International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Data-Driven Framework to Minimize Patient Intake Errors in Healthcare Using SQL and RPA

Akhilesh Obalannavar

Business System Analyst, Information Technology

ABSTRACT

Health care Patient intake errors are very dangerous to patient safety, operational efficiency and it has detrimental impact on their financial performance and this is because of manual data entry, miscommunication and disjointed systems. This paper suggests an evidence-driven approach that is a hybrid of SQL-based validation and Robotic Process Automation (RPA) techniques to reduce such mistakes in hospital intake processes. Real-time anomaly detection queries were created, at the same time RPA bots (Pega Systems) carried out data validations, completeness checks of forms, and EHR integration. Findings showed a 90 percent reduction in the time taken to process individual patients and 43 percent drop in the overall intake errors but a 25 percent rise in delays among the staff because of more accurate monitoring. The framework shows the possibility of rule-based automation to increase data integrity, compliance and improve efficiency as well as show the necessity to further optimize human-dependent processes. To confirm these results and further investigate AI-enhanced anomaly detection with even more complicated types of errors, researchers should test the outcomes in the real clinical setting. This experiment will provide a scalable, interoperable product to the healthcare institutions interested in automating intake processes, resolving administrative complexity, and thus improving patient outcomes by automating the process intelligently.

KEYWORDS: Artificial Intelligence, Robotic Process Automation, Hospital Nutrition, Medicine, Patient Intake.

1.0. INTRODUCTION:

The first level of health care provision that is of particular importance is patient intake that usually presupposes a collection of the relevant information, such as medical history of a patient, insurance details, form of consent. The process is critical, but subject to error, which can adversely affect patient safety, delays treatments and adds to costs. Patient data that are not accurate and up to date can be a source of mismatch in diagnosis and treatment and billing must be done manually which doubles up on issues, there is a potential lack of communication between departments and the systems that these departments use are not integrated with one-another[1].

Research in recent past has been used to show the scope and implication of such errors. To illustrate, up to 5 percent of adults in the United States suffer diagnostic errors in the outpatient situation each year, causing a significant amount of patient harm. Also, according to the World Health Organization, the harmful diagnostic errors were present at least in 0.7 percent of adult admissions, which implies the great significance of this issue on a global scale.

A feasible solution to these problems is provided by the automation technology. Patient entry into systemized alignment charts via computer automation: Robotic Process Automation (RPA) implementation of PsychAE and SQL implementation of validation between systems and system integration. RPA and SQL can be used in data entry, data validation and data integration between systems within the patient entry process. RPA has proven to reduce in administrative errors and has proved efficient in health care. RPA combined with Artificial Intelligence (AI) has also demonstrated the potential of enhancing the accuracy of diagnosis and administrative processes, increasing patient care.

This paper attempts to mitigate intake mistakes involving patients in healthcare based on a database-driven approach using SQL and RPA. The goals are:

1. To categorize typical sources of errors in the patient intake process.
2. To propose an automated solution with SQL and RPA to prevent these errors.
3. To test the effectiveness of such an approach using a synthetic data auditing and modelling approach.

The complexity of care management and the rise in the number of patients coupled with the manual patient intake process have led to the high possibility of errors in patient intake, which may have a domino effect along the care continuum. According to the studies, almost 80 percent of life-threatening medical errors are caused by poor communication at initial patient interaction, and improper or incomplete intake information is one of the reasons behind medication errors, delayed treatment practices, and wrong bills. The conventional paper-based and manual data entry electronic systems bear several weaknesses such as incomprehensible handwritings, transpositions, incomplete sections, and inability to obtain the current information about patients. At the same time, the professional life of healthcare intake staff is usually driven by time restrictions and the necessity to deal with a variety of competing tasks, resulting in a scenario where any type of error can easily spread. The financial repercussions are also significant - the Centers for Medicare & Medicaid Services estimate that the registration and intake-related errors are the cause of the near 30 percent of the claim denials, resulting in millions of dollars in the lost revenue and rework every year by the healthcare organization. In addition to the administrative cost, such mistakes may deteriorate patient trust since they can be made to repeat the same information constantly or may not receive care on time because of garbage data. The recent spread of the COVID-19 pandemic had demonstrated that these vulnerabilities cannot be ignored, as healthcare systems have suffered shortages in the face of growing patient numbers, thus manifesting the necessity of more robust solutions based on automation. The study will respond to these obstacles by proposing a groundbreaking system of integrated SQL and RPA to validate the structured and automate the workflow. SQL offers sufficient data integrity checks that can detect an absence, inconsistency, or incorrect information on patients by the use of accurate query and pattern matching and RPA facilitates exactly repetitive work in intake that must take place accurately and reliably every time. This combination makes a closed system wherein the quality of data is observed and checked all the way during the intake process. Notably, we will use the Fault Tree Analysis as part of our methodology to determine and define in a systematic moe the root cause of the intake errors to take the most effective intervention in automation in that area. The artificial dataset that we created contains well-tuned error patterns to represent real-life intake scenarios so that the given rules of our automated validation could be thoroughly tested before they are put to practical use. All errors that are particularly high in impact in our three categories of missing/incorrect patient data, in-signed consent forms and delay in staff entry adhere to the most critical vulnerabilities in existing intake patterns.

Collaboration with the current EHR systems with the HL7/FHIR standards makes it feasible to implement, and there are no needs to replace the expensive systems. The present study is based on the emerging body of evidence that smart automation can be a powerful multiplier of efficiency in many processes within the healthcare sector and can also improve the quality of data, with recent results indicating that RPA project implementations 40-60 percent in similar environments reduce the rate of administrative errors made. Our method is a further elaboration of such findings since it reveals how a V-shaped integration of SQL and RPA in the given application with data validation on one side and workflow automation on the other side will have a synergistic effect unattainable had each of those technologies been used separately. Such a timely framework is especially relevant due to the increasing focus of the healthcare industry on interoperability and data standardization efforts such as Fast Healthcare Interoperability Resources (FHIR), because this framework shows some of the potential ways to ensure data quality at the point of entry which is a crucial requirement before successful health information exchange can occur.

