature review, PeerJ Comput. Sci. 11 (2025) e2613.

[12] H. Ding, J. Lin, Z. Xu, X. Chen, H.H.X. Wang, L. Huang, et al., A global evaluation of the performance indicators of colorectal cancer screening with fecal immunochemical tests and colonoscopy: a systematic review and meta-analysis, Cancers 14 (4) (2022) 1073.

[13] H. Monteiro, M. Oliveira, J. Reis, F. Tavares, Optimizing colorectal screening in portugal with process mining, Eur. J. Public Health 34 (ement_3) (2024) ckae144.2110.

[14] E. Dekker, P.J. Tanis, J.L. Vleugels, P.M. Kasi, M.B. Wallace, Colorectal cancer, Lancet 394 (10207) (2019) 1467–1480.

[15] M. Pinheiro, D.N. Moreira, M. Ghidini, Colon and rectal cancer: an emergent public health problem, World J. Gastroenterol. 30 (7) (2024) 644–651.

[16] J. Carmona, B. Van Dongen, A. Solti, M. Weidlich, Conformance checking, 56, Springer, Switzerland, 2018, p. 12.

[17] J. Munoz-Gama, N. Martin, C. Fernandez-Llatas, O.A. Johnson, M. Sepúlveda, E. Helm, et al., Process mining for healthcare: characteristics and challenges, J. Biomed. Inf. 127 (2022) 103994.

[18] E. De Roock, N. Martin, Process mining in healthcare–an updated perspective on the state of the art, J. Biomed. Inf. 127 (2022) 103995.

[19] T.R. Gildea, S. DaCosta Byfield, D.K. Hogarth, D.S. Wilson, C.C. Quinn, A retrospective analysis of delays in the diagnosis of lung cancer and associated costs, Clin. Outcomes Res. (2017) 261–269.