

Article

Using Process Mining Techniques to Enhance the Patient Journey in an Oncology Clinic

Ricardo S. Santos ^{1,2,*} , Jaqueline B. Braz ³, Michelle Capelli ³, Alvaro O. I. Rodrigues ³
and José M. Parente de Oliveira ² 

¹ Compumédica Research & Innovation, São Paulo 04269-000, Brazil

² Department of Computer Science, Aeronautics Institute of Technology, São José dos Campos 12228-900, Brazil; parente@ita.br

³ Mogi das Cruzes Oncology Center, Mogi das Cruzes 08730-500, Brazil

* Correspondence: rsantos@cmpx.com.br

Abstract

The cancer care pathway comprises several stages encompassing diagnosis, treatment, and follow-up. Studies show that delays in treatment initiation are associated with worse outcomes, including increased mortality, reduced progression-free survival, and diminished post-treatment quality of life. To address this, patient navigation tools have emerged as a strategy to identify bottlenecks and mitigate delays. In this context, process mining offers a promising approach to discover, model, and optimize workflows using real data from hospital information systems. This paper presents a case study on the application of process mining to analyze care pathways in an oncology clinic. The focus was on identifying critical pathways and delays in the treatment journey to support the patient navigation program. Based on the insights gained, targeted improvement actions were proposed to enhance the patient journey. Using the PM² methodology, event data were extracted and processed from the clinic's information systems to model and analyze two key processes: (i) departmental workflows related to ambulatory care and (ii) longitudinal treatment pathways from initial evaluation to discharge. The results confirm the value of process mining for improving oncology patient journey and highlight its potential as a decision-support tool for healthcare administrators and clinical leaders.

Keywords: process mining; process discovery; cancer care unit; care pathways; patient care management

1. Introduction

The cancer care pathway comprises several stages encompassing diagnosis, treatment, and follow-up. Treatment delays remain a widespread challenge across global health systems and are consistently associated with worse outcomes, including increased mortality, reduced progression-free survival, and diminished post-treatment quality of life [1–4]. For instance, breast cancer therapy initiating more than 90 days after diagnosis has been shown to negatively affect disease-free survival and overall well-being among survivors [3]. Likewise, treatment initiation beyond 60 days is associated with reduced overall survival and a heightened risk of recurrence in head and neck squamous cell carcinoma [4].

To address such delays, health system administrators have implemented policies aimed at reducing the time from diagnosis to treatment, with the goal of improving population-level survival outcomes. One example is Brazil's "Sixty-Day Law", enacted in



Academic Editor: Jiang Bian

Received: 19 September 2025

Revised: 30 January 2026

Accepted: 2 February 2026

Published: 5 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

2012, which mandates that cancer patients begin treatment within sixty days of confirmed diagnosis [5]. Nevertheless, evidence indicates substantial noncompliance with this requirement. Pulido et al. [6] reported that nearly one-third of colorectal cancer patients did not initiate treatment within the legally mandated period, while Marcelino et al. [7] found no reduction in the diagnostic-to-treatment interval for melanoma patients, concluding that the policy was ineffective.

The growing need for patient journey navigation programs [8], combined with the increasing digitalization of clinical and operational workflows, has accelerated the adoption of advanced technologies aimed at improving efficiency and care quality. However, the complexity of the oncology patient journey poses considerable challenges for resource management.

In this scenario, the adoption of innovative technologies such as process mining has emerged as a promising approach for discovering, modeling, and optimizing processes based on real data extracted from hospital information systems [9]. Process mining enables the extraction, modeling, and analysis of actual workflows using event logs, thereby revealing nonconformities and bottlenecks. Evidence indicates that its application in hospitals can significantly enhance patient flow management and resource allocation, resulting in improved efficiency and higher-quality care [10].

Considering that in oncology care the effectiveness and adherence to clinical pathways are critical, given that delays in treatment can negatively affect clinical outcomes [11], the continuous monitoring of the patient journey is essential for improving both care quality and treatment results. In this context, process mining provides a robust analytical tool, enabling managers and clinical teams to detect deviations, identify bottlenecks and delays, and optimize care processes.

Thus, this paper presents a case analysis of the use of process mining to evaluate care pathways in an oncology outpatient clinic. The investigation focused on identifying critical pathways and process variants, as well as detecting delays throughout the treatment journey, with the aim of supporting the clinic's patient navigation program. Based on the insights obtained, targeted improvement actions were proposed to enhance the patient journey.

The study was conducted in a Brazilian oncology clinic that provides treatment for all cancer types and offers chemotherapy, radiotherapy, clinical oncology, and a range of supportive services, including psychotherapy, physical therapy, speech therapy, and nutritional counseling. The clinic receives approximately 150 new cases per month and treats an average of 65 patients per day.

2. Related Works

Currently, process mining applications in healthcare are a widely studied field, as evidenced by numerous literature reviews—both those covering a general domain [12–14] and those focused on specific areas, such as oncology [15].

Regarding the type of process mining application in healthcare, and based on the classification proposed by Santos Garcia et al. [16]—which includes process discovery, conformance checking, process enhancement, and supporting areas—the review by Aversano et al. [12] shows that most studies focus on process discovery (53%), followed by process enhancement (19%), support activities (17%), and, finally, conformance checking (11%).

Although Guzzo et al. [14] adopted a slightly different classification for process mining types, their review presents results quite like Aversano's, showing that the major process mining applications in healthcare focus on process discovery, followed by process analysis (process enhancement) and conformance analysis (conformance checking).

None of the reviews provided a straight justification for the higher frequency of process discovery applications, compared to conformance checking or process enhancement.

One possible hypothesis for this phenomenon is that process discovery is inherently the first step in any process mining approach. To improve a process or check its conformance, it is first necessary to discover and understand it.

This rationale reflects the objective of our case study, which focuses on discovering the processes (and their variations) followed in daily practice and assessing their performance, with the aim of improving and optimizing the efficiency of the patient's journey.

We are not necessarily interested in the official processes designed by departmental management, as these may not be followed as intended. It is important to note that, even to evaluate adherence to institutional processes, it is essential to first identify the processes that are actually being executed.

A second plausible hypothesis is the natural variability of clinical processes, which makes adherence to institutional protocols very difficult. Accounting for all the clinical variables that can affect a patient's journey in a standardized process presents a significant challenge.

We can infer that Guzzo et al. indirectly acknowledge this hypothesis by stating that "guidelines are difficult to accomplish given that, for example, patients suffering from the same disease may require different treatments depending on how their body reacts to certain drugs and medications. For this reason, processes carried out within the same medical context may differ from each other in terms of control-flow, organizational, and temporal perspective."

In this section, we do not aim to conduct an exhaustive literature review, but rather to identify selected case studies related to our work. Likewise, we do not seek to make direct comparisons between the studies reviewed and our own, but instead to highlight relevant similarities and differences.

Oncology was one of the first healthcare fields to apply process mining. In 2008, Mans et al. [17] used process mining to analyze patient flows in gynecological oncology. The overall goal of their study aligns with ours: to gain meaningful insights into care pathways by discovering typical patients' flows in oncological treatment.

However, their work differs from ours in two main aspects. The first is the scope of the data. While our study includes all tumor types, theirs focuses exclusively on gynecological tumors. Additionally, their raw data came from treatment records collected for financial purposes, whereas our data was obtained from clinical information in electronic medical records. This distinction may lead to the discovery of more detailed workflows in our study, as financial records often show group care activities for billing purposes.

The second difference lies in the perspective of process mining application. Mans et al. examined the healthcare process from three perspectives: control flow, organizational, and performance. In contrast, our focus is primarily on the treatment flow beyond the organizational structure of an outpatient clinic. By analyzing treatment flows, we can explore the differences across various cancer types and generate insights to improve care pathways based on these differences.

Another related work on the use of process mining to discover oncology care pathways was conducted by Pijnenborg et al. [18]. In their study, the authors investigated the application of process mining techniques to identify and analyze palliative care pathways for stomach and esophageal cancer. The aim was to develop an evidence-based understanding of which palliative treatments are commonly administered in clinical practice and how these treatments relate to patients' survival times.

Regarding data origin, the dataset used in Pijnenborg et al.'s research was obtained from the Netherlands Cancer Registry (NCR), a national oncology registry that includes

data on all cancer patients across the country. This data source is particularly suitable for their goal of identifying variations in care pathways across multiple hospitals for a specific cancer type.