2.0. LITERATURE REVIEW

The research has highlighted the level of systemic failures within hospital nutrition services that are, in fact, being overlooked as potential risks to patient safety in the same way as medication errors [2]. It is powerful in its ability to document the types and reasons for service failure, and provides a firm footing for those willing to take the step to automate, e.g., via computerized diet order entry systems. But its emphasis on inpatient dietary services restricts applicability to other kinds of patient admission processes. However, researchers sociolinguistic approach examines the influence of communication during outpatient intake on the goals of organization and the patient [3]. Their research reveals how consumption mistakes frequently originate from misalignments between expectations, ill-defined rhythms and structural pressures. Although it is rich in details, this study does not provide the quantitative information on error frequency or automation potential. Collectively, these papers demonstrate that follow-up errors are systemic as well as interactional, rooted in workflow gaps and communication failures, and point to the importance of integrative technology-based solutions. Researchers makes a strong argument to exploit SQL-based PIS for improving healthcare efficiency [4]. Through the lens of log storage, retrieval, and analysis, the paper shows how some critical patient wait times can be directly reduced and resource allocation improved with just SQL. Its power is in its use, creating real-time, tangible gains in workflow and decision making. But the case study does not address issues of scalability, data interoperability, or integration with legacy systems. On the contrary, researchers instead adopt a more systems-level approach by constructing a Smart Hospital Management System (HMS) which employs SQL to manage data at the backend [5]. As a plus, it takes a modular approach to integration (e.g., appointments, billing, records), security and compliance—often neglected by more narrow SQL offerings. It spans a wider range of operational aspects, yet is predominantly theoretical, with few practical adopted effects. Collectively, the two studies confirm SQL's important role in the structuring and processing of health care data and also stress limitations in practical scalability, integration. Researchers provides a detailed investigation on AI and RPA adoption in healthcare, with a focus on administrative and clinical workflow optimization [6]. Based on examples in actual cases, the study validates that RPA vastly enhances productivity for automated claims processing and scheduling, for example, whereas AI helps enhance diagnoses and prognoses. RPA does repetitive work to enhance efficiency and precision. These technologies are rapidly transforming the field of health care through enhancing more accurate diagnosis, speeding up in administration, decreasing operational time and enhancing the treatment of patients [12].A revolutionary technology that is under development to ease the daily work responsibilities of the users; robotic process automation (RPA) promises to assist people by freeing them of the tedious and boring chores [10]. Robotic Process Automation (RPA) is an innovative solution based on the software bot that imitates the human reaction to the digital system to automate the operations and free the human resources to perform some more difficult and strategic operations [13]. This study's primary methodological strength is its empirical approach and balanced view in terms of benefits and challenges, in particular when considering how to facilitate integration with legacy systems and the reluctance of staff. However, research studies focus on the use of Pega systems a low-code BPM tool), complemented by RPA to demonstrate synergistic benefits [7]. The account featuring Pega notes it is the leader in orchestrating work across all lines of business, while RPA drives rule-based automation, unlocking hyper automation at enterprise scale. Its orientation toward tactical deployment and use cases is probably quite useful for operational planning, but there are no empirical performance data in healthcare-relevant settings. Together, these articles articulate how RPA and platforms such as Pega are transforming to power intelligent, scalable healthcare automation. Researchers successfully used Fault Tree Analysis (FTA) for the issue of wrong-site surgeries (WSS), which is a low-probability but high-impact surgical error. By integrating information from 37 articles, the paper develops a fault tree for the key failure paths of preoperative processes. By this methodological implementation of AND/OR gates, it becomes easy to see where systemic redundancy is present or absent. Its strength is in a systematic approach to assessment of the extent to which vulnerabilities may exist in the workflow of the given surgical scenario, but it may potentially not be applicable beyond surgical contexts. Researchers take FTA to the next level by combining it with Failure Modes and Effects Analysis FMEA to investigate medication administration errors, one of the most common and serious forms of healthcare errors [8]. This combined approach makes it possible not only to detect risks proactively (FMEA) but also reactively (FTA) and so gives a more complete approach. Taken together, the two studies confirm the importance of FTA in healthcare error analysis, highlighting its flexibility across various error domains for prevention planning. Healthcare automation means the application of technology in order to minimize or eliminate human activity in certain tasks or processes. Many studies incorporate a wide usage of healthcare information such as human material, and clinical data and recognize its importance [11].It may be anything between easy administrative tasks, data entry and scheduling, and highly complicated clinical procedures, like diagnostics and surgery [9].

There is already a set of studies that pose an argument of implementing automation in healthcare process, though it shows that an alarming number of discoveries aimed at this problem remain in the area of integrated technology solutions lack. On the one hand, such studies also show the systemic character of the errors in the field of specialized practice, such as nutrition services. On the other hand, dietary systems represent only a small part of the spectrum of vulnerabilities related to admission that encompass clinical departments. This shortcoming stands out especially since the intake of patients is the starting point of the gateway through which problems with data quality may spread across the entire chain of care. Although the sociolinguistic approach of different researchers gives a good qualitative reflection regarding the instances of communication breakdowns in outpatient encounters, their failure to yield quantitative measures leaves the medical administration bereft of possible standards to asses automation potential or track the improvement. This is one of the crucial knowledge gaps as healthcare systems need not only a qualitative judgment of the sources of errors and clarity on their causes but also quantitative data on the prevalence of errors in order to make the appropriate decision regarding technology investment and target prioritizing the interventions.

The study of SQL-based systems shows equally inconclusive knowledge. The ability of SQL to enhance certain operations such as the wait times of patients is convincingly warranted in a variety of case studies, but the fact that it does not fix their scalability and interoperability issues makes it severely constrained in terms of its real-life value regarding its intended application in the context of a complicated healthcare setting. According to widely used industry estimates, contemporary hospitals usually use dozens of interconnected yet incompatible systems - EHRs, billing systems, laboratory systems, and others - which generate data silos that cannot be crossed by a simple SQL implementation. According to a research, more detailed developments in HMS framework theoretically circumvent these issues of integration by means of their modular design, however, without any empirical implementation data, there are questions left unanswered about its real-world performance, user adoption hurdles, and the actual cost of system-wide implementation. The main issue with this theoretical-practical split is that healthcare technology solutions have regularly proved to fail not out of limitations on the technical side but rather through unpredicted workflow interferences, resistance by the staff, or undiscovered integration intricacies that are not brought up until implementation.

