

Article

Evaluating Conformance and Performance in Oncology Pathways via Telemedicine Using Process Mining: A Study in a Chilean Hospital Network

Matías Cornejo ¹, Esteban Chiu ^{2,*}, Daniel Capurro ^{3,4} , Sebastián Valderrama ^{4,5} , Sebastián Mondaca ^{5,6},
Tomas Merino ^{5,6}, Steffen Härtel ^{1,7,8}  and Eric Rojas ^{2,9,*} 

- ¹ Centro de Informática Médica y Telemedicina CIMT, Universidad de Chile, Santiago 8820808, Chile; matias.cornejo.t@ug.uchile.cl (M.C.); shartel@uchile.cl (S.H.)
- ² Institute for Biological and Medical Engineering, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile
- ³ School of Computing and Information Systems, University of Melbourne, Melbourne, VIC 3010, Australia; dcapurro@unimelb.edu.au
- ⁴ Department of Internal Medicine, School of Medicine, Pontificia Universidad Católica de Chile, Santiago 8331150, Chile
- ⁵ Cancer Institute, Red de Salud UC CHRISTUS, Santiago 8330033, Chile
- ⁶ Department of Hematology and Oncology, School of Medicine, Pontificia Universidad Católica de Chile, Santiago 8331150, Chile
- ⁷ Laboratory for Scientific Image Analysis SCIAN-Lab, Institute of Biomedical Sciences ICBM, School of Medicine, Universidad de Chile, Santiago 8380453, Chile
- ⁸ Center of Applied Mathematics, CMM, School of Engineering, Universidad de Chile, Santiago 8370456, Chile
- ⁹ Department of Clinical Laboratories, School of Medicine, Pontificia Universidad Católica de Chile, Santiago 8331150, Chile
- * Correspondence: echiu@uc.cl (E.C.); eric.rojas@uc.cl (E.R.)

Featured Application

The proposed process mining workflow can support healthcare organizations in auditing tele-oncology pathways, quantifying conformance to guideline-derived process models, detecting bottlenecks in patient journeys, and producing reproducible metrics for digital health governance and continuous process improvement.

Abstract

Tele-oncology expanded rapidly during the COVID-19 pandemic and remains central to cancer care, yet evidence on guideline adherence in tele-oncology is limited, particularly in Latin America. In Chile, the Explicit Health Guarantees (GES) program mandates oncology care standards but provides no process model for monitoring adherence. We evaluated the conformance and temporal performance of telemedicine-based oncology pathways against a guideline-derived reference model using process mining, based on event logs from 182 patients treated in 2020–2023 in a Chilean hospital network. Conformance was computed using PM4Py with token-based replay and alignment-based fitness, complemented by cohort, abstraction-level and robustness analyses with bootstrap confidence intervals and multiple-comparison-adjusted testing. Overall conformance was low (4.95% of cases fully conforming; mean alignment fitness 0.19, 95% CI 0.16–0.22) against a highly precise reference model (precision 0.90). Censoring-aware analyses showed that low conformance was driven predominantly by genuine structural deviations rather than incomplete observation alone. Event-level cohort signatures were observed, but patient-level analyses suggested that these patterns should be interpreted as exploratory. Crude differences in care duration by out-of-model activity status were largely explained by differential observation time after exposure adjustment. Process-mining conformance checking provides



Academic Editors: Julius Nganji and Kingsley Okoye

Received: 19 June 2026

Revised: 24 July 2026

Accepted: 30 July 2026

Published: 3 August 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.

reproducible indicators for governing tele-oncology pathways when metrics are interpreted alongside methodological caveats.

Keywords: process mining; conformance checking; event logs; tele-oncology; clinical pathways; healthcare process mining; oncology care; performance analysis

1. Introduction

Cancer is one of the leading causes of morbidity and mortality worldwide and represents an increasing challenge for healthcare systems due to the complexity of care and the need to coordinate multiple diagnostic, therapeutic, and follow-up services [1,2]. In Chile, cancer incidence has shown a sustained annual increase, with 59,876 new diagnoses recorded in 2022 and an incidence rate of 188.7 per 100,000 inhabitants. This growing burden has motivated the implementation of public strategies aimed at strengthening continuity of care, improving efficiency, and ensuring conformance to established standards throughout the care process [1].

Oncological care requires close coordination among clinical specialties, strict adherence to standardized protocols, and longitudinal monitoring of patients. In this context, the analysis of clinical pathways has become a key tool for evidence-based healthcare management, as it enables the evaluation of how care processes are executed in relation to expected models [3,4]. This need became particularly evident during the COVID-19 pandemic, when telemedicine was rapidly adopted as an essential mechanism to ensure continuity of care for patients with chronic and oncological conditions [5]. Telemedicine facilitated the maintenance of regular appointments while reducing exposure risks among immuno-compromised patients, improving satisfaction among both users and healthcare teams [6,7]. However, its accelerated implementation also introduced challenges related to interoperability, geographical equity, and the standardization of clinical records across care modalities [1].

In Chile, the Explicit Health Guarantees program (Garantías Explícitas en Salud, GES) regulates the care of 17 oncological pathologies through Clinical Practice Guidelines that define services, maximum waiting times, and follow-up requirements [8]. Despite their regulatory role, these guidelines do not provide a detailed description of the operational flows that structure the care process. This lack of explicit process modeling complicates the assessment of whether patients' real trajectories, both in-person and via telemedicine, align with the expected clinical pathway, thereby limiting the evaluation of adherence to ministerial recommendations and hindering institutional traceability of care [9,10].

Within this framework, process mining (PM) emerges as a particularly relevant methodological approach. PM is a data science discipline that enables the discovery, analysis, and improvement of real-world processes through the examination of event logs generated by information systems [11]. In healthcare, PM has been widely applied to reconstruct clinical pathways based on actual patient data, revealing how care is delivered in practice and allowing comparisons with formal or normative reference models [4,12]. This methodology facilitates the analysis of the sequence, frequency, and duration of clinical activities, as well as the identification of variability patterns and deviations from established medical protocols or guidelines [13]. Recent systematic reviews confirm the expansion of process mining applications in healthcare, while highlighting persistent challenges related to data quality, process heterogeneity, and clinical interpretability [14].

Process mining integrates three complementary perspectives: process discovery, conformance checking, and enhancement [15]. Together, these perspectives provide a comprehensive view of clinical and organizational behavior, strengthening traceability, and

supporting data-driven decision-making. In the Chilean healthcare system, and particularly within the context of the GES program, PM is especially relevant due to the existence of mandatory deadlines and standardized benefits that require systematic monitoring [1,8].

Conformance checking focuses on comparing observed care processes with a predefined reference model to quantify their degree of alignment [16,17]. This approach enables the identification of structural deviations between planned and executed care, as well as their classification according to potential clinical or organizational impact [18]. In healthcare settings, conformance checking has proven effective for validating medical protocols and assessing conformance to clinical guidelines, revealing discrepancies between normative models and real practice [19,20]. These comparisons are commonly performed using algorithms such as token-based replay, which simulate the execution of the reference model and detect deviations when observed event sequences diverge from the expected flow [15,21]. Platforms such as Celonis operationalize this approach by enabling the comparison of real event logs with BPMN reference models through token-based replay analyses [16,17]. Recent work in medical process conformance has further emphasized the relevance of comparing observed clinical behavior with normative process representations to support guideline monitoring and process improvement [22].

In this study, “conformance” refers to the alignment between observed event log behavior and the guideline-derived BPMN reference model, whereas “adherence” refers to the broader clinical interpretation of whether care follows guideline recommendations.

In oncology, conformance checking is particularly useful for assessing adherence to the clinical pathways defined by GES guidelines, as it allows the objective correlation of clinical activities with prescribed care sequences and timelines [8]. Performance analysis further extends this evaluation by incorporating the temporal dimension of care delivery, focusing on the identification of delays, interruptions, and variability in timeliness [15,23,24]. This perspective supports the detection of bottlenecks and the assessment of the healthcare system’s capacity to meet the deadlines established by ministerial programs [8].

In the oncological context, the time intervals between diagnosis, treatment initiation, and follow-up are critical determinants of care quality. Integrating performance analysis with conformance checking enables a more comprehensive assessment of clinical pathways by jointly addressing their structural and temporal dimensions [25]. From this perspective, process mining provides a robust methodological framework for evaluating the real-world implementation of oncological care pathways and for identifying opportunities to improve adherence, efficiency, and quality within complex healthcare systems.

Although conformance checking has been previously applied to oncology pathways (for instance in rectal cancer [19]) and a recent scoping review has mapped its use for guideline adherence more broadly [18], evidence on its application to telemedicine-based oncology care remains scarce, and no prior study has examined tele-oncology pathways within a Latin American regulatory framework such as the Chilean GES program. To our knowledge, this is the first conformance-checking study of telemedicine-based oncology pathways under a mandatory Latin-American guideline framework (the Chilean GES program); its contribution is threefold: (i) it provides empirical evidence on the conformance and temporal performance of tele-oncology pathways under a mandatory national guideline framework; (ii) it demonstrates, at the transition level, how process mining can localize delays that aggregate metrics obscure; and (iii) it illustrates a reproducible conformance-checking workflow that health networks can embed as a governance instrument for digital oncology care.

Beyond evaluating tele-oncology pathways, this study investigates how different conformance metrics and modeling choices influence the interpretation of guideline adherence in highly heterogeneous and right-censored healthcare processes.

This study pursued two objectives. First, to assess the conformance of real telemedicine-based oncology pathways compared to an expected care model, derived from the Clinical Practice Guidelines of the Chilean Ministry of Health and validated by clinical experts from the Red de Salud UC CHRISTUS health network. Second, to perform a time-based performance analysis of the care process to identify relevant variations and generate useful evidence for the evaluation and continuous improvement of efficiency and quality in digital cancer care.

2. Materials and Methods

2.1. Study Environment

The study was conducted at the Red de Salud UC CHRISTUS health network, a private Chilean healthcare institution that provides high-complexity care and integrates hospital, outpatient, and telemedicine services. The network operates a Hospital Information System (HIS) that records healthcare activities in a structured manner, enabling the reconstruction of comprehensive clinical pathways from the recorded data. The analysis focused on patients with cancer treated via telemedicine between 1 June 2020 and 1 June 2023, the period during which telemedicine was incorporated into the follow-up of GES-covered patients with cancer in response to the COVID-19 pandemic. The dataset included 182 patients with various cancer diagnoses who met eligibility criteria during the study period. Data were extracted by the hospital's Information Management Division without involvement of the research team. Eligible patients were those with (i) a confirmed oncological diagnosis, (ii) at least one oncology-related activity recorded in the HIS, and (iii) at least one episode of care delivered through the institutional telemedicine network during the study period. No additional exclusion criteria were applied, and no a priori sample-size calculation was performed, as the study was exploratory and constrained by the volume of tele-oncology records available for the period.