In contrast, our case study focuses on a single healthcare provider—an outpatient clinic—with the objective of discovering different operational workflows across various cancer types.

Although the focus of Savino et al.'s work [19] differs from ours, they also applied process mining to analyze cancer treatment pathways. Specifically, they investigated the adherence of rectal cancer patients treated at an Italian university hospital to the clinical guidelines established by the European Society for Medical Oncology.

Before analyzing the adherence to guidelines, their first step was process discovery—to provide an overview of the real treatment processes and generate an intuitive representation to analyze deviations from the guidelines.

The approach adopted by Savino et al. reinforces the strategic vision of our project: to first discover and analyze the actual processes within the outpatient clinic, as presented in this work, and subsequently identify deviations and delays in these workflows. This will support efforts to improve the patient's journey and enhance care efficiency.

Another relevant study was conducted by Baker et al. [20], whose goal was to discover and quantify patient pathways during chemotherapy treatment to support the development of a decision model for monitoring neutropenic status.

Although the scope and focus of that study differ from those of the present work—since it is limited to chemotherapy treatment and aims to identify process variants related to the analysis of neutropenic status, whereas our study focuses on the overall patient journey encompassing all treatments involved—both studies highlight the huge number of pathway variants present in oncological treatment care.

In the Brazilian context, there are also initiatives applying process mining to oncology. The study by Iachecen et al. [21] is the most closely aligned with ours. Their work aims to uncover the care journey of lung cancer patients within a Brazilian health insurance provider, identifying the main diagnostic tests and treatment modalities used.

While the objectives of both studies are similar, there are key differences. First, the practical focus diverges: Iachecen et al. emphasize identifying the primary diagnostic tests and treatment modalities, whereas our study focuses on identifying critical pathways and variations that highlight inefficiencies in the patient's journey.

Another distinction lies in scope. Iachecen et al. limit their analysis to lung cancer, while our study addresses multiple cancer types. Additionally, the data source differs significantly. The event log used by Iachecen et al. is derived from a sample of beneficiaries of a health insurance provider and primarily consists of data used for billing and reimbursement. A positive aspect of this dataset is that it covers multiple healthcare providers and diverse patient demographics—unlike ours, which is limited to a single cancer center. However, such administrative data may not capture all events necessary for constructing a detailed view of the patient flow, which is critical for our study's purposes. For instance, chemotherapy involves a series of preliminary steps (such as triage, preparing, and dispensing the medication, etc.) that are likely not included in the records submitted to the insurer.

Related work indicates that, despite numerous initiatives applying process mining in the field of oncology, there remains a lack of studies specifically focused on outpatient oncology treatment clinics—underscoring the relevance and justification of the present study.

3. Methodology

Considering that the case study reported in this paper involves the application of process mining, we followed the PM² methodology, a widely used approach for developing process mining projects, proposed by van Eck et al. [22]. This methodology comprises six steps: planning, extraction, data processing, mining and analysis, evaluation, and process improvement (Figure 1).

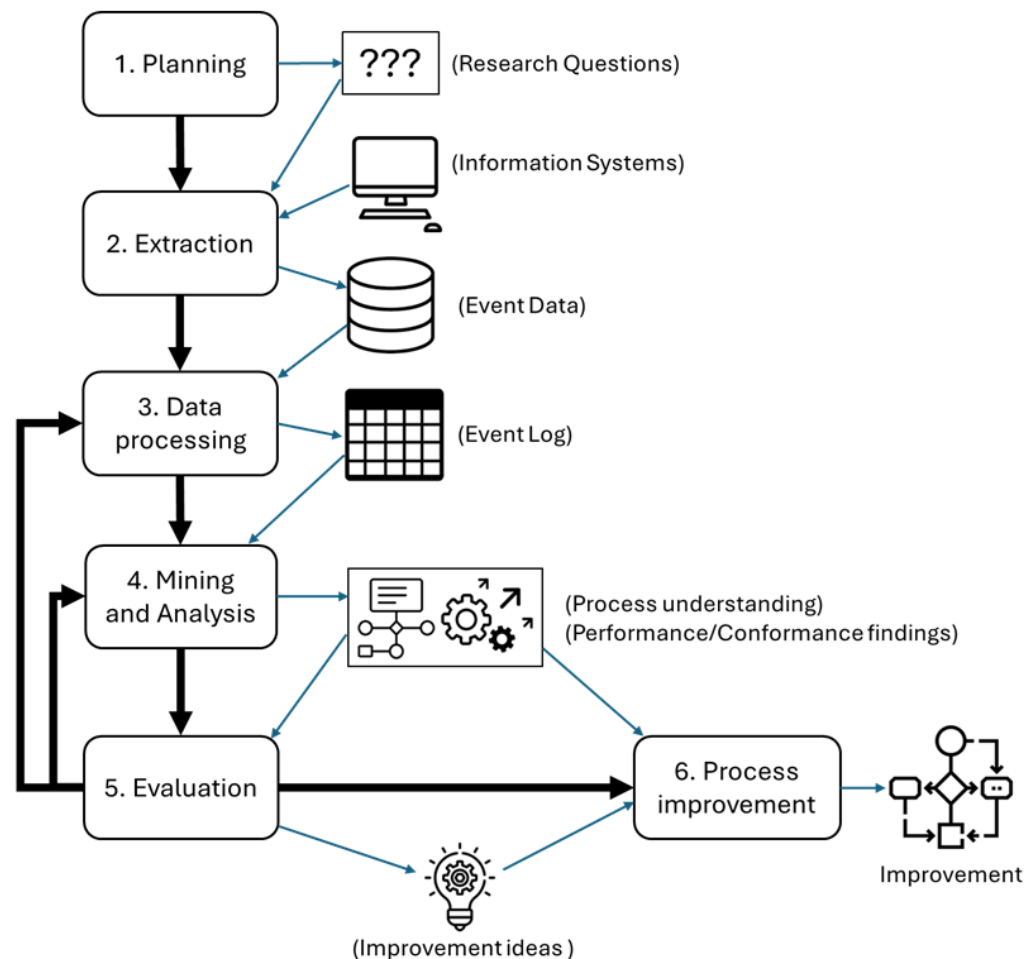


Figure 1. PM2 methodology overview. (Adapted from van Eck et al. [22]).

The essence of the PM² methodology is to transform project goals (for example, achieving a 10% reduction in the execution time of a given process) into research questions that can be answered, generating findings that provide a foundation for improving the selected process.

In the first stage (Planning), the research questions are defined based on the goals of the process mining project. The second stage (Extraction) involves retrieving event data from the relevant information systems. In the third stage (Data Processing), the extracted data are structured and stored in a standardized event log format suitable for analysis.

The fourth stage (Mining and Analysis) applies process mining techniques to the event logs to address the research questions. This stage yields findings related to process performance and compliance or, in the case of more abstract research questions, offers a broader understanding of the process under investigation.

In the fifth stage (Evaluation), the results obtained from the mining and analysis phase are evaluated to derive potential process improvement ideas or to formulate new research questions. The PM² methodology considers the iterations between stages 3 and 5, where the

results of the evaluation stage may indicate the need to develop new data mining models or even return to the data processing stage.

Finally, in the sixth stage (Process Improvement), the proposed improvements are implemented through modifications to the existing processes.

Despite the existence of a healthcare-specific version of PM², known as ClearPath, which addresses aspects such as process simulation and the involvement of clinical specialists [23], we chose to adopt the standard PM² methodology, as this study focuses primarily on the operational processes of the patient journey.

The case study presented in this paper has two practical objectives: (i) to minimize delays in the initiation of each stage of a patient's treatment and (ii) to reduce the patient's length of stay in the clinic during outpatient care services (e.g., chemotherapy or radiotherapy sessions, follow-up appointments). To achieve these objectives, two types of process analyses are required: process discovery and process enhancement (or performance checking).

Through process discovery, the goal is to identify the pathways followed by patients during outpatient care (Department Workflow) and throughout the entire treatment journey (Treatment Workflow). This analysis aims to map critical paths and key process variants. Critical paths warrant particular attention because they are more susceptible to delays and bottlenecks. Therefore, identifying and analyzing these critical paths can provide support for the adoption of monitoring policies and tools, as well as for improving the clinic's operational processes, making them more agile.

In the context of this case study, a critical path is defined as one that concentrates a high volume of patients. The clinic's management established that a path is considered critical if it involves a percentage of patients exceeding 50% above the uniform distribution across all possible paths. For example, if there are five paths, any path that accounts for more than 30% of the patients would be classified as critical.