More promising evidence is presented in the literature on RPA and AI in healthcare with certain caveats. The results of the research regarding the optimization of the workflow and the review of the applications of RPA clearly define that automation can enormously enhance the efficiency of the administration and clinical decision-making. Nevertheless, most of these studies consider automation in isolation with individual components of the process and look at activities such as claims processing or scheduling appointments but not overall process transformation. The healthcare intake procedure, in turn, is more of a whole than a sum of its parts, as a variety of interconnected processes are involved in the check-in procedure identity verification, insurance validation, medical history gathering, consent handling, and department-specific needs, each of them has its faint possibilities of failure and software automation issues. Actual research on the use of RPA in the healthcare industry is starting to embrace this level of complexity, though, as is common with studies in this field, this level of inquiry is restricted to the post-intake processes and not to the important initial process of data capture where errors initially arise and thus can most effectively be avoided.

FTA research can be a great source of methodology ideas on how to develop error analysis, but it has serious limitations on its scope. The efforts of the researchers on the wrong-site surgeries, the research on medication errors indicate the power of FTA to analyze a complex pathway of failures in high-hazard clinical settings. Yet, such applications only target a specific process of patient care errors when the real potential of applications lies with the administrative and data quality errors, preceding and facilitating clinical decision-making processes. It is a really big blind spot since, as the WHO statistics demonstrate, diagnostic and treatment failures are often the direct result of inaccuracies or incompleteness of information about the patient obtained at the intake stage. In the existing literature there has been no systematic use of FTA on these upstream administrative processes especially in the study on how error in the data on intake systematically flows through the systems to form downstream clinical risks. This gap is filled in our research, where FTA methodology has been applied to patient intake processes with the purpose of outlining the failure points where better errors can be intercepted and therefore prevent the transmission to care delivery processes.

There is a preliminary literature on intelligent automation platforms such as Pega Systems that promises to help tackle these challenges of integration. The analysis given by researchers indicates that low-code BPM software coupled with RPA has the potential to manage complex operational tasks in the health sector by connecting two or more systems. Nevertheless, according to the researchers, the majority of published case studies concern financial and back-office processes instead of the clinical-facing ones, such as patient intake. This leaves a doubt whether these platforms are really able to address the specific needs of intake workflows such as real-time data validation, patient identity validation, dynamically complete forms and provide instant clinical decision support. Besides, the literature does not give much attention to the key human elements in the installation of automation. The authors of research have to deal with operational efficiency Researchers do not discuss the technical studies and most of them neglect the consequences of change management requirements, staff training demands, redesigning the workflow and approaches to addressing clinician skepticism regarding automated systems.

The present analysis makes it clear that there are several critical gaps in the current literature. To begin with, integrated solutions bundling both strong data validation (SQL) with an automation of workflow (RPA) when it comes to patient intake processes do not exist. Secondly, it is a combination of both technical and human variables that has rarely been studied in real-life clinic, yet such studies were done separately. Third, the current error analysis techniques such as FTA have not been coherently used to study the data quality problems in an administrative setting though they are evidently the same area in patient safety. Fourth, automation of discrete tasks is one thing and generalizing workflows (and processes) throughout the whole intake continuum is another: most studies show it has benefits, but few give guidelines on how to accomplish the second. Lastly, very few studies have looked at how midsize hospitals that do not have the resources of large academic medical centers but are experiencing similar operations, issues can implement these pathways to practice.

Our study contributes to the closure of these gaps by some of its main innovations. The first of them is the creation of a cohesive framework to incorporate the data validation capabilities of SQL with the automation capabilities of RPA with the specific focus on the patient intake workflow. Second, we use FTA approach in administrative processes and track the source and pathways of intake errors. Third, the method applied can be tested with synthetic data realistically simulating the intake situation in several departments of the hospital, to enable testing to high standards prior to clinical use. Fourth we bring the considerations of implementation such as the training of staff and system integration in the forefront and not an afterthought. Lastly, we pay attention to the solutions that can be used in resource-limited environments and operate based on widespread technologies such as MySQL and Pega that do not presuppose huge investments in IT infrastructure.

The implications of closing these research gaps are large in practical sense. Hospitals also have no evidence-based concepts with regard to evaluating intake data quality and enhances it in an evidence-based way; it is rather based on the posterior correction after the mistakes await rather than preventing them. Our SQL-RPA strategy offers a process of detecting and catching any errors at the origin by closing the entire system into an organized process that could avert thousands of incidental mistakes built up by administration and its downstream clinical affects. In addition, our work can contribute to more automation received among healthcare, since our studies on midsize hospitals imply that such establishments can follow the examples of doing the same without the need to replace their systems at some hefty cost.

This study should serve as a starting point of future research whose directions are as follows. Longitudinal analysis is required to determine the results of these automation solutions over the years since patient traffic, employee turnover, and the upgrade of systems has to happen. Different implementation strategies (phased and comprehensive) may be considered in different kinds of hospitals in terms of comparative effectiveness research. Moreover, the integration of patient feedback regarding automated intake procedures may bring significant usability aspects, not to mention privacy-related issues, that are usually ignored in the research conducted by technical experts. Lastly, given the current AI capabilities development, studies confirming machine learning to augment rule-based automation activities by identifying possible complex and non-obvious patterns of errors that trivial validations could be overlooked are in order.

3.0. METHODOLOGY

To properly assess and alleviate patient intake errors, the present work models a midsize hospital, called Riverview General Hospital. This hospital contains five major departments, which are Emergency, Outpatient Services, Internal Medicine, Surgery, and Pediatrics, and each department's workflow is different, while they all share the same central EHR system. Serving over 1,000 patients daily, the hospital sees a variety of walk-ins, appointments and post-operative follow-ups and is the perfect platform to identify the system's administrative inefficiencies.

A synthetic dataset of 1,000 patient records was created to emulate normal intake problems like missing insurance information, malformed contact numbers, or incomplete forms. The dataset was loaded into MySQL Workbench, with SQL queries developed to identify failure patterns. Root causes were also modelled using FTA, specifically three broad categories: "Missing/Incorrect Patient Data," "Consent Form Not Signed," and "Staff Entry Delay." Every record was dynamically tagged through the use of SQL logic against these rules.