2.2. Ethical Considerations

This study was approved by the Scientific Ethics Committee for Health Sciences of the Pontificia Universidad Católica de Chile (CEC-UC), under resolution ID220310004 dated 9 March 2023, and subsequent updates. Authorization for the use of anonymized clinical data was granted by the Research Directorate of the School of Medicine (DIEMUC). As no patient contact or intervention was required, the study was classified as minimal risk and informed consent was waived.

2.3. Observed Process and Reference Care Model

Process discovery applied to the preprocessed event log yielded a highly heterogeneous, weakly structured (“spaghetti”) as-is model (Figure 1), reflecting the wide variability of the observed care pathways. This heterogeneity motivates the use of a normative reference model against which adherence can be assessed, rather than relying on the discovered model alone.

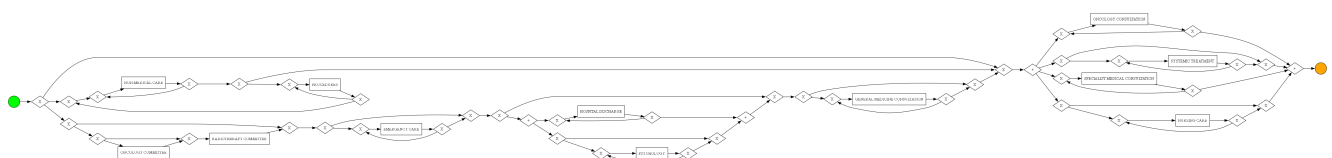


Figure 1. Discovered as-is process model at the category level, obtained by applying process discovery to the event log. The green node denotes the process start and the orange node the process end; rectangles are activities and white diamonds are routing (silent) transitions. The unstructured, highly connected topology illustrates the heterogeneity of the observed care pathways and motivates the guideline-derived reference model used for conformance checking.

A reference care model was developed in Business Process Model and Notation (BPMN) from the activities and waiting-time milestones explicitly represented in the GES Clinical Practice Guidelines of the Chilean Ministry of Health and refined through interviews with nurses and oncologists from the Red de Salud UC CHRISTUS. The model was designed as a guideline-based operational representation of the care pathway rather than an exhaustive representation of every activity that may occur in real-world oncology practice. It comprised the phases of diagnosis and staging, oncology committee, therapeutic decision, treatment, and check-up and follow-up (Figure 2).

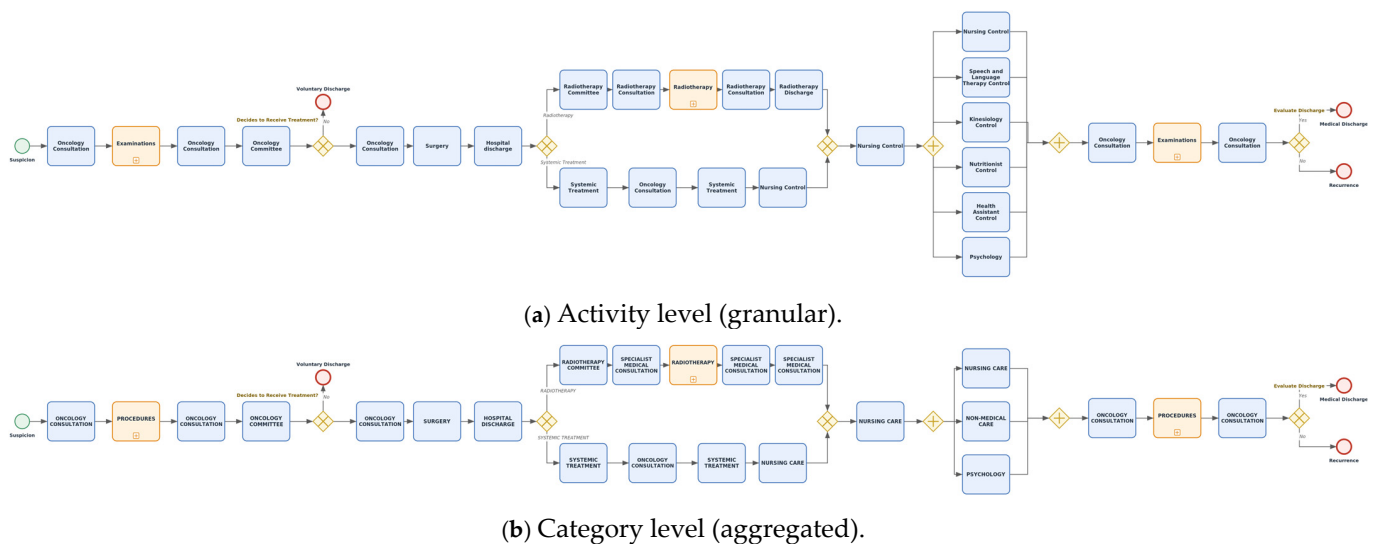


Figure 2. Guideline-derived BPMN reference care model at the two abstraction levels analyzed: (a) the granular activity level and (b) the aggregated category level. Both levels share the same control-flow structure; the category level groups granular activities into broader classes (for example, the non-medical support controls are collapsed into Non-Medical Care, and the specialist consultation variants into Specialist Medical Consultation).

The model was validated qualitatively through individual, semi-structured interviews with four oncology specialists who directly treat these patients. Each interview followed three steps: (i) the expert described the oncological care process; (ii) the expert reviewed the guideline-derived BPMN model iteration by iteration and confirmed or corrected its structure; and (iii) when analyzing the observed variability, the expert indicated which deviations corresponded to genuine clinical variation and which reflected legitimate adjustments of the process. Discrepancies between experts were reconciled by consensus and incorporated into successive model iterations, and pathology-specific differences were accommodated by retaining, at the category level, only the activities common to the modeled GES pathways. We did not conduct a formal Delphi process or compute a quantitative inter-rater agreement coefficient; the validation was an iterative qualitative consensus, consistent with common practice in healthcare process mining, and this is acknowledged as a limitation. Two abstraction levels of the model were considered: a granular activity level (14 modeled labels) and an aggregated category level (10 modeled labels).

Figure 3 presents the overall methodological workflow of the study, which was carried out in the following sequential steps. First, clinical data were extracted from the Hospital Information System and transformed into a standardized event log through expert-guided activity mapping and normalization (Section 2.4). Second, and in parallel, a guideline-derived BPMN reference care model was developed and validated by oncology specialists (Section 2.3). Third, the event log was compared against the reference model through conformance checking, applying token-based replay and alignment-based techniques (Section 2.5). Fourth, conformance

results were statistically characterized using bootstrap confidence intervals and cohort-level comparisons (Section 2.6). Fifth, a time-based performance analysis was conducted to identify bottlenecks and assess adherence to the GES deadlines (Section 2.7). Across these steps, Celonis EMS v3.1 was used for process exploration, visualization and performance analysis, while PM4Py v2.7.0 was used to obtain reproducible conformance metrics and perform robustness analyses; the outputs from both analytical streams were then integrated to evaluate pathway conformance and identify bottlenecks.

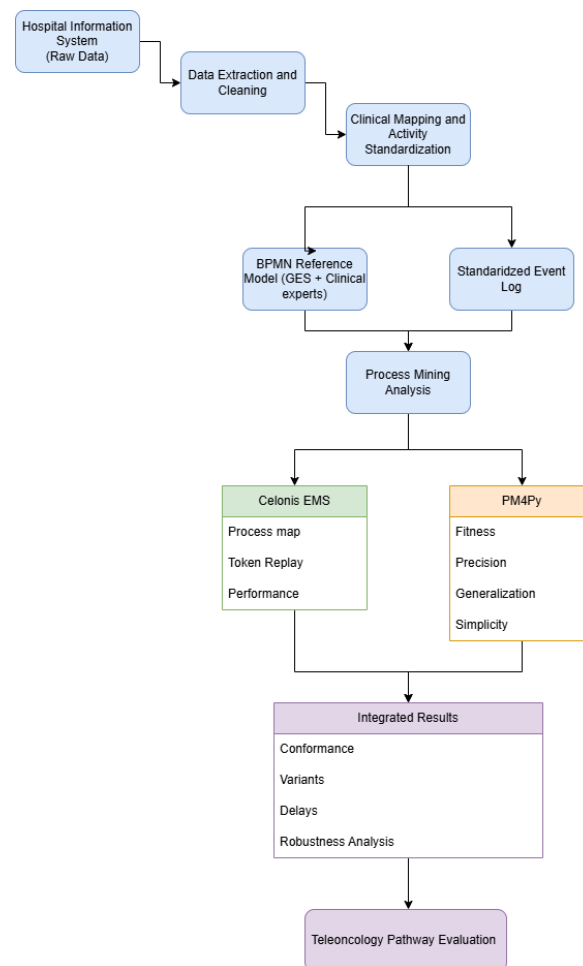


Figure 3. Methodological workflow of the proposed process mining framework. Clinical and telemedicine records were extracted from the Hospital Information System and transformed into a standardized event log through expert-guided activity mapping. A BPMN reference model derived from national clinical guidelines and validated by oncology specialists was analyzed in parallel with the event log. Celonis EMS supported process exploration and performance analysis, whereas PM4Py provided reproducible conformance metrics and robustness assessments. The resulting indicators were used to evaluate pathway adherence, identify bottlenecks and support tele-oncology governance.

2.4. Data Preparation and Activity Standardization

An event log was created from the clinical information system. Each record included case ID (patient identifier), activity name, timestamp, and resource (executing practitioner). The curated log comprised 182 cases, 8296 events and 172 distinct trace variants. Because of terminology differences between the activities recorded in clinical systems and those defined in the ministerial guidelines, a mapping and standardization process was performed with clinical experts, considering the profession and specialty of the personnel, the location of the activity, and the specific actions performed in each event (Appendix A). Event log preprocessing is a critical step in process mining studies, particularly when raw

information system data must be transformed into analyzable case identifiers, activity labels, timestamps, and resources [26]. Recent work on guideline adherence supported by process mining highlights the need for explicit clinical activity coding and standardized representation when conformance checking is applied to healthcare processes [24].

The delivery modality (in-person vs. telemedicine) was treated as a mode of care rather than a distinct pathway activity, so telemedicine variants were collapsed onto their base activity to align the log vocabulary with the reference model. After standardization, 97% of category-level events were covered by the reference model. Patients were assigned to cohorts by their principal ICD-10 diagnosis (colon C18, rectum C20, lung C34, pancreas C25, breast C50, and others).

The preprocessing pipeline comprised the following reproducible steps: (1) timestamps were parsed to a common UTC datetime and events ordered chronologically within each case; (2) free-text activity labels were normalized (accent and case folding, whitespace trimming); (3) ICD-10 diagnostic codes were cleaned (removal of punctuation and de-concatenation of multiple codes) to assign each patient a principal diagnosis; (4) the resulting vocabulary was mapped to the reference-model labels at two abstraction levels (activity and category) using expert-defined rules based on the practitioner's profession and specialty, the location of the activity and the action performed; (5) each case was assigned to a tumor cohort by its principal ICD-10 code and to a calendar quarter by its first event; and (6) the standardized log was validated against both BPMN models. After standardization, 97% of category-level events were covered by the reference model. A public reproducibility package, including pseudocode, a synthetic event log, and an executable reference implementation, is provided to support methodological reproducibility (Appendix C).