Process discovery analysis was performed by the Fluxicon Disco[®] v. 4.2.6. (Fluxicon BV, Eindhoven, The Netherlands) process mining tool.

The performance checking analysis focused on identifying delays within the clinic's institutional operational workflow (Figure 2). In this case study, delays were identified using two metrics: average activity execution time and average patient waiting time. The clinic's quality management team established target times and upper limits for each metric. For example, for the Consultation/Evaluation activity, the target execution and waiting times were 30 and 20 min, respectively, with corresponding limits of 45 and 30 min. Any activity whose average execution or waiting time exceeded these limits was classified as a delay.

The same approach was applied to the treatment workflow (Figure 3), with the difference that time metrics were defined in days rather than hours.

Performance checking was conducted using the ONCOPATHWAYS[®] v. 0.5.1. (Com-pumedica Pesquisa e Inovação Ltda, São Paulo, Brazil) tool, a software application developed within this project to monitor delays in the patient's journey in real time and to present some other performance metrics to users.

The following sections describe the methods applied in each PM² process activity of the case study.

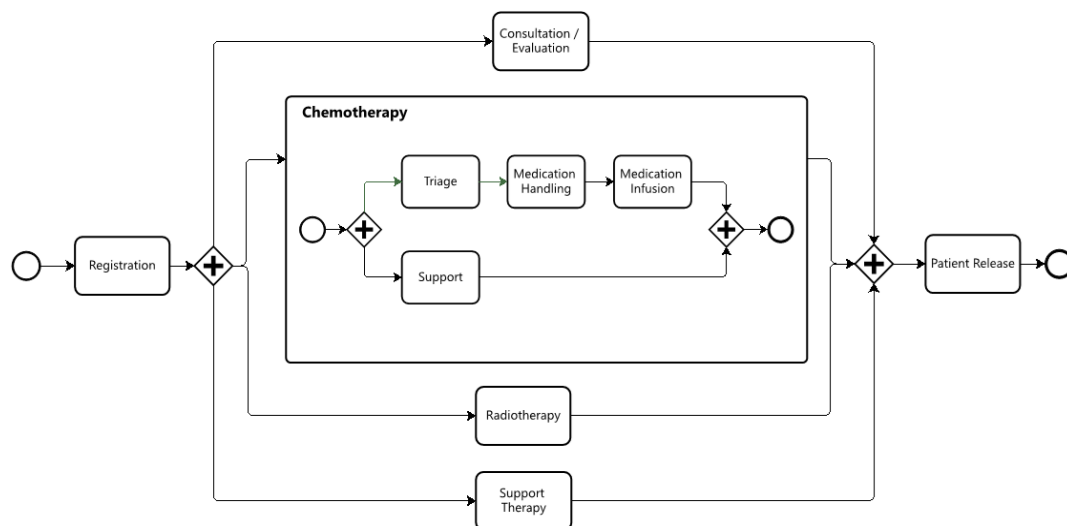


Figure 2. Institutional operational workflow (Department Workflow).

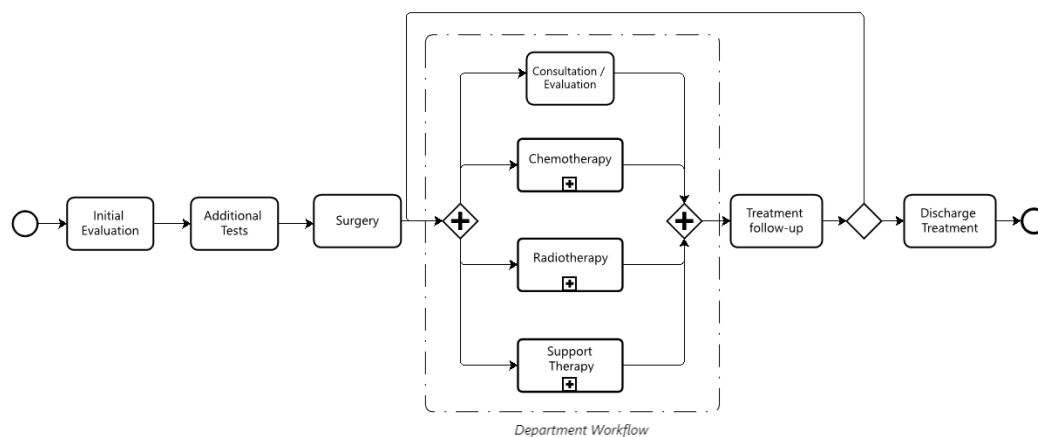


Figure 3. Treatment workflow.

3.1. Planning

The inputs to this phase are the organization's business processes, and the outputs include the definition of research questions and the identification of information systems from which event data can be extracted.

This case study examined two distinct business processes relevant to the oncology patient journey. The first concerns patient flow within the clinic during chemotherapy or radiotherapy sessions, referred to as the Department Workflow (Figure 2). The research questions formulated for this workflow aimed to identify critical paths, those with the highest patient volume.

The second business process analyzed is the patient's treatment journey (Figure 3). Some research questions for this process sought to examine variations in patient flow based on cancer type. Understanding these process variants can support more effective resource planning for the care activities associated with each treatment pathway. Additional research questions aimed to identify patients who experienced significant delays or remained in waiting status at various stages of the process.

3.2. Extraction

All events related to the business processes are stored in the clinic's management system, which encompasses everything from the electronic health record (EHR) to the

billing module. This is a proprietary system developed by the clinic's own IT team, and its database is managed using Firebird.

The event data were extracted using SQL scripts validated and executed by the clinic's IT team. No personal data from patients or healthcare professionals were extracted from the system. Only the data necessary for constructing the event log were retrieved: event identifier, executed activity (which may be a clinical or administrative procedure), execution start and end timestamps, scheduled date and time, patient arrival timestamp, executor identifier and resources used to execute the task.

The extraction period for events related to the treatment workflow was from January 2020 to May 2025, and for the department workflow, January to May 2025. No inclusion or exclusion criteria were applied other than the period of analysis.

The extracted event data were stored in a PostgreSQL (v 16) database.

3.3. Data Processing

From the extracted event data, two event logs were created: one containing the events related to the department workflow and the other to the treatment workflow.

To enable more detailed analyses of the clinic's processes, additional columns were included in the event logs. The added data includes: the ICD-10 code, to identify the type of tumor; the first-line therapeutic approach (surgery, chemotherapy, radiotherapy, etc.); the patient's age group; and gender.

These event logs were stored in a PostgreSQL database and also exported to CSV format, which is one of the formats supported by Fluxicon Disco.

3.4. Mining and Analysis

For each research question defined during the planning stage, a process mining model was generated using one of the tools to address the corresponding question.

The process discovery analysis was also applied across different data segments, including time periods, cancer types, age groups, and patient gender, to assess stability and consistency across these samples. For the analysis of variants by cancer type, the investigation was limited to three groups—breast, prostate, and digestive organ cancers—which together account for nearly 70% of the cases treated at the clinic.

Performance-checking metrics are obtained directly from the ONCOPATHWAYS monitoring tool. For activities in which delays were identified, the data were exported to an Excel spreadsheet, enabling individual analysis of cases that exceeded the defined time limits. Table 1 presents the target times defined for the activities in the Department Workflow, while Table 2 shows the target times for the stages of the Treatment Workflow. As shown in Table 2, the analysis of the treatment workflow's performance is limited to waiting times.

Table 1. Target time for department workflow.

Department Task	Task Time (min)		Waiting Time (min)	
	Target	Limit	Target	Limit
Registration	10	15	10	15
Consultation/Evaluation	30	45	20	30
Medication Handling	9	10	45	30
Medication Infusion	330	360	45	60
Chemo—Support	45	60	45	60
Radiotherapy	30	60	30	60
Support Therapy	30	60	30	60

Table 2. Target time for treatment workflow.

Treatment Stage	Waiting Time (Days)	
	Target	Limit
Initial Evaluation	30	45
Additional Tests	30	45
Chemotherapy	30	45
Radiotherapy	30	45
Support Therapy	30	45
Treatment Follow-up	--	--

3.5. Evaluation

The evaluation of the models and indicators generated by the process mining tools was performed manually by three clinic professionals: the administrative manager, the quality manager, and the nursing manager.

The evaluation procedure consisted of three main activities: (i) verifying the accuracy of the results, (ii) analyzing the root causes of deviations and delays, and (iii) consolidating these causes into a structured table.