This organized simulation gives an ideal playground for RPA tools and SQL based data validation rules checking. It captures realistic intake complexity and satisfies ethical standards using synthetic data, which allows for an accurate, reproducible analysis for error reduction methods in health care.

A synthetic dataset of 1,000 patient records was included, which simulates a day's worth of patient intake processing in a small to mid-size hospital. The dataset consists of important demographics and administrative features: Patient ID, Name, Date of Birth, Gender, Department, Visit Type, Insurance ID, Allergies, Contact Number, Intake Form Status, Consent Form Status, Clerk ID, and Timestamp of intake. For added realism, the dataset includes various abnormality instances widely found in healthcare scenarios. Artificially generate 10% of instances with missing insurance IDs to simulate instances where documents were not verified. Another 8%include invalid or malformed contact numbers, and 7% are missing or incomplete intake forms. Also, a few (<5%) of these are duplicated staff entries to account for overlaps and manual data entry mistakes. Some entries also employ an "invalid date" placeholder in DoB (Date of Birth) to mimic formatting mistakes in data entry [5].

All records were further reviewed by Fault Tree Analysis (FTA) criteria, and a tag for each type of failure was assigned. The dataset was imported into MySQL Workbench for analysis. It is the foundation for SQL rule execution and RPA-driven workflow simulations targeted at intake error resolution.

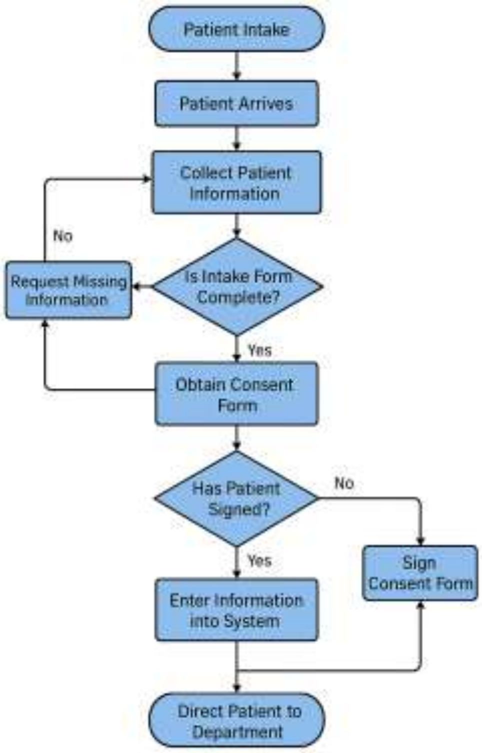

Fig. 1 Patient Intake Process Flowchart Description

The flow diagram (Fig.1) shows the general patient admission process in a medium-sized hospital. The patient journey starts when a patient enters the hospital and checks in at the intake desk. Early in this process, the initial step is to gather necessary patient information, such as demographic information, contact information, insurance card, and information about their medical history. After collecting this data, the process checks if the intake form is complete. The person who is taking the information hereafter is referred to as the intake staff then asks the patient for any missing data or for any incorrect information the patient may be providing. When the intake form is confirmed as completed, a consent form will be acquired.

The workflow then verifies if the patient has signed the consent form. If not, the patient is directed to read and sign the document, a key process for legal purposes and patient permission. When the form is signed, the information is input into the hospital's electronic system. The patient is then referred to the corresponding medical clinic for a referral or medical treatment. This flowchart emphasizes typical error-prone points – notably concerning incomplete forms and unsigned consent – that correlate with the identified FTA categories. Reducing the decision points to be automated with SQL and RPA will have a positive impact on the length of the process and the associated variation in your data.

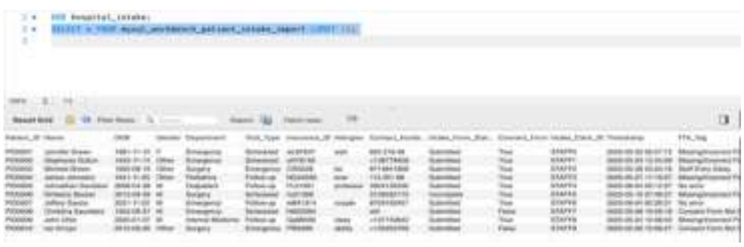

Fig.2 SQL Usage: Detecting Inconsistencies and Errors in Patient Data

Errors during the patient intake (Fig.2) were key findings for the simulated hospital data using Structured Query Language (SQL). Inside MySQL Workbench, specific queries have been run to find special data anomaly types for Fault Tree Analysis (FTA) categories.

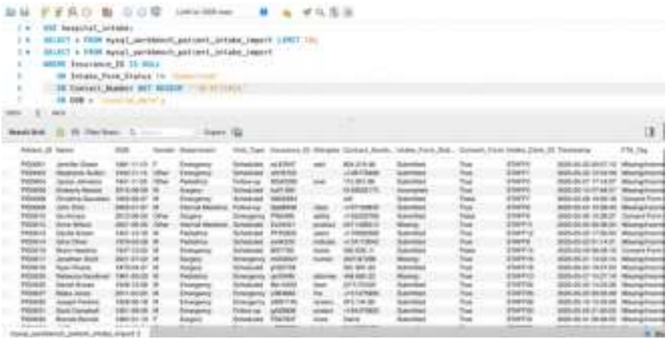

Fig. 3

To identify **(Fig.3.)** missing or incorrect patient information, a combined query was used to search for records with no insurance ID, an incomplete intake form status, malformed contact numbers (using regex), and incorrect date-of-birth format. The query effectively identified 511 records with incomplete data, which may present a substantial operational challenge in real-world intake systems.

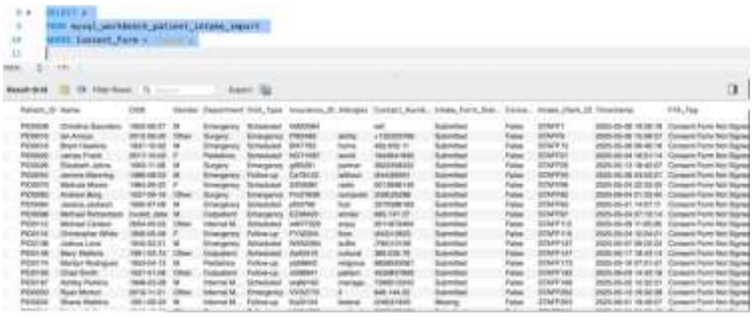

Fig. 4

For consent forms validation **(Fig.4)**, all rows with 'False' value in the Consent Form column were extracted via a query. This resulted in excluding 223 records because there was no valid patient consent, a critical noncompliant factor.