Because this mapping relies on expert judgment, it may introduce a degree of subjectivity. To mitigate this, standardization followed explicit, documented rules rather than case-by-case decisions, and ambiguous labels were resolved consistently: for example, the multiple specialty-consultation labels recorded in the HIS (e.g., gastroenterology, coloproctology or surgical-oncology consultations) were mapped to a single Specialist Medical Consultation category; telemedicine variants of an activity were collapsed onto their in-person base activity; and duplicated or malformed labels (e.g., "Psychology Psychology") were normalized to their canonical form. Residual subjectivity in this mapping is acknowledged as a limitation.

2.5. Conformance Checking

Conformance and performance were analyzed using process mining. Initial process exploration and visualization were carried out on the Celonis EMS v3.1 platform, which integrates the institutional event log with the BPMN reference model. To obtain reproducible quantitative metrics and to enable systematic comparisons across cohorts, abstraction levels and cohort-definition rules, the analysis was complemented with the open-source PM4Py library. PM4Py has been increasingly used in healthcare process mining studies to support reproducible analyses based on multiple mining and conformance techniques [27]. Two conformance techniques were applied. Token-based replay yielded a case-level binary indicator: a case was classified as conforming when its trajectory could be replayed within the reference model without missing or remaining tokens. Under this definition, conformance was assessed as a strict end-to-end criterion, meaning that any unresolved token discrepancy was sufficient to classify a trace as non-conforming. To avoid penalizing clinically expected follow-up activities not explicitly represented in the reference model, oncology consultations and oncology-related nursing activities were incorporated through the activity standardization process and included in an allowlist before replay. Cases with out-of-model activities were analyzed separately from fully conforming cases, because the

presence of an out-of-model activity captures a specific type of deviation and does not necessarily represent the full binary conformance status of the trace. Activities belonging to other clinical specialties that were not represented in the BPMN model were not allowlisted and were therefore treated as deviations. This decision was implemented following input from clinical experts from the same healthcare network. This approach was intended to preserve the guideline-based scope of the model while accommodating routine oncology follow-up practices.

Alignment-based conformance provided a continuous log-fitness measure that is more robust to concurrency and complex deviations [17]. Model quality was additionally characterized by precision (the proportion of model-allowed behavior that is actually observed), generalization, and simplicity. To test the robustness of the cohort definition, results were recomputed under an alternative rule based on the presence of an exact diagnostic code.

2.6. Statistical Analysis

Uncertainty in fitness and in the percentage of conforming cases was quantified with 95% bootstrap confidence intervals (1000 resamples). For each cohort, the distribution of activities was compared against the overall population using a chi-square test, with standardized residuals (z-scores, $|z| \geq 1.96$ considered significant) identifying over- and under-represented activities and Cramer's V quantifying effect size. *p*-values were corrected for multiple comparisons using the Benjamini–Hochberg false discovery rate (FDR) [28]. Temporal stability was assessed by computing fitness per calendar quarter.

2.7. Performance Analysis

Time-based performance was analyzed using directly follows graphs annotated with mean transition times, and the slowest transitions per cohort were ranked to identify bottlenecks. Where available, mean transition times were compared against the maximum time limits established by the GES Clinical Practice Guidelines. As pathway timings are pathology-specific, the time-based analysis was performed separately for the colon and breast cancer cohorts.

3. Results

A total of 182 patient cases (8296 events) were included. The population was highly heterogeneous: 172 distinct activity variants were observed, and the most common variant accounted for only 4.4% of cases. This indicates a high degree of heterogeneity in tele-oncology care pathways. Overall pathway conformance was low and concordant across techniques: 4.95% of cases (9/182) were fully conforming by token-based replay, with a mean alignment-based fitness of 0.19 (95% CI 0.16–0.22). The reference model was highly precise at both abstraction levels (precision 0.90 at the category level and 0.98 at the activity level), indicating that it permits very little behavior beyond what is actually observed. Generalization (0.52) and simplicity (0.81) were moderate (Table 1).

Table 1. Overall conformance and model-quality metrics (category level).

Metrics	Value [95% CI]
Conforming cases (token-based replay)	4.95% (9/182)
Mean alignment fitness	0.19 [0.16–0.22]
Precision (category/activity level)	0.90/0.98
Generalization	0.52
Simplicity	0.81
Distinct trace variants	172

Comparison across abstraction levels. The two abstraction levels of the reference model yielded consistent conformance conclusions. At the aggregated category level (10 modeled labels), 4.95% of cases were fully conforming by token-based replay, with a mean alignment fitness of 0.19 (95% CI 0.16–0.22). At the granular activity level (14 modeled labels), conformance was marginally higher (6.6% of cases fully conforming) with a comparable mean alignment fitness of 0.17 (95% CI 0.13–0.20). Precision was high at both levels (0.90 at the category level and 0.98 at the activity level), generalization was nearly identical (0.52 vs. 0.54 at the category vs. activity level), and simplicity was comparable (0.81 vs. 0.77). The high precision at both abstraction levels confirms that neither model is over-permissive: in both cases the reference admits little behavior beyond what is actually observed, so the low fitness reflects genuine deviations and right-censoring rather than a loose specification. Because the conformance findings were stable across abstraction levels, the category level was retained for the primary cohort and performance analyses, as it offers the most interpretable mapping to the guideline activities.

Conformance was uniformly low across cohorts, with no cohort markedly exceeding the overall population (Table 2; Figure 4). This is consistent with the heterogeneity observed within cohorts (e.g., colon: 33 distinct variants among 34 cases). The breast cohort reached 0% binary conformance yet a continuous alignment fitness (0.16) comparable to that of cohorts that did record conforming cases (e.g., colon, 0.17; overall, 0.19), indicating that the 0% reflects right-censored, incomplete trajectories rather than the least adherent care.

Table 2. Conformance and model-quality metrics by cohort at the category level.

Cohort	<i>n</i>	Conforming Cases (Token-Based Replay)	Fitness [95% CI]	Deviations vs. General (FDR)
General	182	4.95%	0.19 [0.16–0.22]	-
Colon	34	5.88%	0.17 [0.10–0.25]	Significant ($p < 0.001$)
Breast	31	0.0%	0.16 [0.12–0.21]	Significant ($p < 0.001$)
Others	86	8.14%	0.23 [0.17–0.29]	Significant ($p < 0.001$)
Lung (suppl.)	9	0.0%	0.19 [0.10–0.33]	Significant ($p < 0.001$)
Pancreas (suppl.)	11	0.0%	0.07 [0.05–0.09]	Significant ($p < 0.001$)
Rectum (suppl.)	11	0.0%	0.14 [0.07–0.25]	Significant ($p < 0.001$)

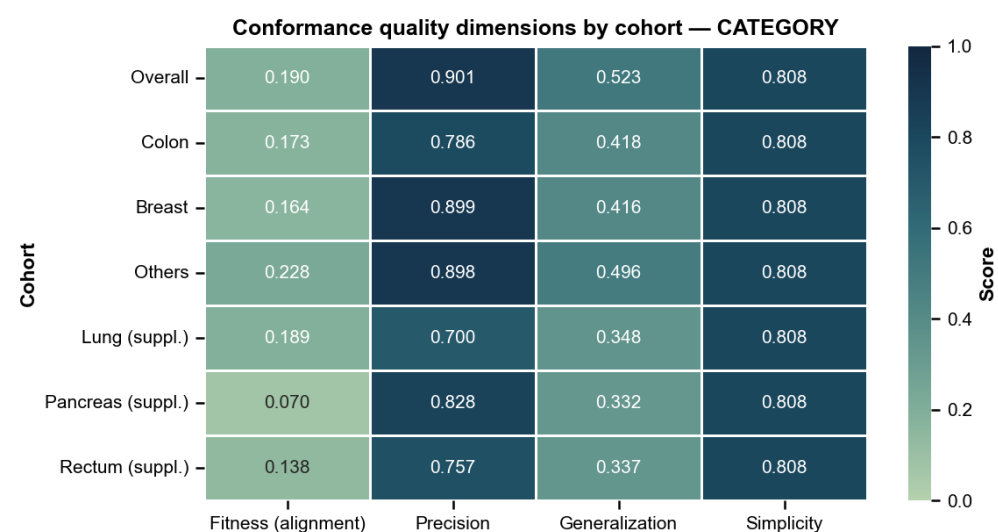


Figure 4. Conformance and model-quality dimensions by cohort.

This low conformance is explained partly by right-censoring inherent to the study window and partly by genuine structural deviations from the reference model. Only 2 of 182 cases (1%) ended in a discharge event, and only ~40% of cases began at the model's

entry activity. Because the 2020–2023 window truncates trajectories that are still ongoing, strict replay classifies most incomplete trajectories as non-conforming. Binary conformance and continuous fitness should therefore be read together: binary conformance reflects whether a full guideline trajectory was observed within the window, while alignment fitness captures the degree of partial adherence.

These findings are consistent with prior work reporting low adherence to clinical guidelines in oncology, including breast, colon and cervical cancer, where standardized guidelines often fail to anticipate complex patients, comorbidities, or disease progression [29,30]. In our cohort (Figure 5), cohort-specific signatures were apparent at the event level (e.g., over-representation of systemic treatment and nursing care in colon cancer, and of non-medical care in breast cancer). However, when the analysis was repeated at the patient level—aggregating each patient’s activity profile to avoid the pseudo-replication that arises from treating repeated events as independent—the differences in activity presence across cohorts were no longer statistically significant after false-discovery-rate correction (Cramér’s V 0.08–0.15; all $q > 0.35$; Figure A1b). These cohort signatures should therefore be read as descriptive and exploratory rather than as confirmed cohort effects. Some deviations are nonetheless consistent with clinically justified adaptations and resource availability rather than failures of care, in line with evidence that remote-care teams adjust the frequency and ordering of activities according to clinical evolution [31].

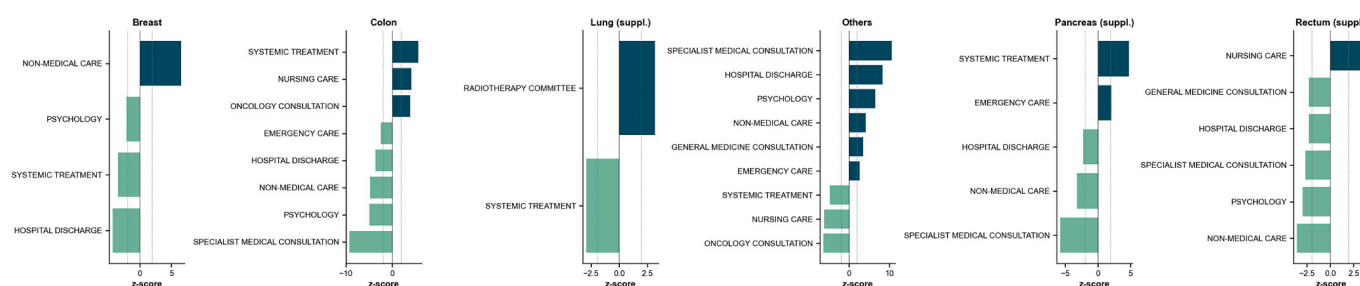


Figure 5. Most over- and under-represented activities per cohort (standardized residuals). Dark (navy) bars denote over-represented activities (positive standardized residual) and green bars under-represented activities (negative residual); dashed vertical lines mark the ± 1.96 significance threshold.