For the process discovery analysis, the assessment involved determining whether the identified critical paths reflected the clinic's routine practice and whether the observed variations actually occurred. Variations classified as deviations from the institutional process were subsequently selected for further investigation, which at times required internal auditing procedures.

For the performance analysis, the evaluation focused on confirming whether delay alerts had been triggered appropriately. For each workflow activity flagged as a delay, all instances exceeding the predefined thresholds were examined to identify the underlying causes of delays. These causes were then systematically stored in an Excel spreadsheet.

Given that the data sample used by the process mining tools was anonymized, the analysis of a specific case of deviation or delay required evaluators to review all cases from that day to identify the case flagged by the tool.

3.6. Process Improvement

Based on the spreadsheet consolidating issues related to deviations and delays, the clinic's management team, comprising managers from the involved departments, identified a set of improvement actions. The complete set of improvement proposals was subsequently submitted to the company's board of directors for approval and implementation.

4. Results and Discussion

The results are presented and discussed according to the steps of the PM² methodology.

4.1. Planning

The research questions were developed in collaboration with the clinic's clinical and operational management teams.

For the Department Workflow, the following questions were defined:

- What is the most common patient flow in the clinic's daily operations?
- What is the percentage of significant deviations from the institutional workflow? (Significant deviations are considered those that should not occur; for example, a patient undergoing a procedure without prior registration at the front desk.)
- What is the average process execution time?
- Which activities have average patient waiting times exceeding the defined thresholds?

For the Treatment Workflow, the following questions were formulated:

- What is the most common treatment pathway in the clinic?
- Are there significant differences among the treatment pathways for breast, prostate, and digestive organ cancers?
- Considering these three types of cancer, which treatment stages have starting times that exceed the defined thresholds?

4.2. Extraction and Data Processing

Two SQL scripts were developed to extract event data from the clinic's management system: one for events related to the Department Workflow and another for events related to the Treatment Workflow. The first resulted in a database table with 13,400 rows, while the second extracted 67,469 rows.

These tables were exported as CSV files. The table related to Department Workflow events generated six CSV files: one containing the entire period and five with records separated by month, from January to May 2025. The table related to Treatment Workflow events produced seven CSV files: one containing the full period from 2020 to 2025, and six with records separated by year.

The creation of multiple log files aimed to facilitate comparative analysis over time, as the analysis tool supports reading and processing multiple datasets simultaneously.

The CSV files were utilized for analyses conducted with Fluxicon Disco, whereas ONCOPATHWAYS performed its analyses by directly processing data stored in the PostgreSQL database.

4.3. Mining and Analysis

Based on the analysis of the Department Workflow, a process model was generated covering the entire period from January to May 2025 (Figure 4). This model addresses the research question: "What is the most common patient flow in the clinic's daily operations?"

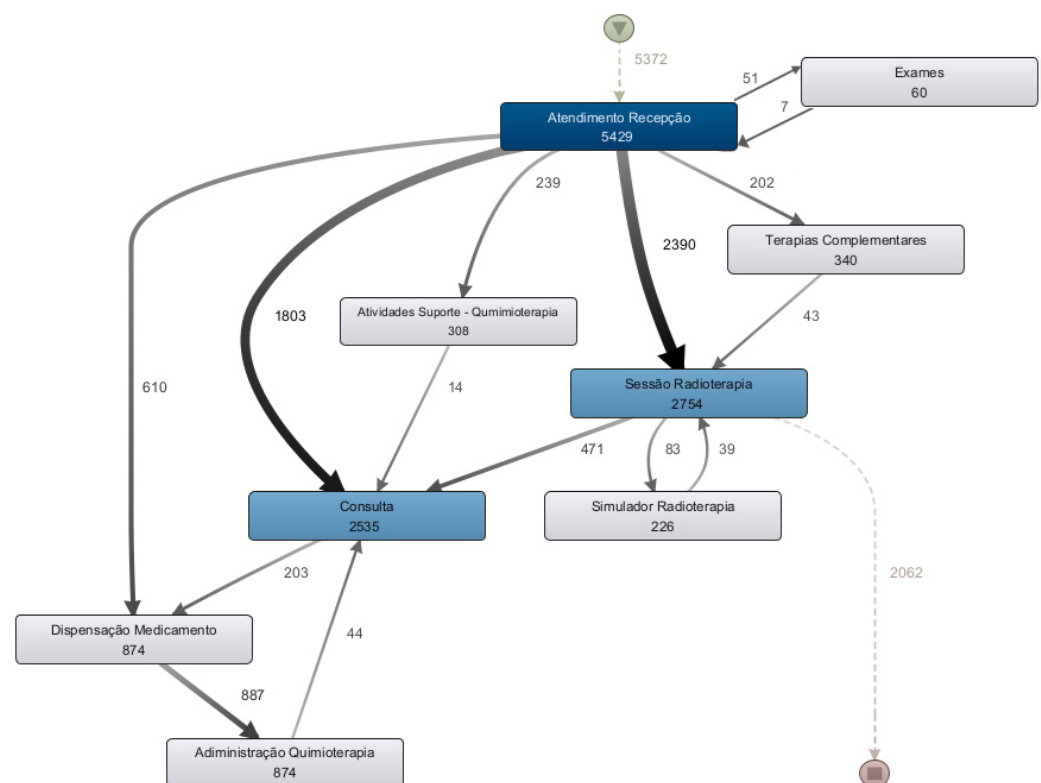


Figure 4. Department process model—January 2025 to May 2025 (main paths).

The most frequent activity flow observed in the clinic corresponds to Registration → Radiotherapy, in which the patient checks in at the front desk (“*Atendimento Recepção*”) and then proceeds to the radiotherapy session (“*Sessão Radioterapia*”), with 2754 occurrences.

The second most frequent path is the Registration → Consultation/Evaluation, where the patient registers at the reception and then proceeds to a medical consultation (“*Consulta*”), with 2535 occurrences.

Chemotherapy infusion represents the third most common flow at the clinic, with 874 occurrences. This flow includes patient check-in at the reception, followed by medication dispensing by the pharmacy (“*Dispensação Medicamento*”) and subsequent drug administration to the patient (“*Administração Quimioterapia*”).

A closer examination of the process model shown in Figure 4 reveals the presence of several flow variants. For example, when analyzing the standard chemotherapy administration sequence (Registration → Medication Handling → Medication Infusion)—shown in Figure 4 in Portuguese as *Atendimento Recepção* → *Dispensação Medicamento* → *Administração Quimioterapia*—610 occurrences followed the direct path from Registration to Medication Handling. However, a total of 874 Medication Handling events were recorded. This discrepancy indicates the existence of a process variant: in some cases, patients check in at reception, then undergo a medical consultation before proceeding to medication handling (dispensing) and, subsequently, to medication infusion.

The process model presented in Figure 4 also allows us to address the research question: “What is the percentage of significant deviations from the institutional workflow?”. A closer analysis of the model reveals some discrepancies that correspond to deviations from the institutional process. For example, it is observed that the process starts with 5372 occurrences; however, there are 5429 check-in records at reception (*Atendimento Recepção*). This indicates that some cases (57) checked in more than once, which represents a deviation from the expected workflow. In this case, the deviation rate is only 1%.

Another deviation can be seen in the medication handling (*Dispensação Medicamento*) activity. While there are 874 occurrences of dispensing, the model shows 887 outgoing transitions from this activity leading to drug administration. This also represents a small deviation, with a similar rate of approximately 1%.

Another way to identify process deviations is by expanding the visualization of process variants. When displaying the process model with 50% of all variants (Figure 5), it is possible to detect a deviation from the institutional workflow: patients who bypassed reception and proceeded directly to medication dispensing (24 cases) or to a medical consultation (36 cases).

To address the research question “What is the average execution time of the process?”, we used the “Global Statistics” feature of the Fluxicon Disco tool (Figure 6). The observed average execution time was 92.8 min. The chart shows that most cases have an execution time of less than 1 h and 43 min, which aligns with the clinic’s most frequent workflow: radiotherapy and medical consultations. Both activities are typically quick, usually completed in under an hour.

The longest workflow in an oncology outpatient setting is chemotherapy administration, which ranges from 3 to 6 h. However, these cases represent only 16% of the total occurrences.

Figure 6 also shows that the Department Workflow process contains 148 variants. Additionally, different sample segments were analyzed, including period, cancer type, age group, and patient gender. The process findings remained consistent across all segments, indicating that these variables do not significantly impact the clinic’s operational workflow.