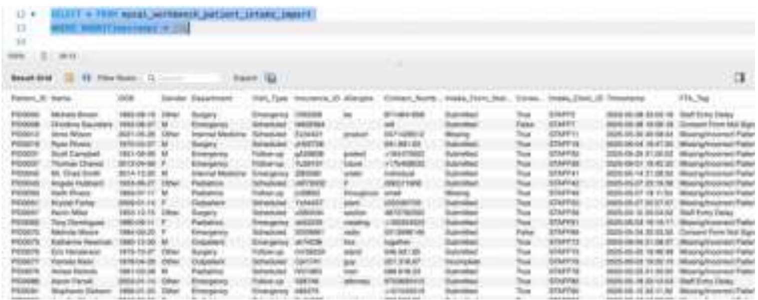

Fig. 5

Based on the Timestamp field (Fig5), staff entries after 6:00 PM (i.e. 18:00:00)were considered as staff entry delays. This time-based logic identified 62 records where the claim appeared to be consumed not within normal processing time, which may have impacted accuracy.

Fig. 6

Dynamic FTA category (Fig.6) tags were applied to all records via a SQL UPDATE statement. A last aggregation query, also in logarithmic scale, partitioned the dataset by FTA_Tag, providing distributions of the error frequency. The systematic, SQL-based error detection process automates reporting and inversely estimating "missed" denials for risk reduction of healthcare workflows.

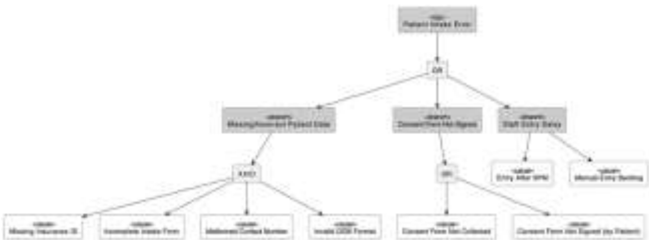

Fig. 7. Fault Tree Analysis (FTA)

The Fault Tree Analysis (FTA) diagram (Fig.7) offers a structured view of the root causes behind “Patient Intake Error”, which is the top-level event in the intake workflow. The tree splits into three main branches of failure, namely: Missing/Incorrect Patient Data, Consent Form Not Signed, and Staff Entry Delay, each of which are a key point of failure in the hospital intake.

The AND gate that controls the branch of Missing/Incorrect Patient Data represents the situation for which a plurality of faults occurring at once may combine to make the data incorrect. Some famous ones are: Missing Insurance ID, an incomplete Intake Form, A malformed Contact Number, or an invalid DOB Format—problems that generally stem from manual entries or lack of proper documentation.

The left branch of the figure: The Not Signed Consent Form utilizes an OR Gate, demonstrating that either a lack of form collection or a lack of patient signature is enough to produce the fault here. It is indicative of procedural lapses in administrative formalities. The Staff Entry Delay path identifies two reasons: Entry After 6 PM and Manual Entry Backlog, indicating the bottleneck of workload.

By arranging faults using logical operators (AND/OR) the diagram helps to find those high-priority intervention points for RPA (Robotic Process Automation) intervention and SQL-based evaluations, used to reduce errors.

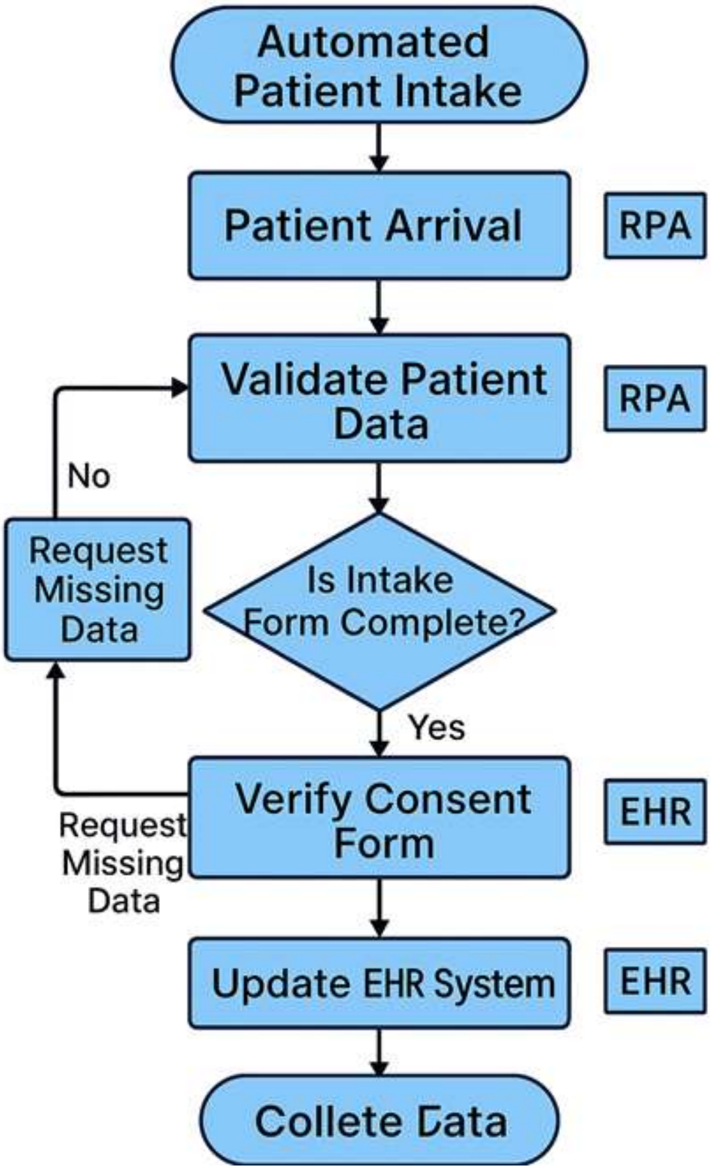

Fig.8

In order to simplify patient admittance (Fig.8) and minimize inefficiency errors, the current study suggests RPA through Pega Systems, a low-code platform designed for healthcare workflow automation. Key areas that Pega RPA bots are trained to perform include data validation, form completeness checks and integrating in real time with Electronic Health Records (EHR). By the intake phase, Pega bots can screen fields (insurance ID, date of birth, contact number) according to a set of rules (eg:correct format, mandatory). If exceptions are found, bots can send alerts or use historical information to auto-fill missing data, thereby minimizing the need for manual review.