Cases containing at least one out-of-model activity showed a longer observed care duration than cases without such activities (crude, unadjusted comparison): the mean total duration was 680.9 days (median 808.2; $n = 87$) among cases with at least one out-of-model activity, compared with 399.6 days (median 296.1; $n = 95$) among cases without them. This crude difference of approximately 281 days is confounded by observation time: as shown in Appendix B, the exposure-adjusted rate of out-of-model events per 100 patient-days does not increase with the observation window, so the excess reflects longer follow-up rather than more deviation per unit of time. The crude comparison is therefore reported for descriptive completeness only and should not be read as evidence that out-of-model activity prolongs care. Figure 6 shows the corresponding duration distribution (crude, unadjusted distribution; see Appendix B for the exposure-adjusted analysis that removes this association).

Beyond structural conformance, the time-based analysis of directly follows graphs localized delays to specific transitions rather than across the whole pathway. The slowest transitions involved telemedicine-mediated specialist and general-medicine interactions: in the overall cohort the specialist-to-general-medicine transition had a median of 10.6 days [IQR 5.6–30.2], and specialist-to-psychology transition had a median of 6.9 days [IQR 0.8–9.9]. In the pancreas cohort, the emergency-care-to-oncology transition had a median of 10.4 days [IQR 2.0–14.2] (Table 3). For colon cancer (Figure 7), delays were concen-

trated in repeated specialist consultations; for breast cancer (Figure 8), transitions were comparatively shorter and better aligned with guideline timeframes. Overall, delays were concentrated in telemedicine-based interactions rather than uniformly distributed across the pathway.

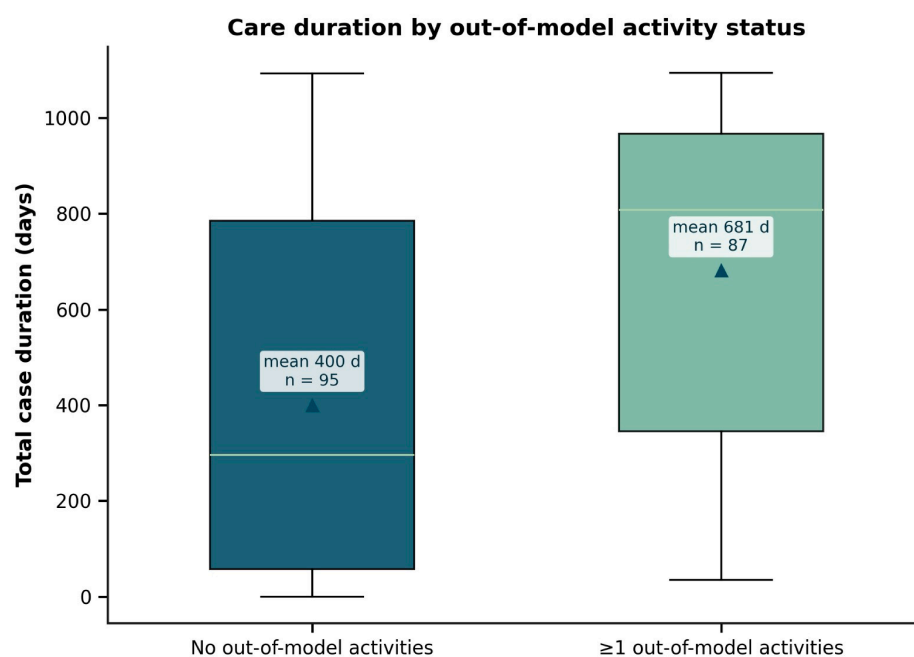


Figure 6. Total care duration according to the presence or absence of out-of-model activities.

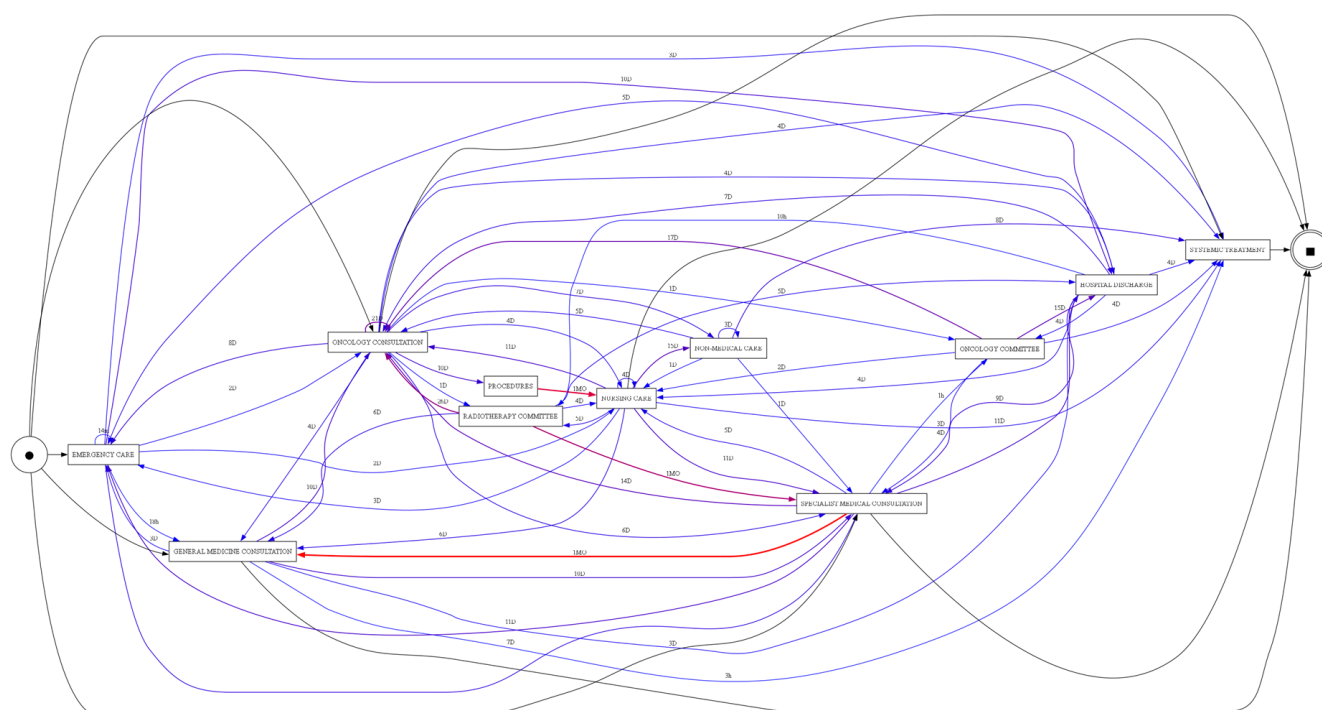


Figure 7. Directly follows graph (mean transition times) for the colon cancer cohort. Edge color encodes the magnitude of the mean transition time (blue = shorter, purple-to-red = longer); black edges connect the artificial start (circle) and end (square) nodes. Labels give mean transition times (m = minutes, h = hours, D = days, MO = month).

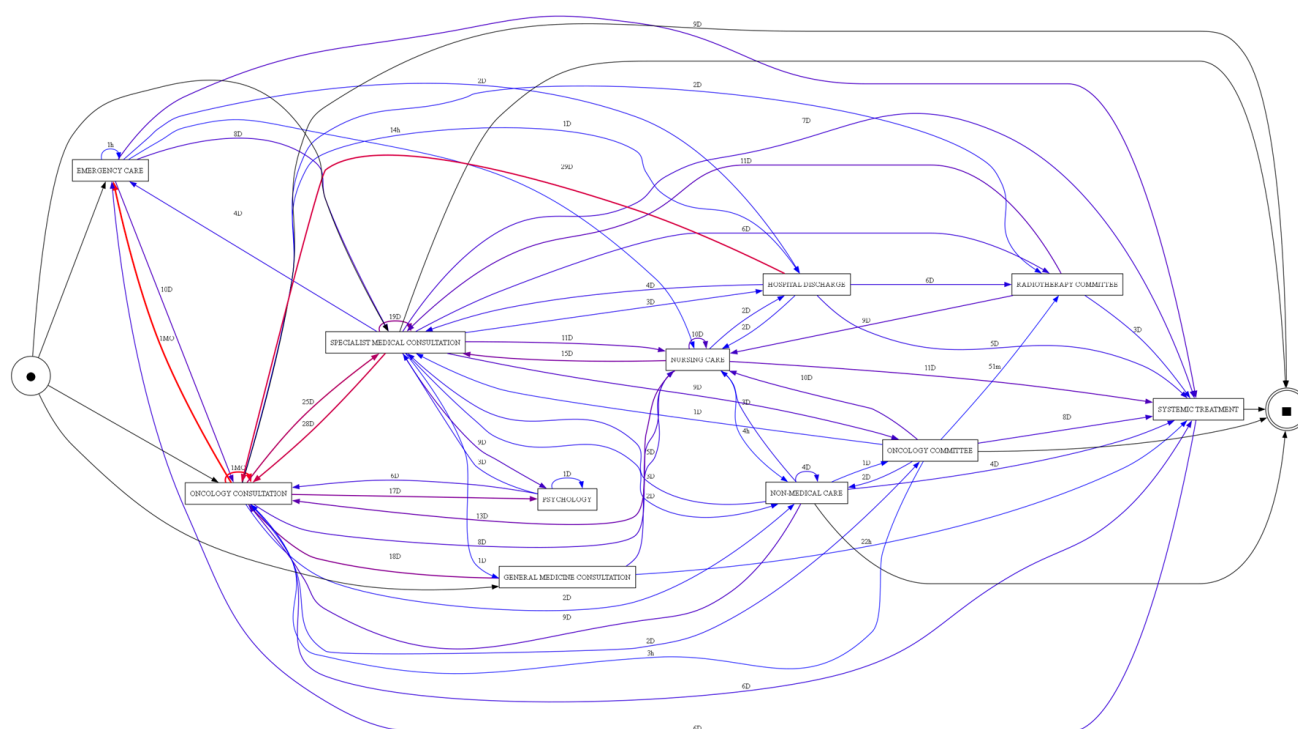


Figure 8. Directly follows graph (mean transition times) for the breast cancer cohort. Edge color encodes the magnitude of the mean transition time (blue = shorter, purple-to-red = longer); black edges connect the artificial start (circle) and end (square) nodes. Labels give mean transition times (m = minutes, h = hours, D = days, MO = month).