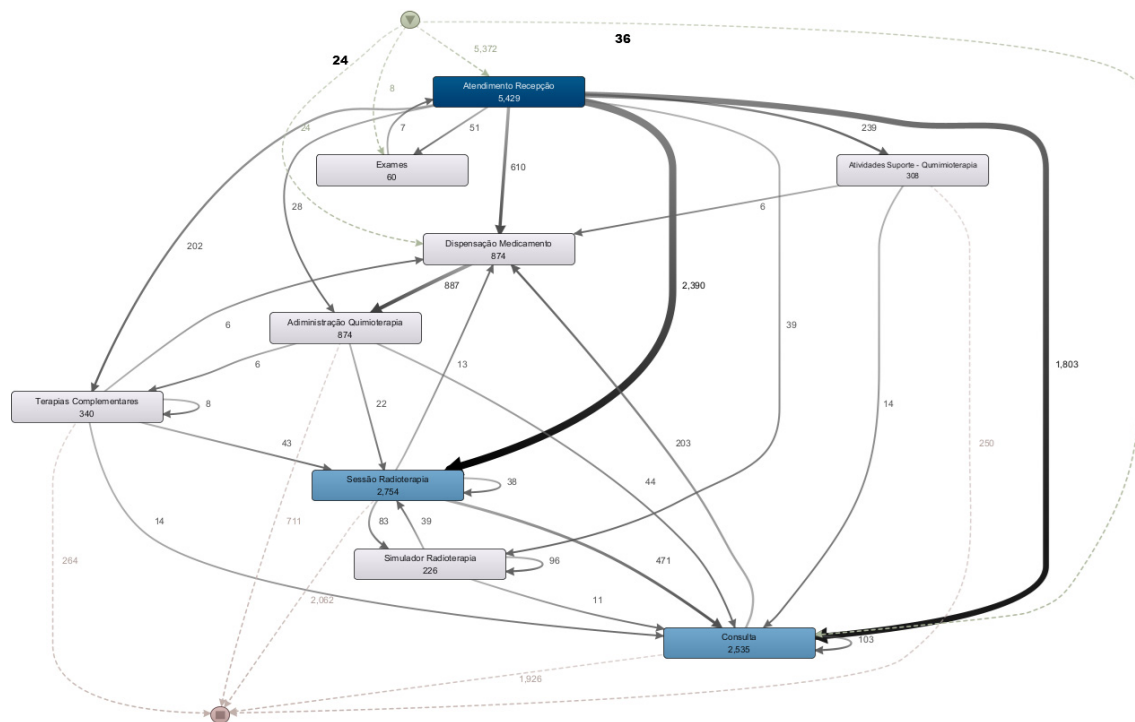


Figure 5. Department process model—January 2025 to May 2025 (top 50% of variants).

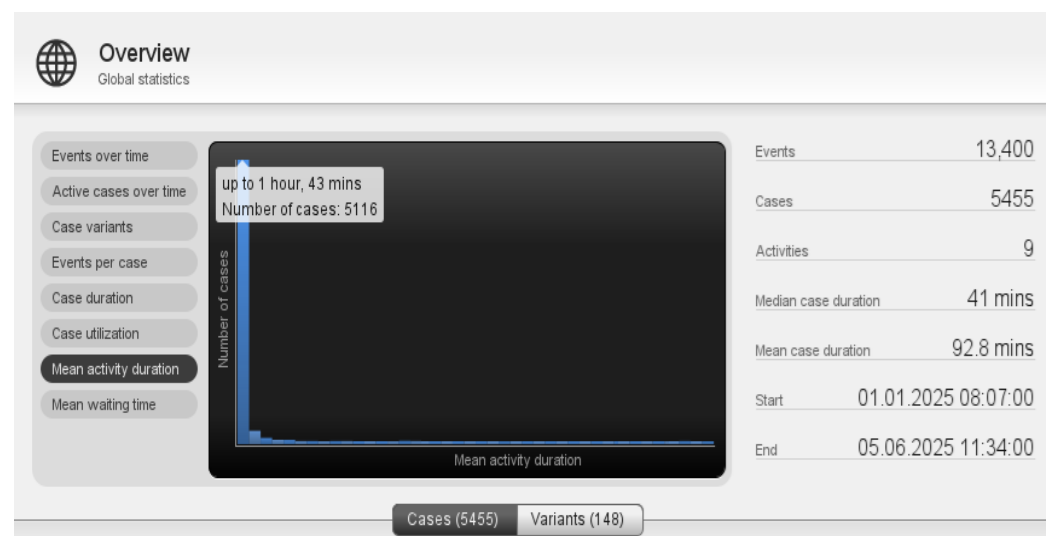


Figure 6. Global statistics of the department process.

From the analysis of the Treatment Workflow, a process model was generated covering the period from 2020 to 2025 (Figure 7). This model addresses the research question: “What is the most common treatment pathway in the clinic?” The results revealed that the most frequent treatment was radiotherapy (*Radioterapia*), undertaken by 1355 patients, followed by chemotherapy (*Quimioterapia*), administered to 490 patients.

The prevalence of radiotherapy as the most frequent treatment in the model is explained by the clinic’s role as a regional reference center for this service, receiving many patients who undergo oncological treatment at other institutions but attend this clinic specifically for radiotherapy.

The second research question regarding treatment workflow to be addressed is: “Are there significant differences among the treatment pathways for breast, prostate, and digestive organ cancers?” The results of the process model analysis for each of these cancer types

indicate that there are. When comparing the treatment process model for prostate cancer (Figure 8) with that for colon cancer (Figure 9), it becomes evident that the former presents a more clearly defined flow, with radiotherapy as the predominant treatment approach. In contrast, chemotherapy is more prevalent in the treatment of colon cancers. Another noteworthy aspect is the higher incidence of supportive therapy (*Terapias Suporte*) in cases of colon tumors compared to prostate cancer cases.

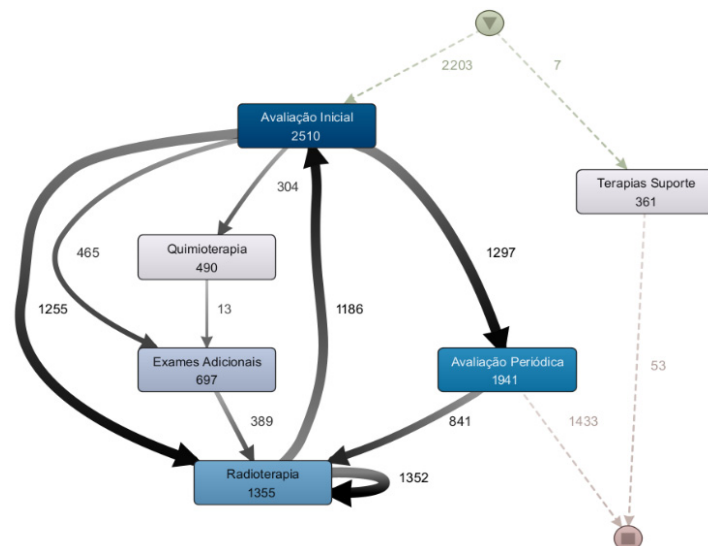


Figure 7. Treatment process model—2020 to 2025.

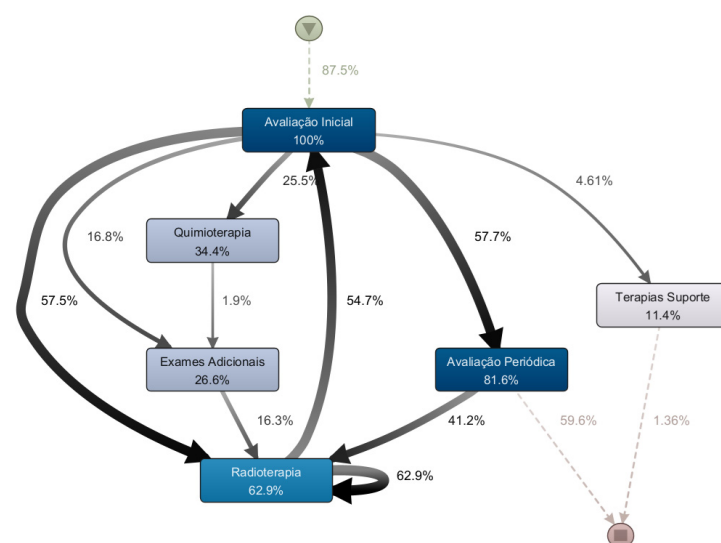


Figure 8. Treatment process model—prostate cancer.