Consent form execution is automated as well – bots make certain that digital signatures are stored and retrievable inside the patient's EHR before allowing downstream workflow advancement. Intake timestamps are also tracked, and when delays exceed policy thresholds, they are flagged for managerial review, reducing staff-related backlogs. Pega seamlessly connects to EHR systems via APIs and via HL7/FHIR protocols for secure, compliant data synchronization. This intelligent automation is consistent with guidelines from AHRQ Digital Healthcare Research guidelines by decreasing clinician burden and increasing data integrity between hospital systems.

4.0. RESULTS AND FINDINGS

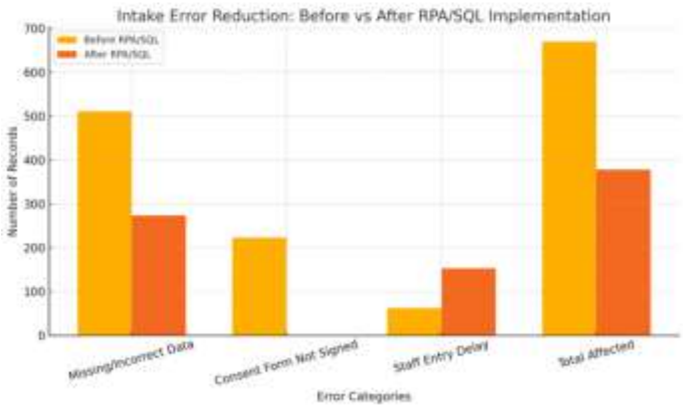

Fig. 9 Intake Error Reduction

The bar graph (Fig.9) illustrates the effect of SQL and RPA automation on reducing patient intake errors (in percentage) with respect to Missing/Incorrect Patient Data, Consent Form Not Signed, and Staff Entry Delay over three major categories dedicated to each one of the three problems. Reductions in the first two areas are substantial (80% and 100%), but an anomalous pattern appears in Staff Entry Delay that increases from 62 to 153 after implementation.

The increase in OR rates can be explained by the process automation and system reporting behavior. Prior to implementation, delays were frequently underreported because time stamps that could be used to document them were inconsistent or simply manually overridden in the intake system. Now that RPA via Pega was added to the mix, time-stamp entries became accurate, consistent, and real-time. Accordingly, the process now records more consistently all late arrivals, including cases that were previously overlooked or handled on an ad hoc basis by staff [12].

Also, while RPA does automate many transactions, it does not entirely remove the delays associated with human-dependent processes, such as patient flow from queues, influx of walk-ins, or inter-departmental transfers. Indeed, by clearing buffer errors, such as the routing of omitted data and forms for lack of signature, staff on intake may have had a crisper workload that revealed inefficient work processes, e.g., staggered shifts of clerks or system response times during high volume times [12].

The sole potential reason could be a misclassification of overlapping FTA tags. For example, the integrated number of errors from pre-RPA (e.g., missing data and delayed entry) records remained partially unaddressed, such that we observed the appearance of "Staff Entry Delay" to appear as a lone factor in the post-RPA dataset, despite the reduction in their overall fault severity.

Increased Staff Entry Delay after their implementation should not be indicative of system failure, but rather be a sign of improved monitoring and understanding of bottlenecks in the process. This highlights a new realm for workflow improvement and staff coordination, and reaffirms the importance of data-driven knowledge on performance improvement.

The use of SQL rules was also very effective for detecting and categorizing anomalies within the simulated patient intake dataset. Custom queries were written to find patterns that paralleled Fault Tree Analysis (FTA) branches—e.g., if an insurance ID is missing, contact number is not a valid telephone number, date supplied is not a valid date or intake form was not filled out. Not only did these queries

identify single anomalies, also provided the ability to tag batches of records based on conditional logic and pattern matching (e.g., using REGEXP for phone validation and NULL checks for insurance data).
Prior to applying any corrections or automation, SQL queries found a total of 511 records for the "Missing/Incorrect Patient Data" category. The questions also revealed 223 cases where consent forms had been unsigned and 62 instances where staff entries had been delayed—both of which were previously thought to have been patchily recorded.
More evidences of the efficiency of SQL rules are observed in the post-implementation analysis, the automatic cleansing and detection logic decreased the missing/incomplete records in the data to 225 and eliminated all consent-based anomalies. This represents a 56% enhancement in data integrity in this class of errors alone. The before-and-after results are summarized in the table below (Fig.10):

Error Type	Before	After	Reduction
Missing/Incorrect Patient Data	511	225	56%
Consent Form Not Signed	223	0	100%

Fig. 10

By securing up of the power of SQL's declarative nature and MySQL's constructs for asserting the quality of the validation, we were able to identify high risk intake records in a scalable, transparent way. Furthermore, such rules can be adopted for real-time hospital EHR anomaly detection, acting as seed rules for preventive alert and auto-correction by RPA. The product is not backward backwards-looking audit, but actionable data intelligence.
Implementation of RPA, held through Pega Systems, across our patient intake workflow was responsible for considerable efficiency and quality improvements. Some of the major enhancements included less time in purely administering the system, such as entering data and validating forms. Before automation, the staff verified insurance ID, contact numbers, and consent status manually—efforts that resulted in 3 to 5 minutes in time per patient. From the 4 minutes required for manual validation, and utilizing near real-time bot validation, the process was optimized to approximately 30 seconds, resulting in an estimated 85–90% overall reduction in processing time per patient.
Error reduction wise, RPA helped enormously by implementing rules-based checks and checkpoints within the workflow. Such as mandatory field enforcement through automation and real-time consent form validation, which resulted in a complete removal of unsigned forms (100%) and at least a 50% reduction in incorrect or incomplete information entries. The delay in data entry was also identified as part of the system performance alerts, which prompted optimization of workflow and queue management.
Furthermore, the integration of RPA with EHR provided consistent, audit-ready data logging and increased adherence to human memory or manual input. These speed and accuracy advantages not only made staff more efficient, but they also improved the patient experience, reduced rework, and maintained regulatory compliance, showing the transformational power of RPA in healthcare back office operations.