Table 3. Slowest transitions (selected) by cohort, median transition time [IQR], in days.

Cohort	Transition	n	Median Days [IQR]
General	Specialist -> General Medicine	20	10.6 [5.6–30.2]
General	Specialist -> Psychology	14	6.9 [0.8–9.9]
General	General Medicine -> Specialist	21	10.8 [6.0–24.5]
Colon	Specialist -> Specialist	91	6.9 [1.7–25.5]
Breast	Oncology -> Emergency	8	12.6 [4.9–55.7]
Pancreas	Emergency -> Oncology	8	10.4 [2.0–14.2]

To assess whether the low conformance was an artifact of right-censoring, the token-based replay was complemented with prefix-alignments and a decomposition of the alignment cost. Because the observation window truncates trajectories that were still in progress at the extraction date, a strict replay can misclassify not-yet-observed steps as deviations; prefix-aware alignment guards against this by scoring only the portion of each trajectory that had actually been reached. Allowing incomplete trajectories raised conformance only marginally (from 4.40% to 4.95%), and the cost decomposition attributed only about 0.1% of the total misalignment to unreached (censored) steps, with the remaining 99.9% corresponding to genuine deviations. A sensitivity analysis that varied the definition of the censored tail confirmed that genuine deviation remained dominant (at least 95% of the misalignment cost) and that no case became conforming under any tail definition, indicating that the low conformance is a real property of the observed pathways rather than an artifact of the observation window (Figure 9).

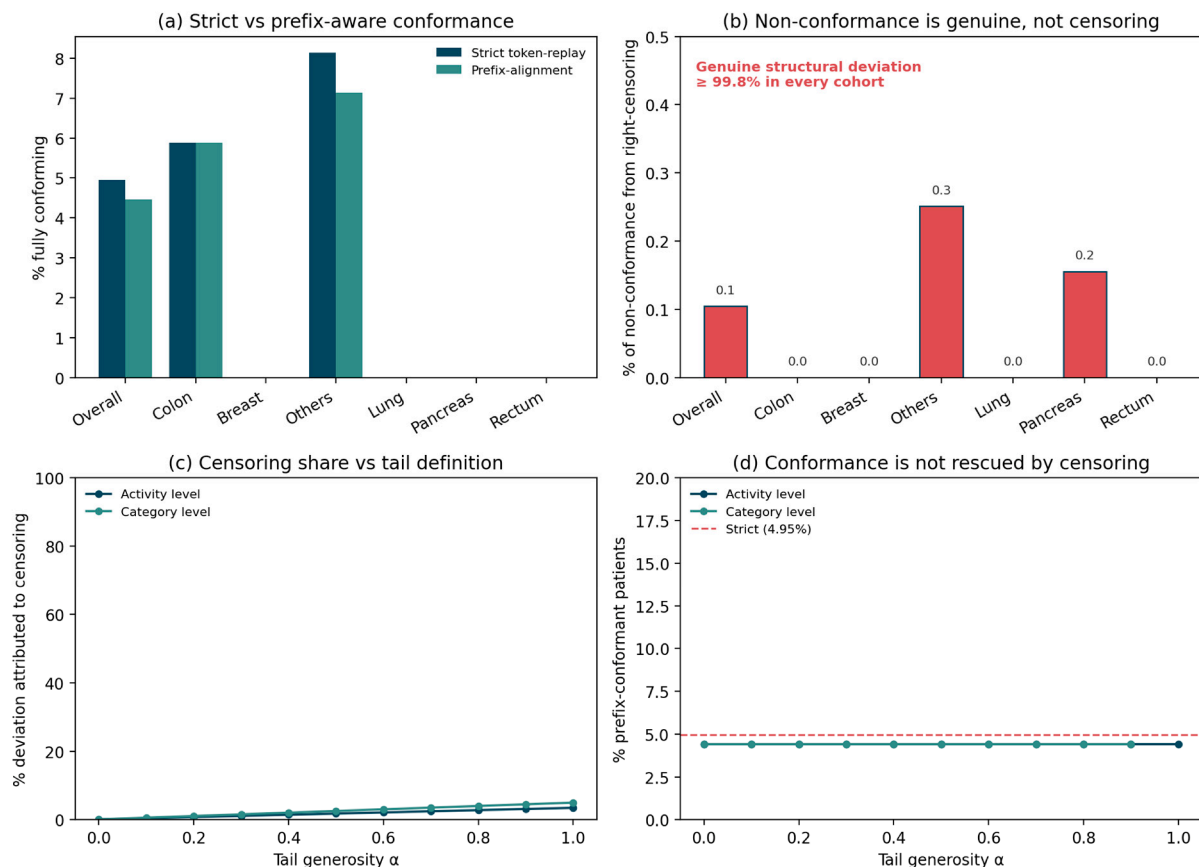


Figure 9. Censoring-aware conformance analysis: strict token-based replay versus prefix-aware conformance by cohort (a); decomposition of non-conformance into genuine deviation and right-censoring (b); and a sensitivity analysis showing that the censoring share (c) and the proportion of prefix-conformant patients (d) are robust to the definition of the censored tail.

Having established that the low conformance was not an artifact of the observation window, we further examined patient- and pathway-level predictors of adherence, the exposure-adjusted performance, and delivery-modality effects. These confirmatory statistical analyses are reported in Appendix B (Figures A1 and A2).

Conformance results were robust to the cohort definition rule: the difference in fitness between the principal-diagnosis rule and the exact-code rule was negligible in all cohorts ($|\Delta \text{fitness}| < 0.012$). The negligible differences observed between cohort-definition rules indicate that the conformance findings are robust to diagnostic assignment strategy and are unlikely to be driven by classification artifacts.

A quarterly analysis showed apparently increasing fitness toward 2023 (from ~0.20 in 2020–2022 to 0.65 in 2023 Q2); however, this trend is an artifact of right-censoring, since cases entering near the end of the observation window have short and necessarily more complete observed trajectories, and should not be interpreted as a genuine improvement in adherence.

Taken together, three patterns emerged across the analysis: (i) low overall conformance, driven predominantly by genuine structural deviations rather than incomplete deviations alone; (ii) event-level cohort-specific signatures that should be interpreted as descriptive and exploratory after patient-level correction; and (iii) transition-specific delays concentrated in identifiable pathways segments. The association between deviations and longer care duration was reproduced, and key conformance estimates are accompanied by uncertainty quantification.

4. Discussion

This study evaluated the conformance and temporal performance of telemedicine-based oncology care pathways against a guideline-derived BPMN reference model, using process mining over 182 patients treated between 2020 and 2023. The central finding is that overall conformance was low (4.95% of cases fully conforming by token-based replay; mean alignment fitness 0.19, 95% CI 0.16–0.22) even though the reference model itself was highly precise at the category level (precision 0.90). The combination of low fitness and high precision is informative: it indicates that the discrepancy does not arise from an over-permissive or poorly specified model but from observed trajectories that are systematically incomplete relative to the full guideline pathway. The joint interpretation of fitness, precision, generalization, and simplicity is increasingly recognized as necessary for assessing whether low conformance reflects model weakness, behavioral variability, or incomplete observations [32].

Interpreting the four model-quality dimensions jointly reinforces this reading. Precision is high at both abstraction levels (0.90 at the category level and 0.98 at the activity level), indicating that the reference model permits almost no behavior beyond what is actually observed regardless of label granularity; the low fitness therefore cannot be ascribed to an over-permissive specification at either level. Moderate generalization (0.52) is consistent with a guideline-derived model that encodes the normative pathway while only partially anticipating the diversity of real trajectories, which is the intended behavior of a normative rather than exhaustive model. The moderate-to-high simplicity (0.77–0.81 across levels) indicates that the model is compact and free of redundant structure, so its limited fitness reflects genuine divergence between observed and expected behavior rather than an unnecessarily convoluted reference. Taken together, precision, generalization and simplicity describe a tight, parsimonious model at both abstraction levels, supporting the interpretation of low conformance as a property of incomplete and heterogeneous observed data measured against a well-specified reference, rather than as evidence that the model is itself deficient. In practical terms these dimensions have direct implications for guideline monitoring. High precision means that the reference model rarely admits behavior that was not actually observed, so activities it flags as out-of-model can be trusted as genuine departures worth auditing rather than modeling artifacts. Moderate generalization implies that the model does not absorb every legitimate real-world variant, so novel but clinically valid pathways may surface as deviations and should be reviewed before being treated as non-adherence. High simplicity indicates that the model is compact and can be communicated to and audited by clinical teams. Taken together the metrics support using the model as an interpretable auditing instrument, provided its outputs are read as screening signals rather than definitive judgments of care quality.

Right-censoring inherent to the study design contributes to this gap. Only 2 of 182 cases (1%) reached a discharge event, and roughly 40% of cases began at the model's entry activity, so most pathways were observed mid-course rather than end-to-end, and a strict end-to-end replay necessarily classifies these partial trajectories as non-conforming. However, censoring is not the main driver of the low conformance. When incompleteness was modeled explicitly with prefix-alignments—which allow an observed trajectory to be a valid, unfinished prefix of a conforming path—the proportion of conforming cases increased only marginally, from 4.40% to 4.95%, showing that incompleteness alone does not rescue conformance. Decomposing the alignment cost showed that unreached future steps (censoring) accounted for only about 0.1% of the total misalignment, whereas about 99.9% corresponded to genuine deviations: observed activities inserted, repeated or reordered relative to the acyclic guideline pathway. A sensitivity analysis that progressively reclassified interior model steps as censoring—up to the extreme of attributing

every unreached step to censoring—still left genuine deviation dominating (at least 95% of the misalignment, bootstrap 95% CI) and did not move a single case into conformance (Figure 9). A complementary patient-level regression showed that a higher number of recorded events was associated with lower fitness and a higher share of telemedicine with higher fitness, whereas tumor cohort was not significant (Figure A1). Right-censoring is a well-recognized challenge in healthcare process mining [33]; our results show that although it must be accounted for when reading conformance figures, the low conformance in this cohort chiefly reflects genuine structural divergence from the reference pathway rather than incomplete observation.

The exceptionally high number of observed variants (172 variants across 182 cases) further illustrates the intrinsic heterogeneity of oncology care pathways. Unlike many operational processes, cancer trajectories are strongly influenced by differences in disease progression, treatment response, comorbidities, complications, and individualized clinical decision-making. As a result, patients with the same diagnosis may follow substantially different care sequences over time. This variability is often obscured in guideline-based or assumption-driven pathway models, which necessarily represent an abstraction of expected care rather than the full spectrum of real-world clinical behavior. Similar challenges have been reported in real-world oncology process mining, where interactive and incremental discovery techniques have been proposed to incorporate domain knowledge and manage highly variable clinical trajectories [34].