Different sample segments were analyzed, including period, age group, and patient gender; however, the treatment process findings remain consistent across these segments.

The ONCOPATHWAYS tool was used to address questions related to performance checking. Figure 10 presents the ambulatory journey monitoring screen, which provides an answer to the question: “Which activities have average patient waiting times exceeding the defined thresholds?”.

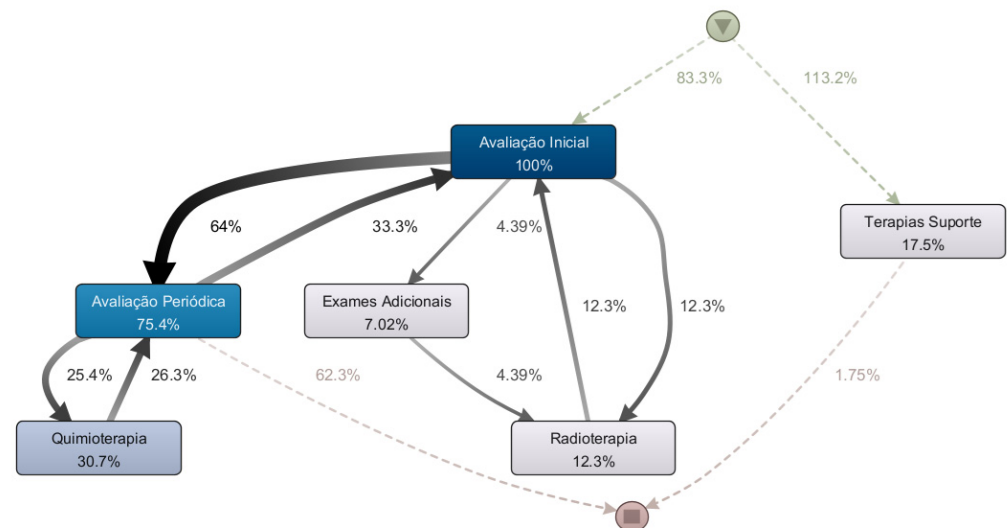


Figure 9. Treatment process model—colon cancer.

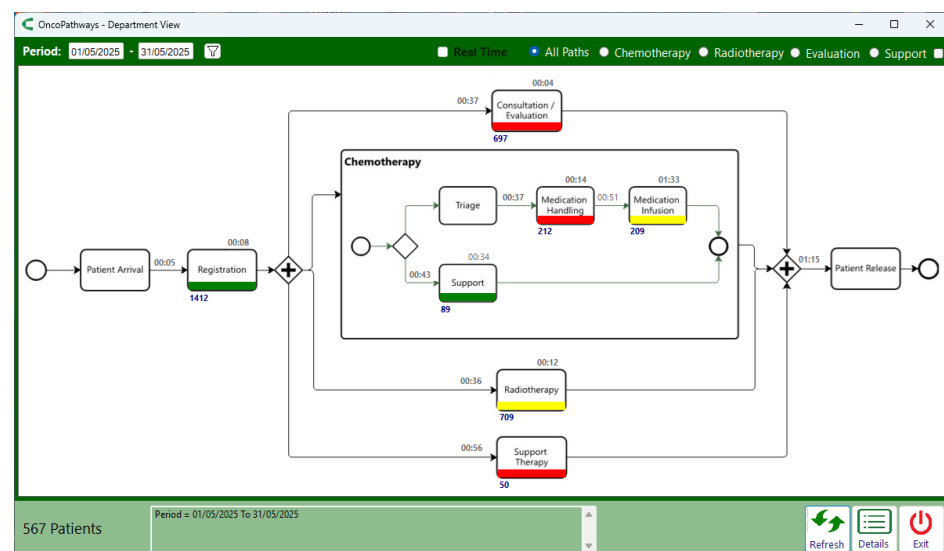


Figure 10. Department workflow monitoring. The colored bars represent: green = within the target time; yellow = outside the target time but within the limit time; red = outside the limit time.

Delays were identified in three activities: Consultation/Evaluation, Medication Handling, and Support Therapy. In the case of Medication Handling, the delay was flagged due to task execution time exceeding the defined limit, whereas in the other two activities the delays were attributed to patient waiting times. Task execution time is represented by the number inside each box, while patient waiting time corresponds to the number above the arrows connecting the boxes. For the Consultation/Evaluation activity, this indicates that the average waiting time for patients to be called for a medical evaluation was 37 min. The number below each box represents the qty of patients who underwent that activity; for Consultation/Evaluation, this amounted to 697 patients.

The treatment journey monitoring screen (Figure 11) addresses the question: “For these three types of cancer, which treatment stages have start times that exceed the defined thresholds?”. Considering the period represented in Figure 11 (January–May/2025) and selecting the three most frequent cancer types (breast, prostate, and digestive organ), no delays were flagged, and treatment initiation times were generally favorable, with an average of 28 days for chemotherapy and 27 days for radiotherapy. However, the longest

waiting times—shown above the boxes—indicate that at least one patient waited 62 days to begin chemotherapy, while another waited 127 days to begin radiotherapy.

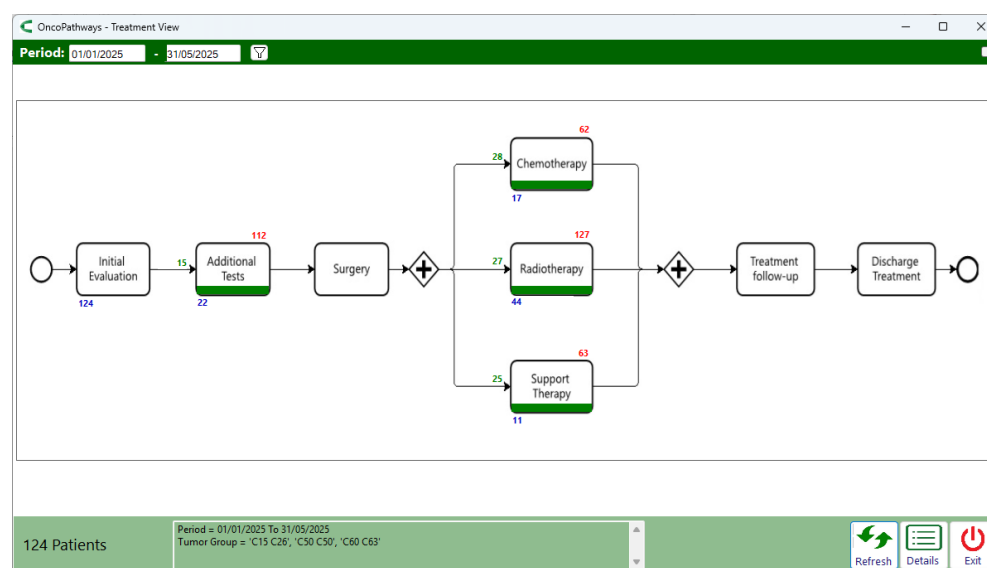


Figure 11. Treatment workflow monitoring. The colored bars represent: green = within the target time; yellow = outside the target time but within the limit time; red = outside the limit time.

The chemotherapy details screen (Figure 12) shows that 6 of the 17 patients who underwent chemotherapy began treatment after the target threshold of 30 days. In the case of radiotherapy, one patient started after 127 days; this was the same patient who had previously been flagged for a 118-day delay in the Additional Tests activity, which directly contributed to the postponement of radiotherapy.

Desc	Initial Evaluation	Date Event (min)	Date Sec. Event	Waiting Time (days)	Tumor Type
QT	2025-01-22	2025-03-26		063	C189
Cir+QT+VO	2025-02-12	2025-04-14		061	C509
QT+Imuno+ Cir	2025-01-10	2025-02-26		047	C509
HT+QT	2025-02-13	2025-03-27		042	C610
QT	2025-04-11	2025-05-13		032	C509
QT	2025-02-28	2025-03-31		031	C509
Cirurgia + HT	2025-02-26	2025-03-27		029	C509
HT	2025-04-09	2025-05-06		027	C610
Cir+ QT	2025-03-18	2025-04-14		027	C189
Cir+ QT	2025-03-07	2025-03-31		024	C509
Cir+HT	2025-02-03	2025-02-25		022	C610
HT	2025-02-27	2025-03-19	2025-04-14	020	C610
HT	2025-02-19	2025-03-07		016	C610
Cir+QT+RT	2025-01-22	2025-02-03		012	C259
QT + RT	2025-03-19	2025-03-26	2025-04-22	007	C190
HT + QT	2025-04-01	2025-04-08		007	C610
OncoBca + Cir+RT	2025-04-23	2025-05-12	2025-04-14		C610

Figure 12. Treatment workflow monitoring—Details.