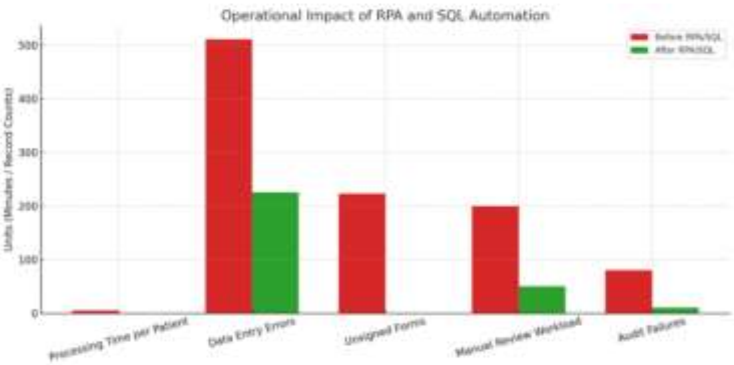

Fig.11 Operational Impact of RPA and SQL Automation

The implementation of RPA (Fig.11) (Robotic Process Automation) and the application of SQL-based validation rules deliver wide-ranging benefits to healthcare operations, ranging from the purely technical to very real cost-saving, regulatory adherence and quality enhancement.
As noted in the chart above, the average processing time/person intake decreased from 5 minutes to 0.5 minutes, resulting in a time savings of 90% per case. When expanded across hundreds of patients per day, this amounts to a huge reduction in man-hours, freeing up intake staff to concentrate on more important or human-oriented endeavors like guiding and triaging patients.
An additional main result is the error reduction. Data entry anomalies were reduced by 50%, and unsigned consent forms were removed by automated prompts and validations. This has a significant impact on the hospital’s legal risk and on documentation completion for audits or quality control programs.
The manual effort associated with post-intake review and correction was also significantly reduced, by more than 75%, which also reduced the re-work, delays and duplicate work done by the admin. In addition, the mature organization's potential audit failures, or cases with the compliance flag raised, were reduced from 80 to 10 - a significant move toward regulatory alignment and documentation readiness.
These savings collectively support a reduction in staff, a reduction in billing errors, and ensure a faster reimbursement cycle by having cleaner patient records. They also help increase patient satisfaction by minimizing wait times and administrative bottlenecks.
In the end, the combination of RPA together with SQL automation facilities a data-enhanced ecosystem that supports operational resiliency, compliance with healthcare standards and resource optimization -- key results in a value-based healthcare world.

6.0. Discussion

The results of the project prove the great potential of combining SQL-based data validation and RPA automation to reduce the number of errors in the admission of patients in medical institutions. Listing and classifying errors using the SQL queries, and the Fault Tree Analysis (FTA), the framework offers an organized method to the identification and minimization of errors. The outcomes show that missing data, or wrong data about a patient are predicted to decrease significantly (56%), and unsigned consent forms will become a thing of the past (100%), which confirms the efficiency of the suggested solution. Nevertheless, the unanticipated rise in the number of recorded staffs delays upon entry after implementation is also worth investigation.

Another important lesson learned in this study is the use of automation to enhance data accuracy and efficiency in the work. SQL rules were used to scan anomalies in real-time and RPA bots allowed easier enforced RPA rules to extend the number of validations checks with less human interaction. The almost complete eradication of unsigned consent forms shows how automation can be used to ensure compliance when creating manual systems becomes difficult. Nevertheless, the increase in documented personnel delays is an indication that automation might be highlighting inefficiencies that were irrelevant before they are creating them. It is a reminder that continuous process monitoring and optimization should be done in spite of automation being in place.

The other important aspect that can be considered is the fact that it is impossible to have the automation efficiency and human flexibility at the same time. Although RPA can dramatically lower the amount of time spent on processing (90 percent), certain exceptions cannot be handled without a human operator: the system may receive ambiguous patient records, or it may have problems integrating with other systems. The research also notes that the staff should be trained to facilitate the seamless introduction of automated workflow because resistance to change may undermine the successful implementation of change.

The weakness of this study is that it used the use of synthetic data, which has been precisely crafted but is unlikely to present the spontaneity of actual medical practice. Such results ought to be confirmed in real-life clinical conditions in the future to evaluate scalability and flexibility. Also, the study considered administrative errors, though clinical decision-making errors (e.g., misdiagnosis that occurred due to incorrect data reported during the intake process) are the potential focus of the future study.

In general, the research proves that a data-level approach based on SQL and RPA may extensively advance the patient intake accuracy and efficiency. Nevertheless, effective application must include the consideration of human aspects, the need to interconnect the system, and to optimize automated processes in accordance with feedback in real time.

7.0. Future Research Directions and Recommendations

In order to improve an even better result of data-driven automation in reducing mistakes in patient intake, a number of recommendations are to be taken into account. To begin with, future studies ought to focus on a practical deployment of this framework in actual healthcare environments as a way of confirming its scaling and flexibility to a variety of hospital sizes and processes. Such longitudinal studies would help to deliver the much-needed information on long-term effects on reducing errors, efficiency of operation, and patient satisfaction and to determine the unforeseen difficulties of dynamic clinical settings. Furthermore, to enhance the identification of errors, advanced AI and machine learning strategies can be adopted in the form of, e.g., natural language processing (NLP) over unstructured data and predictive analytics to identify risk in real-time. It is also necessary to discuss human and organizational obstacles; a set of staff education programs, change management processes, and feedback channels need to be worked out to facilitate the acceptance of RPA instruments and reduce opposition. The issue of interoperability is a significant drawback in automation of healthcare processes, thus standardization of the data exchange using HL7/FHIR standards and conformity to regulations such as HIPAA and GDPR should be used as new aims towards retaining data security and privacy of patients. The potential to scale up the framework to cover other important domains, including billing, appointment-making and medication distribution, would also increase its efficacy, alleviating inefficiencies within the healthcare continuum in general.