One of the strengths of process mining is its ability to make this variability explicit, allowing researchers and healthcare organizations to identify, quantify, and compare both frequent and infrequent pathways [35,36]. In this context, the large number of variants observed should not be interpreted solely as process instability, but also as evidence of the complexity and diversity inherent to oncology care delivery [3,33].

These findings are consistent with prior work reporting low adherence to clinical guidelines in oncology, including breast, colon and cervical cancer, where standardized guidelines often fail to anticipate complex patients, comorbidities, or disease progression [29,30]. At the event level, cohort-specific activity patterns were apparent, including over representation of systemic treatment and nursing care in colon cancer and of non-medical care in breast cancer. However, after aggregating activity profiles at the patient level to avoid pseudo-replication, these differences were no longer statistically significant after false-discovery-rate correction. These signatures should therefore be interpreted as exploratory process patterns rather than confirm cohort effects. This finding is consistent with the possibility that some deviations reflect clinically justified adaptations and resource availability rather than failures of care, in line with evidence that remote-care teams adjust the frequency and ordering of activities according to clinical evolution [31].

Importantly, the breast cancer cohort illustrates why a single binary metric can mislead. Breast cases showed 0% binary conformance yet a continuous alignment fitness (0.16) in line with cohorts that did record conforming cases, indicating that the 0% reflects right-censoring of incomplete trajectories rather than the least adherent care. Reporting a continuous fitness measure alongside the binary indicator, together with bootstrap confidence intervals and cohort-level statistical testing, provides a more faithful and reproducible picture of adherence than a single percentage. This methodological point, pairing an interpretable case-level indicator with a robust continuous measure and explicit uncertainty, is one of the contributions of the present analysis and is readily transferable to other guideline-monitoring settings.

The temporal analysis added a complementary, actionable layer. Transition times were strongly right-skewed, so we report medians with interquartile ranges rather than means, which a few outliers inflate (for example, the specialist-to-general-medicine tran-

sition averaged 45.5 days but had a median of only 10.6 days, IQR 5.6–30.2; Figure A2). Delays were not uniformly distributed but concentrated in specific transitions. Repetition of activities—the main source of deviation—was not specific to the telemedicine modality (Cramér's $V = 0.042$ between repetition and remote delivery; Figure A2), so bottlenecks should not be attributed to telemedicine per se. Clinically, the longest delays concentrated in the pancreas and colon cohorts. In pancreatic cancer, prolonged emergency-to-oncology and specialist transitions are consistent with the aggressive course of the disease, frequent decompensations that route patients through emergency care, and the need for multidisciplinary staging before treatment; possible interventions include fast-track emergency-to-oncology referral, pre-allocated multidisciplinary-committee slots and nurse-navigator coordination. In colon cancer, the repeated and prolonged specialist consultations suggest iterative pre-treatment work-up and scheduling frictions, which structured referral protocols, protocolized pre-operative pathways and telemedicine triage could help consolidate. Because process mining localizes these bottlenecks at the transition level, it can direct quality-improvement efforts to concrete points in the pathway that aggregate indicators would obscure [37].

From a governance perspective, the use of process mining requires not only analytical methods but also explicit data governance structures that define how process data are prepared, interpreted, and used for organizational decision-making [38]. In this context, these results illustrate how conformance checking could serve as a scalable, reproducible instrument for monitoring tele-oncology pathways against a national guideline framework such as the Chilean GES program, provided the metrics are interpreted with their methodological caveats and the approach is validated across multiple institutions. Although cases containing at least one out-of-model activity showed a longer crude average duration (about 673 versus 398 days), this association is largely an artifact of differential observation windows: after adjusting for exposure (modeling out-of-model events as a rate per patient-day) the association reversed (Poisson incidence-rate ratio about 0.64 per unit of duration, 95% CI 0.54–0.76) and the correlation between duration and the exposure-adjusted rate was weak and non-significant (Spearman $\rho = 0.13$, $p = 0.10$; Figure A2). Out-of-model activity should therefore not be read as a cause of longer care; rather, longer-observed patients simply accrue more events of every kind. This reinforces the value of jointly analyzing structural deviations and temporal performance when evaluating real-world oncology care. Taken together, the findings argue for guideline and monitoring frameworks that explicitly accommodate hybrid in-person/telemedicine care and that account for incomplete, right-censored trajectories when adherence is assessed from real-world data [39].

Limitations

This study has several limitations. First, the event log is affected by right-censoring, as many patient trajectories were still ongoing within the 2020–2023 observation window. Consequently, conformance metrics, particularly binary token-based replay, may underestimate adherence to the reference pathway. Second, the BPMN model was derived from Clinical Practice Guidelines and therefore represents a normative abstraction of the expected process rather than a complete representation of all activities occurring in routine oncology care. Some detected deviations may therefore reflect clinically relevant activities or local practices that are not explicitly represented in the model. Third, the analysis was conducted in a single private healthcare network and during a period strongly influenced by the COVID-19 pandemic, which may limit the generalizability of the findings. Future studies should evaluate alternative modeling strategies, incorporate longer observation windows, and replicate the analysis across different institutions and clinical contexts. The findings derive from a single private healthcare network and a pandemic-influenced period; they are representative of this specific setting

and should not be generalized to the national implementation of the GES program without multi-institutional, public-sector validation. The reference model was validated through an iterative qualitative consensus with four experts rather than a formal Delphi process with a quantified inter-rater agreement, and the expert-guided activity mapping carries a residual risk of subjective bias. Finally, the lung, pancreas and rectum cohorts are small ($n = 9\text{--}11$), so their cohort-level results are exploratory.

5. Conclusions

This study demonstrates the feasibility and utility of process mining for evaluating the conformance of telemedicine-based oncology pathways against standardized Clinical Practice Guidelines. Rather than confirming uniformly high adherence, the analysis shows that real-world tele-oncology trajectories are highly variable and, within a 2020–2023 observation window, predominantly incomplete: low conformance reflects genuine structural deviation from the guideline pathway far more than right-censoring—which sensitivity analyses show contributes only marginally—rather than uniform clinical non-adherence. Quantifying adherence with paired binary and continuous metrics, uncertainty estimates, and cohort-level testing enables an objective and reproducible characterization of where and how real pathways diverge from the normative model.

Methodologically, the approach provides a reproducible workflow that health institutions can embed to audit their processes, localize transition-specific delays, and monitor patient trajectories against government requirements. Clinically, the results suggest that tele-oncology pathways can be monitored as part of longitudinal cancer care, underscoring the need for standards that accommodate hybrid care and that integrate activities not explicitly represented in guideline-derived process models. Overall, the study contributes to consolidating conformance checking via process mining as a tool for evaluation, traceability, and continuous improvement in digital oncology care, while clarifying how its metrics should be interpreted in the presence of censored, real-world data.

Author Contributions: M.C.: Conceptualization; Methodology; Data curation; Formal analysis; Investigation; Visualization; Writing—original draft; Writing—review and editing. E.C.: Conceptualization; Methodology; Formal analysis; Validation; Writing—original draft; Writing—review and editing. D.C.: Conceptualization; Supervision; Writing—review and editing. S.V.: Data curation; Investigation; Formal analysis. S.M.: Formal analysis; Validation. T.M.: Formal analysis; Validation. S.H.: Methodology; Supervision; Writing—review and editing. E.R.: Conceptualization; Methodology; Supervision; Project administration; Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research and Development Agency (ANID) through the National Fund for Scientific and Technological Development (FONDECYT), project number 11230708. Steffen Härtel receives support by MINEDUC grant RED 21994, BASAL FB210005, DAAD 57662855, Centro CTI220001, and NCN2024_068.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Scientific Ethics Committee for Health Sciences of the Pontificia Universidad Católica de Chile (CEC-UC) (protocol code ID220310004, approved 9 March 2023).

Informed Consent Statement: Patient consent was waived because the study used anonymized clinical data extracted retrospectively, with no patient contact or intervention, and was classified as minimal risk by the approving ethics committee.

Data Availability Statement: The clinical and administrative data analyzed in this study contain sensitive patient information and are subject to institutional and national regulations governing the protection of personal health data. For ethical reasons and in accordance with the data governance policies of the Pontificia Universidad Católica de Chile and the approved study protocol, these data

cannot be made publicly available. Access to the data set is restricted and may be granted only under formal institutional agreements and with appropriate ethical approval. Due to Pontificia Universidad Católica regulations, the source code used in this study cannot be publicly shared. Nevertheless, the complete pseudocode, a synthetic de-identified event log and an executable reference implementation are openly available in the public repository referenced in Appendix C.

Acknowledgments: We would like to thank the health network, Red de Salud UC CHRISTUS and the UC Cancer Center for their collaboration in this research and for allowing us to use their patient information. In addition, we thank the Universidad de Chile and the Pontificia Universidad Católica de Chile for their academic support during the undertaking of this study. Finally, we extend our gratitude to the National Research and Development Agency (ANID) for the financial support provided through the National Fund for Scientific and Technological Development (FONDECYT) project no. 11230708, “Improving the Patient Journey in Telemedicine using Process Mining”.

Conflicts of Interest: The authors declare no conflicts of interest. The commercial process mining platform used (Celonis EMS) was accessed through an academic license, and the vendor had no role in study design, data analysis, interpretation of results, or the decision to publish.

Abbreviations

The following abbreviations are used in this manuscript:

BPMN	Business Process Model and Notation
EHR	Electronic Health Record
EMS	Execution Management System
GES	Garantías Explícitas en Salud (Explicit Health Guarantees)
IRB	Institutional Review Board
MINSAL	Ministry of Health of Chile
PM	Process Mining

Appendix A

Table A1. Activity mapping used for event standardization.

Event Log Activity	BPMN Activity
Hospital discharge	Hospital discharge
Radiotherapy discharge	Not mentioned in the BPMN model
Telemedicine Radiotherapy Discharge	Not mentioned in the BPMN model
Oncology Care	Oncology Consultation
Emergency Care	Not mentioned in the BPMN model
Surgery	Surgery
Radiotherapy Committee	Radiotherapy Committee
Oncology Committee	Oncology Committee
Face to face Specialist Medical Consultation	Specialist Medical Consultation
Telemedicine Specialist Medical Consultation	Specialist Medical Consultation
Nursing Care	Nursing Care
Speech therapy Care	Non-medical Care
Physiotherapy Care	Non-medical Care
Nutritionist Care	Non-medical Care
Nursing Assistant Care	Nursing Care
Endoscopy	Procedures
Outpatient otolaryngology examination	Procedures
Interventional Radiology Procedure	Procedures
Psychology	Psychology
Telemedicine Psychology	Psychology
Systemic Treatment	Systemic Treatment
Radiotherapy	Radiotherapy

Appendix B. Statistical Validation and Robustness

This appendix collects the confirmatory statistical analyses that support the Results without altering the process-mining narrative of the main text: the patient-level predictors of alignment conformance (Figure A1), and the exposure-adjusted performance and delivery-modality analyses (Figure A2). It also reports the time-to-completion and care-duration models that formally quantify the right-censoring (Figure A3).