All delays flagged in the workflow monitoring screen (Figures 10–12) were individually analyzed to identify their root causes. The results of this analysis are presented in the following section.

4.4. Evaluation and Process Improvement

The validation of the process mining models confirmed the accuracy of the system across all measured results. In cases where the tool indicated undue delays, the issue stemmed not from algorithmic error but from inaccuracies in data entry within EHR.

Three process deviations identified through the process discovery analysis were examined in detail by the quality team. The first involved a higher number of outgoing transitions from medication dispensing (887) compared to drug administration (874). Investigation revealed that this discrepancy resulted from medications being dispensed in separate batches, an accepted practice when different lots of the same drug are used. Consequently, this deviation was not considered problematic.

The second deviation concerned cases in which patients completed reception registration more than once (1% of occurrences). This occurred when care activities were delivered in different physical areas of the clinic, for example, a medical consultation followed by a physical therapy session requiring check-in at two separate reception desks. Although not a major issue, an improvement was recommended.

The third deviation involved instances in which patients received a care service, such as consultation or medication dispensing, without first checking in at reception. Analysis conducted by the quality team revealed that 90% of these cases were linked to delivery services offered by the Clinic (e.g., oral medications sent directly to patients' homes). However, some cases resulted from physicians calling patients directly for appointments before they registered at reception. For this latter situation, a process improvement was proposed.

The review of delay indicators from the department flow monitoring screen revealed that the start and end times of activities recorded in the EHR are often inaccurate, as the system does not require these timestamps to be entered at the exact moment the events occur. For example, a physician may call a patient into the office without immediately entering the start time in the system or may call the next patient without first recording the end time of the previous consultation. This limitation makes it impossible to accurately analyze delays in these activities. Nonetheless, process improvement suggestions were submitted to the EHR development team.

Regarding the delay identified in the medication handling activity, the issue did not stem from inaccurate EHR data, which had been correctly entered, but from an inadequately defined time limit (10 min). More complex drug formulations may legitimately require additional preparation time, an aspect not considered when the limit was established. The quality team submitted a recommendation to address this issue.

Although the treatment workflow monitoring screen did not have flag delays based on average times, the tool did identify individual cases of delay, as shown in the red box in Figure 12. Each of these cases was subsequently investigated. Delays attributable to a patient's personal decisions were excluded from the analysis. Among the remaining cases, most were linked to delays or denials in authorization of care procedures by healthcare insurance companies, which required patients to seek alternative solutions and ultimately postponed treatment initiation. Based on these findings, the quality team proposed several recommendations to mitigate delays in treatment initiation.

The problems investigated, their causes, and the corresponding suggestions for process improvement are presented in Table 3.

The clinic manager initiated a project to implement improvements based on the recommendations presented in Table 3. By the time of paper submission, the clinic had already implemented two of these improvements, corresponding to the recommendations listed in the second and third rows of the table.

Table 3. Problems and suggestions for process improvement.

Problem	Cause	Improvement
More than once check-in at reception desk	Care activities take place in different physical locations	Improve the EHR scheduling algorithm to optimize the patient flow within the clinic
No check-in at reception desk	Patients called for an appointment before registering at the reception desk	Include a constraint in the EHR to forbid the start of a care task without registration at the reception desk or without entering the end time of the previous appointment
Inaccuracy in the record of start and end times	EHR system does not require this information to be entered at the exact moment the events occur	Require this information as mandatory, entering it automatically using the system timestamp with a single click by the care professional
Inadequacy in the definition of the task's time limit	Complex formulas require additional preparation time	Allow multiple time limits to be set for each activity according to its different classes
Delays in starting treatment	Delays or denials of procedures authorization by the healthcare insurance	Send advance alerts to patient navigators to inform them of pending approval cases Develop a predictive model to estimate the likelihood of delays and/or denials for procedures authorization

The case study highlights several lessons for the study site that can also be applied to other outpatient oncology clinics.

First, the analysis demonstrated that even in a highly complex and variable clinical environment, a relatively small number of care pathways concentrate most patient flows. Making these dominant paths explicit allowed the clinic's managers to distinguish between acceptable clinical variability and true operational deviations. This distinction is essential for prioritizing improvement actions, separating rare or clinically justified exceptions from patterns that represent systematic inefficiencies and delays.

A second key lesson concerns the central role of data quality and information system design. Many of the delays identified in departmental workflows were not caused by resource shortages or clinical decision-making, but rather by inaccuracies in the recording of start and end times in the EHR. This finding was not a surprise, as it is a well-documented issue in healthcare process mining projects, as noted by Muñoz-Gama et al. [24]. As a result, the clinic recognized that investments in improving data capture mechanisms—such as enforcing real-time timestamping and embedding workflow constraints into the EHR—constitute high-impact process improvement measures.

Finally, the study showed the importance of moving from aggregate performance indicators to case-level monitoring. Although average treatment initiation times met institutional targets, the identification of individual patients experiencing excessive delays revealed vulnerabilities that would have remained hidden under traditional reporting approaches. For the case study site, this reinforced the importance of integrating process mining insights into patient navigation activities, enabling earlier intervention in high-risk cases and supporting more proactive coordination across clinical, administrative, and external actors such as health insurers.

5. Conclusions

This study demonstrates the applicability and value of process mining techniques in uncovering and understanding the operational workflows of an outpatient oncology

clinic, as well as identifying delays. By applying the PM² methodology, we successfully mapped real-world treatment and departmental workflows, revealing the most frequent care pathways and identifying significant process variants and deviations.

Our findings show that radiotherapy and medical consultations are the most common daily activities in the oncology clinic analyzed in this case study, with chemotherapy following closely behind. Furthermore, the analysis of treatment pathways for breast, prostate, and digestive organ cancers revealed distinct patterns in care delivery, emphasizing the need for tailored operational strategies based on tumor type.

The application of process mining for performance analysis revealed issues related to data entry in the EHR system, as well as in the healthcare insurance approval process. These issues were examined, and improvement actions were proposed to optimize the patient's journey and increase efficiency, particularly regarding delays in the initiation of treatment.

One particularly important aspect highlighted by this work is the shift from aggregate performance indicators to case-level monitoring. This shift is crucial because, even when managerial reports present favorable aggregated metrics, such as average treatment initiation times, individual patients may still experience excessive delays. Such cases reveal vulnerabilities in care delivery that would remain hidden under traditional, aggregate-based reporting approaches.

Although the present study concentrated primarily on process discovery and performance analysis, it establishes a robust foundation for subsequent conformance checking.

While the case study was conducted within a single clinical setting, we considered that the methodological approach presented here is readily transferable to other institutions. This is largely attributable to the high degree of standardization in ambulatory oncology care, which is guided by internationally recognized guidelines such as those developed by the NCCN (National Comprehensive Cancer Network), ASCO (American Society of Clinical Oncology), and ESMO (European Society for Medical Oncology), among others. Despite minor variations among these protocols, the core activities underpinning ambulatory cancer treatment remain substantially consistent across settings.

Future works should extend the analysis to additional clinical sites, pursue the detection of operational inefficiencies, refine and streamline clinical workflows, and conduct systematic evaluations of adherence to evidence-based guidelines.

The results not only confirm the utility of process mining techniques to enhance the patient journey in an oncology clinic but also highlight its potential as a decision-support tool for healthcare administrators and clinical leaders.

Author Contributions: Conceptualization, R.S.S.; methodology, R.S.S.; software, R.S.S.; validation, J.B.B., M.C. and A.O.I.R.; formal analysis, J.M.P.d.O.; investigation, R.S.S.; resources, R.S.S.; data curation, R.S.S.; writing—original draft preparation, R.S.S.; writing—review and editing, R.S.S. and J.M.P.d.O.; visualization, R.S.S.; supervision, J.M.P.d.O.; project administration, R.S.S. and A.O.I.R.; funding acquisition, R.S.S. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Compumedica Research & Innovation.

Institutional Review Board Statement: This study was approved by the National Research Ethics Committee (CEP/CONEP) under number: CAAE:75225723.7.0000.5497 on 2 December 2023.