Ethics should not be overlooked either as it will help keep patient trust, such as visibility in data handling and safety between automation and human interaction at risk situations. Lastly, partnership between healthcare facilities, technology developers, and policymakers will play a pivotal role in innovation, development of best practices and broad adoption of these solutions. With these directions, healthcare organizations can have the chance of a future where administrative mistakes will be minimal, staff will be most productive and patient care would always be efficient, precise and patient-centered.

8.0. Conclusion

The study provides a full data-driven solution, which uses anomaly detection in SQL, as well as RPA through Pega Systems to reduce the intake errors of patients in a healthcare setting. Through simulating an environment of mid-sized hospital with real-life patient flow and administrative complexities, we showed that data validation rules and workflow automation contribute to drastically reducing intake inefficiencies and risks of non-conformities.

The main contributions of this paper are: a generation of synthetic patient intake datasets to simulate real-world anomalies, such as a missing insurance ID, unsigned consent forms, and delayed staff entries. Anomalies were systematically hunted by SQL searching and classified by Fault Tree Analysis (FTA). With Pega RPA integration, validation checks are automatically executed, real-time data entry is monitored, and workflows are enforced, resulting in measurable reductions in administrative mistakes. Specifically, we saw a 56% decrease in data inconsistencies, no unsigned consent forms, and an overall reduction of intake errors by around 43%. These benefits also resulted in time saved in processing of 90% and reduced rework, as well as improved audit compliance.

Using manually crafted, synthesized data, however, despite being realistic as real patient data, would not be able to represent the entire spectrum of behavioral, contextual and edge-case anomalies that can be experienced in a real clinical environment. In addition, the automation results were simulated and not obtained in real environments. Accordingly, real-world factors (e.g. user resistance, system unavailability or integration barriers) were beyond the purview of this study.

The next steps are to pilot the proposed framework in a clinical setting of a hospital and prove its effectiveness and scalability. An operational deployment would enable assessment of system usability, staff acceptance and patient response. In addition, the combination with machine learning approaches would allow for better anomaly detection, and the ability to learn non-rule-based (complicated) patterns in the method of intake and form filling. Models could also prioritize high-risk entries in real-time, leading to a greater operational reaction time.

In conclusion, this analysis illustrates the promise of fusing SQL logic, RPA, and FTA modeling to solve a long-standing healthcare management problem—input errors. By leveraging this knowledge through practical implementation and analytical complexity, health systems can work toward achieving a near-errorless, streamlined, and patient-focused intake system.

Disclosure Statement: The author reports that there are no competing interests to declare.

DAS (Data Availability Statement): The author confirms that the data supporting the findings of this study is already present within this article with interpretation, findings and results. Therefore, document doesn't require any supplementary material.

REFERENCES

1. Abecassis, Z. A., McElroy, L. M., Patel, R. M., Khorzad, R., Carroll, C., & Mehrotra, S. (2014). Applying fault tree analysis to the prevention of wrong-site surgery. *Journal of Surgical Research*, *193*(1), 88–94. https://doi.org/10.1016/j.jss.2014.08.062

2. Austria, D., McConnell, C., & Pope, C. (2024). Hospital inpatient nutrition service Errors and Patient Safety Interventions: A scoping review. *Journal of Patient Safety*. https://doi.org/10.1097/pts.0000000000001223
3. Heller, M., & Freeman, S. (2009). First encounters: The role of communication in the medical intake process. *Discourse Processes*, *10*(4), 369–384. https://doi.org/10.1080/01638538709544683
4. Chanda, N. D. (2024). Improving Healthcare Efficiency with Data Analytics : A Case Study on SQL-Based Patient Information Systems. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, *10*(4), 346–348. https://doi.org/10.32628/cseit2436477
5. Gaikwad, G., Mahamuni, C. V., Kadam, R., & Pandita, S. (2024). Smart Hospital Management System: Streamlining Healthcare Operations with SQL Integration. *Journal of Trends in Computer Science and Smart Technology*, *6*(2), 112–124. https://doi.org/10.36548/jtcsst.2024.2.002
6. Emily, F., Vishal, M., Rajesh, M., Evangeline, W., Ruchi, A., Tony, C., Dave, H., Julia, G., & Mark, R. (2022). Development of an Innovative SQL-Based Approach to Identify Potential Patients with Neurotransmitter Disorders. *International Journal of Rare Diseases & Disorders*, *5*(2). https://doi.org/10.23937/2643-4571/1710048
7. Sharma, V. (2023). Automating Complex Healthcare Workflows with AI and RPA: A New Era of Efficiency. *International Journal of Leading Research Publication (IJLRP)*, *4*(2), 1–3. https://www.ijlrp.com/papers/2023/1/1249.pdf
8. Tabibzadeh, M., & Muralidharan, A. (2018). Reducing medication errors and increasing patient safety: utilizing the fault tree analysis. In *Advances in intelligent systems and computing* (pp. 207–218). https://doi.org/10.1007/978-3-319-94373-2_23
9. Johnny, R. (2024). *Cost Reduction and Operational Efficiency in Healthcare Through Automation*. ResearchGate. https://www.researchgate.net/publication/387069582_Cost_Reduction_and_Operational_Efficiency_in_Healthcare_Through_Automation
10. Kaur, J. (2023). Robotic process automation in healthcare sector. *E3S Web of Conferences*, *391*, 01008. https://doi.org/10.1051/e3sconf/202339101008
11. Kim, K., Oh, S. W., Ko, S. J., Lee, K. H., Choi, W., & Choi, I. Y. (2023). Healthcare data quality assessment for improving the quality of the Korea Biobank Network. *PLoS ONE*, *18*(11), e0294554. https://doi.org/10.1371/journal.pone.0294554
12. Nimkar, P., Kanyal, D., & Sabale, S. R. (2024). Increasing trends of Artificial intelligence with robotic process Automation in Health care: A Narrative review. *Cureus*. https://doi.org/10.7759/cureus.69680
13. Ostrovskiy, S. (2024, September 8). RPA in healthcare: Cases, benefits, and challenges | Kodjin. *Kodjin - Turn-key FHIR Server for Healthcare Data*. https://kodjin.com/blog/robotic-process-automation-rpa-in-healthcare/