At the patient level, an ordinary-least-squares model of alignment fitness showed that a higher number of recorded events was associated with lower fitness, whereas a higher telemedicine share was associated with higher fitness (coefficients with 95% confidence intervals; Figure A1a). Tumor cohort was not a significant predictor, and the association between activity presence and tumor cohort was likewise non-significant after false-discovery-rate correction, with effect sizes remaining small throughout (Cramér's V 0.08–0.15; Figure A1b). Taken together, these results indicate that conformance is shaped primarily by trajectory length and care modality rather than by the underlying diagnosis, which is consistent with a single reference model applying across cohorts.

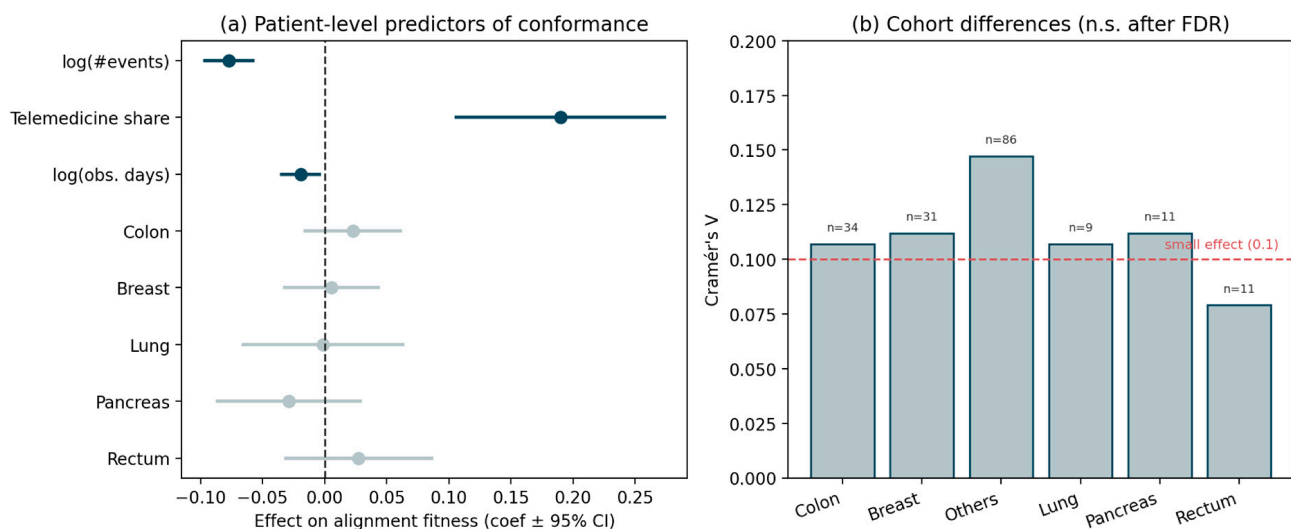


Figure A1. Patient-level analysis: predictors of alignment conformance ((a) ordinary least squares, coefficients with 95% confidence intervals) and association between activity presence and tumor cohort, dark (navy) markers indicate statistically significant predictors (95% CI excluding zero) and light gray markers non-significant ones (95% CI including zero) ((b) Cramér's V , non-significant after false-discovery-rate correction). “#” denotes ‘number of’ (e.g., #events is the number of events per patient).

Finally, the crude association between out-of-model activity and longer care duration reversed after adjusting for exposure: the exposure-adjusted rate of out-of-model events per 100 patient-days declined as the observation window lengthened (Poisson incidence-rate ratio about 0.64), showing that the apparent excess reflected longer follow-up rather than more deviation per unit of time (Figure A2a,b). Because transition times were right-skewed, they are summarized as medians with interquartile ranges rather than means, which would otherwise be inflated by a small number of long delays (Figure A2c). Telemedicine accounted for 18.5% of all events and was concentrated in oncology consultations, confirming that remote care is embedded at specific, identifiable points of the pathway rather than distributed uniformly across activities (Figure A2d).

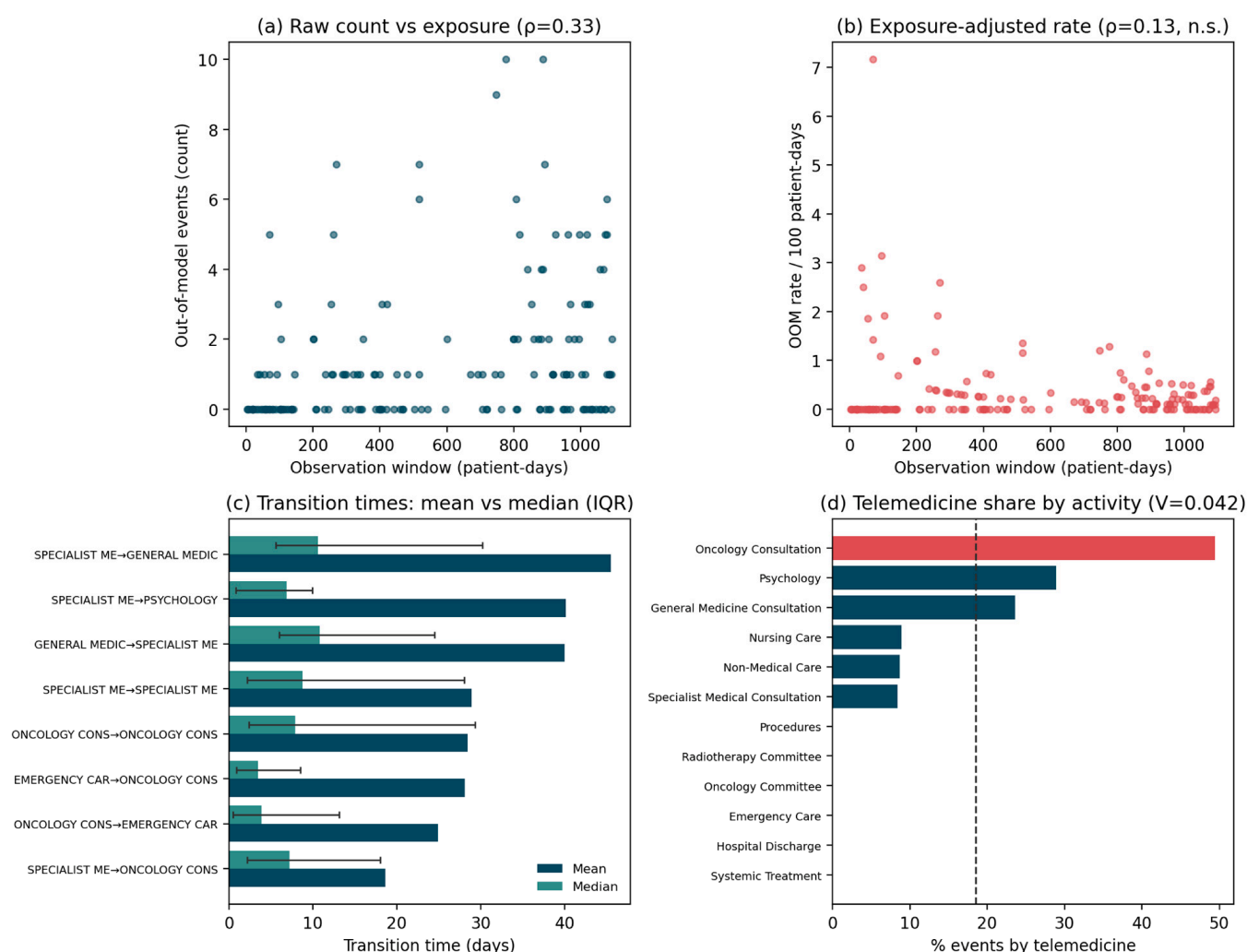


Figure A2. Performance and exposure analyses: out-of-model event count (a) and exposure-adjusted rate per 100 patient-days (b) versus the observation window; mean versus median transition times with interquartile ranges (c), dark (navy) bars show the mean and green bars the median transition time; error bars indicate the interquartile range; and telemedicine share by activity type (d), the dashed vertical line marks the overall telemedicine share and the highlighted (red) bar marks the activity with the highest telemedicine share.

To characterize the incompleteness of the observed trajectories formally, we fitted a Kaplan–Meier time-to-completion model (event = terminal hospital discharge) on the 173 patients with a positive observation window; nine single-event cases were excluded because a time-to-completion is undefined for them. Only 2 of these 173 patients reached a terminal discharge, the median time-to-completion was not reached, and the cumulative incidence of completion was 1.7% at the end of follow-up (Figure A3), which quantifies the right-censoring inherent to the 2020–2023 study window. A complementary Gamma generalized linear model (log link) of total care duration, fitted on the same cohort, showed that a higher number of recorded events was associated with a longer duration ($\exp(\beta) = 1.010$ per additional event, 95% CI 1.007–1.013, $p < 0.001$), whereas telemedicine share and tumor cohort were not significant predictors; the full model is reported in Table A2.

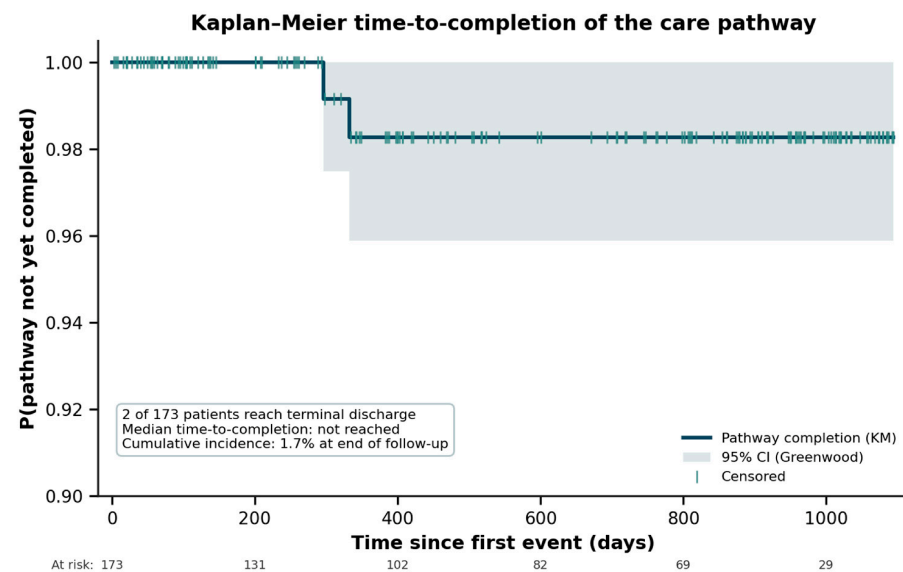


Figure A3. Kaplan–Meier time-to-completion of the care pathway (event = terminal hospital discharge; $n = 173$ patients with a positive observation window). The curve is nearly flat (2 completions, median not reached, cumulative incidence 1.7% at end of follow-up), formally quantifying the right-censoring of the observed trajectories.