Informed Consent Statement: Patient consent was waived because only secondary, anonymized data was used. The study does not involve clinical aspects but is limited to examining the operational flow of the treatment plan. Therefore, only information pertaining to the execution of activities was collected, including the type of activity, date of execution, start and end times, resources utilized, and the roles of those involved.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank the staff at the Mogi das Cruzes Oncology Center who are evaluating the process mining solution.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hanna, T.P.; King, W.D.; Thibodeau, S.; Jalink, M.; Paulin, G.A.; Harvey-Jones, E.; O’Sullivan, D.E.; Booth, C.M.; Sullivan, R.; Aggarwal, A. Mortality due to cancer treatment delay: Systematic review and meta-analysis. *BMJ* **2020**, *371*, m4087. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Ungvari, Z.; Fekete, M.; Buda, A.; Lehoczki, A.; Munkácsy, G.; Scaffidi, P.; Bonaldi, T.; Fekete, J.T.; Bianchini, G.; Varga, P.; et al. Quantifying the impact of treatment delays on breast cancer survival outcomes: A comprehensive meta-analysis. *GeroScience* **2025**, *Epub ahead of printing*. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Williams, F. Assessment of Breast Cancer Treatment Delay Impact on Prognosis and Survival: A Look at the Evidence from Systematic Analysis of the Literature. *J. Cancer Biol. Res.* **2015**, *3*, 1071. [\[PubMed\]](#)
4. Liao, D.Z.; Schlecht, N.F.; Rosenblatt, G.; Kinkhabwala, C.M.; Leonard, J.A.; Ference, R.S.; Prystowsky, M.B.; Ow, T.J.; Schiff, B.A.; Smith, R.V.; et al. Association of Delayed Time to Treatment Initiation with Overall Survival and Recurrence Among Patients With Head and Neck Squamous Cell Carcinoma in an Underserved Urban Population. *JAMA Otolaryngol. Head Neck Surg.* **2019**, *145*, 1001. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Lombardo, M.S.; Popim, R.C. Access of the patient to the cancer network under the ‘Sixty-Day Law’: Integrative Review. *Rev. Bras. Enferm.* **2020**, *73*, e20190406. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Pulido, Z.; Sogame, L.C.M.; Aleixo, S.B. The 60-Days Law: Reality of Timely Treatment in the Analysis of a Series of Colorectal Cancer Cases. *Rev. Bras. Cancerol.* **2023**, *69*, 4. [\[CrossRef\]](#)
7. Marcelino, B.A.; Soares, F.H.; Curi, M.L.D.C.; Wainstein, A.J.A.; Drummond-Lage, A.P. Is a law enough to provide a better prognosis for Brazilian melanoma patients or education is still a gap? *Health Plan. Manag.* **2022**, *37*, 2479–2484. [\[CrossRef\]](#)
8. Chen, M.; Wu, V.S.; Falk, D.; Cheatham, C.; Cullen, J.; Hoehn, R. Patient Navigation in Cancer Treatment: A Systematic Review. *Curr. Oncol. Rep.* **2024**, *26*, 504–537. [\[CrossRef\]](#) [\[PubMed\]](#)
9. van der Aalst, W. *Process Mining: Data Science in Action*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2016. [\[CrossRef\]](#)
10. Wicky, A.; Gatta, R.; Latifyan, S.; De Micheli, R.; Gerard, C.; Pradervand, S.; Michielin, O.; Cuendet, M.A. Interactive process mining of cancer treatment sequences with melanoma real-world data. *Front. Oncol.* **2023**, *13*, 1043683. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Lauritzen, B.B.; Jensen, J.S.; Grønhoj, C.; Wessel, I.; Von Buchwald, C. Impact of delay in diagnosis and treatment-initiation on disease stage and survival in oral cavity cancer: A systematic review. *Acta Oncol.* **2021**, *60*, 1083–1090. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Aversano, L.; Iammarino, M.; Madau, A.; Pirlo, G.; Semeraro, G. Process mining applications in healthcare: A systematic literature review. *PeerJ Comput. Sci.* **2025**, *11*, e2613. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Al-Badarnah, H.; Arif, A. A Systematic Review of Process Mining in Healthcare. In *Proceedings of the 2025 International Conference on New Trends in Computing Sciences (ICTCS)*, Amman, Jordan, 16–18 April 2025; IEEE: New York, NY, USA, 2025; pp. 485–492. [\[CrossRef\]](#)
14. Guzzo, A.; Rullo, A.; Vocaturo, E. Process mining applications in the healthcare domain: A comprehensive review. *WIREs Data Min. Knowl.* **2022**, *12*, e1442. [\[CrossRef\]](#)
15. Kurniati, A.P.; Johnson, O.; Hogg, D.; Hall, G. Process mining in oncology: A literature review. In *Proceedings of the 2016 6th International Conference on Information Communication and Management (ICICM)*, Hatfield, UK, 29–31 October 2016; IEEE: New York, NY, USA, 2016; pp. 291–297. [\[CrossRef\]](#)
16. Garcia, C.D.S.; Meinheim, A.; Junior, E.R.F.; Dallagassa, M.R.; Sato, D.M.V.; Carvalho, D.R.; Santos, E.A.P.; Scalabrin, E.E. Process mining techniques and applications—A systematic mapping study. *Expert Syst. Appl.* **2019**, *133*, 260–295. [\[CrossRef\]](#)
17. Mans, R.S.; Schonenberg, M.H.; Song, M.; Van Der Aalst, W.M.P.; Bakker, P.J.M. Application of Process Mining in Healthcare—A Case Study in a Dutch Hospital. In *Biomedical Engineering Systems and Technologies*; Fred, A., Filipe, J., Gamboa, J., Eds.; Communications in Computer and Information Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 25, pp. 425–438. [\[CrossRef\]](#)
18. Pijnenborg, P.; Verhoeven, R.; Firat, M.; Laarhoven, H.V.; Genga, L. Towards Evidence-Based Analysis of Palliative Treatments for Stomach and Esophageal Cancer Patients: A Process Mining Approach. In *Proceedings of the 2021 3rd International Conference on Process Mining (ICPM)*, Eindhoven, The Netherlands, 31 October 2021–4 November 2021; IEEE: New York, NY, USA, 2021; pp. 136–143. [\[CrossRef\]](#)

19. Savino, M.; Chiloire, G.; Masciocchi, C.; Capocchiano, N.D.; Lenkowicz, J.; Gottardelli, B.; Gambacorta, M.A.; Valentini, V.; Damiani, A. A process mining approach for clinical guidelines compliance: Real-world application in rectal cancer. *Front. Oncol.* **2023**, *13*, 1090076. [[CrossRef](#)] [[PubMed](#)]
20. Baker, K.; Dunwoodie, E.; Jones, R.G.; Newsham, A.; Johnson, O.; Price, C.P.; Wolstenholme, J.; Leal, J.; McGinley, P.; Twelves, C.; et al. Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. *Int. J. Med. Inform.* **2017**, *103*, 32–41. [[CrossRef](#)] [[PubMed](#)]
21. Iachecen, F.; Popadiuk, G.S.; Batista, B.R.; Dallagassa, M.R.; Carvalho, D.R.; Portela Santo, E.A.; Ioshii, S.O. Process mining applied to the patient journey: A case study in lung cancer. *RBSS—Rev. Bras. Saúde Supl.* **2023**, *1*, 1.
22. Van Eck, M.L.; Lu, X.; Leemans, S.J.J.; Van Der Aalst, W.M.P. PM2: A Process Mining Project Methodology. In *Advanced Information Systems Engineering*; Zdravkovic, J., Kirikova, M., Johannesson, P., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9097, pp. 297–313. [[CrossRef](#)]
23. Johnson, O.A.; Dhafari, T.B.; Kurniati, A.; Fox, F.; Rojas, E. The ClearPath method for care pathway process mining and simulation. In *International Conference on Business Process Management*; Springer International Publishing: Cham, Switzerland, 2018; pp. 239–250.
24. Muñoz-Gama, J.; Martin, N.; Fernandez-Llatas, C.; Johnson, O.A.; Sepúlveda, M.; Helm, E.; Galvez-Yanjari, V.; Rojas, E.; Martinez-Millana, A.; Aloini, D.; et al. Process mining for healthcare: Characteristics and challenges. *J. Biomed. Inform.* **2022**, *127*, 103994. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.