Table A2. Gamma generalized linear model (log link) of total care duration (person-days) on the 173 patients with a positive observation window. $\exp(\beta)$ is the multiplicative change in expected duration; cohort terms are relative to the Others group.

Predictor	$\exp(\beta)$	95% CI	p
Intercept	322.81	247.57–420.91	<0.001
Number of events	1.010	1.007–1.013	<0.001
Telemedicine share	1.519	0.838–2.753	0.168
Colon (vs. Others)	0.763	0.555–1.049	0.096
Breast (vs. Others)	0.983	0.721–1.340	0.914
Lung (vs. Others)	0.641	0.381–1.079	0.094
Pancreas (vs. Others)	0.853	0.531–1.371	0.511
Rectum (vs. Others)	0.717	0.443–1.162	0.177

Appendix C. Reproducibility

The clinical event log contains protected patient information, and the institutional implementation of the analysis code cannot be redistributed. To make the methodology fully auditable, a public reproducibility package is provided. It contains: (i) the complete pseudocode of the analysis pipeline (event-log preprocessing and vocabulary mapping; censoring-aware conformance and its decomposition into right-censoring versus genuine deviation; exposure-adjusted performance and delivery-modality analysis; and the sensitivity analysis of the decomposition); (ii) a synthetic, de-identified event log that reproduces the structure of the study data without any real patient records; and (iii) an executable reference implementation that runs the whole analysis on the synthetic log and regenerates every result type reported in this study (conformance decomposition, modality effect sizes, exposure-adjusted rates, patient-level tests, robust transition times, regression and survival models, and the sensitivity analysis).

The reproducibility package is openly available at: <https://github.com/evchiu/Oncology-pathways-via-telemedicine-using-process-mining-A-study-in-a-Chilean-hospital-network> (accessed on 29 July 2026).

References

- Ministerio de Salud. *Plan Nacional de Cáncer 2018–2028*, 1st ed.; Ministerio de Salud: Santiago, Chile, 2018.
- Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2024**, *74*, 229–263. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rojas, E.; Munoz-Gama, J.; Sepúlveda, M.; Capurro, D. Process Mining in Healthcare: A Literature Review. *J. Biomed. Inform.* **2016**, *61*, 224–236. [\[CrossRef\]](#)
- Santos, A.; Leal, G.; Balancieri, R. Process mining in healthcare: A tertiary study. *BMC Med. Inform. Decis. Mak.* **2025**, *25*, 306. [\[CrossRef\]](#)
- World Health Organization. *WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening*; World Health Organization: Geneva, Switzerland, 2019.
- Mfouth Kemajou, P.; Mbanya, A.; Coppieters, Y. Digital Approaches in Post-COVID Healthcare: A Systematic Review of Technological Innovations in Disease Management. *Biol. Methods Protoc.* **2024**, *9*, bpae070. [\[CrossRef\]](#)
- West, H.; Barzi, A.; Wong, D. Telemedicine in Cancer Care Beyond the COVID-19 Pandemic: Oncology 2.0? *Curr. Oncol. Rep.* **2022**, *24*, 1843–1850. [\[CrossRef\]](#)
- División de Prevención y Control de Enfermedades, Ministerio de Salud. Guías Clínicas AUGÉ—DIPRECE. Available online: <https://diprece.minsal.cl/le-informamos/auge/acceso-guias-clinicas/guias-clinicas-auge/> (accessed on 1 June 2025).
- Shirke, M.M.; Shaikh, S.A.; Harky, A. Implications of Telemedicine in Oncology During the COVID-19 Pandemic. *Acta Biomed.* **2020**, *91*, e2020022. [\[CrossRef\]](#)
- Shirke, M.M.; Shaikh, S.A.; Harky, A. Tele-Oncology in the COVID-19 Era: The Way Forward? *Trends Cancer* **2020**, *6*, 547–549. [\[CrossRef\]](#)
- van der Aalst, W.M.P. *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2011.
- Rebuge, Á.; Ferreira, D.R. Business Process Analysis in Healthcare Environments: A Methodology Based on Process Mining. *Inf. Syst.* **2012**, *37*, 99–116. [\[CrossRef\]](#)
- Partington, A.; Wynn, M.; Suriadi, S.; Ouyang, C.; Karnon, J. Process Mining for Clinical Processes: A Comparative Analysis of Four Australian Hospitals. *ACM Trans. Manag. Inf. Syst.* **2015**, *5*, 19. [\[CrossRef\]](#) [\[PubMed\]](#)
- 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. [\[CrossRef\]](#) [\[PubMed\]](#)
- van der Aalst, W.M.P. *Process Mining: Data Science in Action*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2016.
- Rozinat, A.; van der Aalst, W.M.P. Conformance Checking of Processes Based on Monitoring Real Behavior. *Inf. Syst.* **2008**, *33*, 64–95. [\[CrossRef\]](#)
- Adriansyah, A. Aligning Observed and Modeled Behavior. Ph.D. Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands, 2014. [\[CrossRef\]](#)
- Oliart, E.; Rojas, E.; Capurro, D. Are We Ready for Conformance Checking in Healthcare? Measuring Adherence to Clinical Guidelines: A Scoping Systematic Literature Review. *J. Biomed. Inform.* **2022**, *130*, 104076. [\[CrossRef\]](#) [\[PubMed\]](#)
- Savino, M.; Chiloiro, 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. [\[CrossRef\]](#) [\[PubMed\]](#)
- Benevento, E.; Pegoraro, M.; Antoniazzi, M.; Beyel, H.H.; Peeva, V.; Balfanz, P.; van der Aalst, W.M.P.; Martin, L.; Marx, G. Process Modeling and Conformance Checking in Healthcare: A COVID-19 Case Study. In *Process Mining Workshops*; Montali, M., Senderovich, A., Weidlich, M., Eds.; Springer Nature: Cham, Switzerland, 2023; pp. 315–327.
- Lu, X.; Mans, R.S.; Fahland, D.; van der Aalst, W.M.P. Conformance Checking in Healthcare Based on Partially Ordered Event Data. In Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), Barcelona, Spain, 16–19 September 2014; pp. 1–8. [\[CrossRef\]](#)
- Leonardi, G.; Montani, S.; Striani, M. Checking Medical Process Conformance by Exploiting LLMs. *Appl. Sci.* **2025**, *15*, 10184. [\[CrossRef\]](#)
- del Río-Ortega, A. On the Definition and Analysis of Process Performance Indicators. Ph.D. Thesis, Universidad de Sevilla, Seville, Spain, 2012.
- Arthofer, K. Adherence to Guidelines Supported by Process Mining. *Stud. Health Technol. Inform.* **2025**, *324*, 27–28. [\[CrossRef\]](#) [\[PubMed\]](#)
- 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, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 425–438.
- Marin-Castro, H.M.; Tello-Leal, E. Event Log Preprocessing for Process Mining: A Review. *Appl. Sci.* **2021**, *11*, 10556. [\[CrossRef\]](#)

27. Rabbi, F.; Banik, D.; Ibne Hossain, N.; Sokolov, A. Using process mining algorithms for process improvement in healthcare. *Healthc. Anal.* **2024**, *5*, 100305. [[CrossRef](#)]
28. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [[CrossRef](#)]
29. Tchelebi, L.T.; Shen, B.; Wang, M.; Potters, L.; Herman, J.; Boffa, D.; Segel, J.E.; Park, H.S.; Zaorsky, N.G. Nonadherence to Multimodality Cancer Treatment Guidelines in the United States. *Adv. Radiat. Oncol.* **2022**, *7*, 100938. [[CrossRef](#)] [[PubMed](#)]
30. Bierbaum, M.; Rapport, F.; Arnolda, G.; Tran, Y.; Nic Giolla Easpaig, B.; Ludlow, K.; Clay-Williams, R.; Austin, E.; Laginha, B.; Lo, C.Y.; et al. Rates of Adherence to Cancer Treatment Guidelines in Australia and the Factors Associated with Adherence: A Systematic Review. *Asia-Pac. J. Clin. Oncol.* **2023**, *19*, 618–644. [[CrossRef](#)] [[PubMed](#)]
31. Knudsen, K.E.; Willman, C.; Winn, R. Optimizing the Use of Telemedicine in Oncology Care: Postpandemic Opportunities. *Clin. Cancer Res.* **2021**, *27*, 933–936. [[CrossRef](#)] [[PubMed](#)]
32. Leewis, S.; Smit, K.; van de Hoef, A. Precision, Fitness, Generalization, and Simplicity as Quality Dimensions for Decision Discovery Algorithms. *Appl. Sci.* **2025**, *15*, 11060. [[CrossRef](#)]
33. Munoz-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](#)] [[PubMed](#)]
34. Schuster, D.; Benevento, E.; Aloini, D.; van der Aalst, W.M.P. Analyzing Healthcare Processes with Incremental Process Discovery: Practical Insights from a Real-World Application. *J. Healthc. Inform. Res.* **2024**, *8*, 523–554. [[CrossRef](#)] [[PubMed](#)]
35. Rojas, E.; Fernández-Llatas, C.; Traver, V.; Munoz-Gama, J.; Sepúlveda, M.; Herskovic, V.; Capurro, D. PALIA-ER: Bringing Question-Driven Process Mining Closer to the Emergency Room. In *Proceedings of the BPM Demo Track and BPM Dissertation Award 2017, Co-Located with the 15th International Conference on Business Process Management (BPM 2017), Barcelona, Spain, 13 September 2017*; CEUR Workshop Proceedings: Aachen, Germany, 2017; Volume 1920.
36. Rojas, E.; Capurro, D. Characterization of Drug Use Patterns Using Process Mining and Temporal Abstraction Digital Phenotyping. In *Business Process Management Workshops*; Daniel, F., Sheng, Q.Z., Motahari, H., Eds.; Springer Nature: Cham, Switzerland, 2019; pp. 187–198. [[CrossRef](#)]
37. Rojas, E.; Cifuentes, A.; Burattin, A.; Munoz-Gama, J.; Sepúlveda, M.; Capurro, D. Analysis of Emergency Room Episodes Duration Through Process Mining. In *Business Process Management Workshops*; Daniel, F., Sheng, Q.Z., Motahari, H., Eds.; Springer Nature: Cham, Switzerland, 2019; pp. 251–263. [[CrossRef](#)]
38. Goel, K.; Martin, N.; ter Hofstede, A. Demystifying data governance for process mining: Insights from a Delphi study. *Inf. Manag.* **2024**, *61*, 5. [[CrossRef](#)]
39. Zon, R.; Kennedy, E.; Adelson, K.; Blau, S.; Dickson, N.; Gill, D.; Laferriere, N.; Lopez, A.M.; Mulvey, T.; Patt, D.; et al. Telehealth in Oncology: ASCO Standards and Practice Recommendations. *JCO Oncol. Pract.* **2021**, *17*, 9. [[CrossRef](#)] [[PubMed](#)]

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.