

MedPeer Publisher

Abbreviated Key Title: MedPeer

ISSN : 3066-2737

homepage: <https://www.medpeerpublishers.com>

FAILURE MODE, EFFECTS AND CRITICALITY ANALYSIS OF THE MEDICATION STOCK CIRCUIT IN A PEDIATRIC HOSPITAL PHARMACY IN MOROCCO: A BEFORE–AFTER STUDY WITH UNIVARIATE AND MULTIVARIABLE ANALYSIS

DOI: 10.70780/medpeer.000QGYN

AUTHOR AND AFFILIATION

EDDAHOU MI Y^{1,2}, BENNANI I³, BELACHHAB I², BOUATIA M^{1,2}¹ Faculty of Medicine and Pharmacy of Rabat, Mohammed V University, Rabat, Morocco² Children's Hospital of Rabat, CHU Ibn Sina, Rabat, Morocco³ Faculty of Medicine, Pharmacy and Dental Medicine of Fez, Fez, MoroccoCorresponding author: EDDAHOU MI Y (eddahoumiyassine@gmail.com).

ABSTRACT

Background: Reception, storage and ordering of medicines in the hospital pharmacy are high-risk steps, yet quantitative evaluations of corrective actions are rare in Moroccan pediatric hospitals.

Objective: To identify and prioritize failure modes in the medication stock circuit using failure mode, effects and criticality analysis (FMECA), to describe the change in estimated criticality after preventive actions, and to explore factors associated with the magnitude of this change.

Methods: Uncontrolled before–after study (March–June 2025) in the pharmacy of the Children's Hospital of Rabat. Failure modes were grouped into 17 operations (reception 8, storage 3, ordering 6) and scored by consensus for detectability, occurrence and severity (1–10; criticality = $D \times O \times S$; threshold 100) before and after preventive actions. Analyses comprised paired Wilcoxon and exact McNemar tests, univariate analysis and exploratory multivariable linear regression of the relative reduction in criticality.

Results: Before actions, 14/17 operations (82.4%) were at or above the threshold (median criticality 168 [IQR 120–210]); none was afterwards (median 48 [36–63]; $p < 0.001$). Estimated criticality was reduced by a mean 69.1% (95% CI 64.3–73.9), mainly through lower occurrence (50.2% of the log-reduction) and better detectability (46.0%) rather than severity (3.9%). The reduction was larger in ordering than in reception (median 80.0% vs 62.4%; crude difference +15.6 points, 95% CI 5.0–26.2); the adjusted estimate (+17.1 points, 95% CI 4.1–30.2) is imprecise and exploratory. Sensitivity analyses gave consistent results.

Conclusions: Estimated FMECA criticality was markedly lower after the implementation of preventive actions, especially in the ordering process. This uncontrolled design with consensus-based scoring cannot demonstrate a causal effect or improved patient safety; findings are descriptive and hypothesis-generating and warrant confirmation with independent scoring and objective indicators.

KEYWORDS

Failure mode and effects analysis, FMECA, Hospital pharmacy, Medication safety, Stock management, Risk management, Pediatric hospital, Morocco

MAIN ARTICLE

INTRODUCTION

Medication-related harm remains one of the leading causes of avoidable injury in health care, and the safety of the medication-use system has been identified as a global priority [1,2]. In hospitals this system starts long before the bedside: medicines must be ordered, received, stored, prepared for distribution and traced by the hospital pharmacy. Failures at these upstream stages, such as late or incomplete deliveries, confusion between similar products, loss of traceability, inadequate storage conditions or stock-outs, can interrupt treatment or allow non-conforming products to reach the wards. Children are particularly exposed to the consequences of such failures because of weight-based dosing and narrow dosing margins.

Quality risk management frameworks such as ICH Q9(R1) and ISO 31000 encourage a prospective and structured approach to risk [3,4]. Failure mode and effects analysis (FMEA), and its extension with a criticality analysis (FMECA), is among the most widely used tools in this respect. Adapted to health care by the United States Department of Veterans Affairs [5], it breaks a process into steps, identifies how each step may fail, and scores each failure mode for severity, occurrence and detectability; the product of these scores yields a criticality index used to rank priorities. A systematic review of its application to the medication-use process reported that participants considered FMEA an effective group activity for identifying system errors, although time-consuming and subjective [6]. In hospital pharmacy, FMECA has been used, for example, to anticipate failures during the relocation of a pharmacy [7] and of hospital medication processes [8]; in the latter study, average criticality indices fell by 76–80% after mitigation measures in the relocation scenario [8].

Many FMECA reports in this field are, however, essentially descriptive: they list failure modes and criticality scores but seldom test statistically whether corrective actions changed criticality, or which characteristics predict the magnitude of improvement. In addition, the validity and reliability of FMEA outputs have been questioned, in part because scores are ordinal and their product is sensitive to small scoring differences [9,10]. To our knowledge, quantitative before–after evaluations of FMECA in the stock circuit of a Moroccan pediatric hospital pharmacy have not been reported.

The objectives of this study were (i) to identify and prioritize the failure modes of the reception, storage and ordering processes of the medication stock circuit in a pediatric hospital pharmacy; (ii) to quantify the change in criticality after preventive actions; and (iii) to explore, through univariate and multivariable analyses, the factors associated with the magnitude of risk reduction.

METHODS

Study design and setting

We conducted a prospective, single-center, before–after quality improvement study from March to June 2025 in the pharmacy of the Children’s Hospital of Rabat, which belongs to the Ibn Sina University Hospital Center (CHU Ibn Sina), Morocco. An FMECA approach was applied to the part of the medication circuit that concerns stock management: (i) reception and verification of deliveries, (ii) storage, and (iii) preparation and distribution of internal orders (hereafter “ordering”). Each process was described by a flow diagram built with the pharmacy team.

Participants and data sources

Participants were selected by purposive (non-probabilistic) sampling and comprised pharmacists in charge of stock management and pharmacy technicians directly involved in reception, storage, distribution or control, with at least six months of experience in their current post. Trainees, newly assigned staff and staff from other pharmacy units were excluded. The final sample comprised 6 pharmacists and 12 pharmacy technicians ($n = 18$). Data were collected by field observation of the circuit, interviews and a structured survey using three process-specific conformity grids (reception, storage and ordering), complemented by documentary review (internal stock-management procedures, reception checklist, medicine storage-conditions sheet, non-conformity reports and delivery notes). Potential causes of failure were explored with an Ishikawa diagram.

FMECA procedure

For each process, failure modes, their effects and probable causes were listed and grouped by operation (for example, delivery schedule, documentation, storage area). Each operation was scored by consensus among the evaluators, namely the pharmacists and pharmacy technicians involved in the circuit, under the supervision of the head of the pharmacy department on three 1–10 scales: severity (S; impact on patient safety, product quality or continuity of care), occurrence (O; probability that the failure occurs) and detectability (D; ability to detect the

failure before it has consequences, a higher score indicating lower detectability). The criticality index was computed as $C = D \times O \times S$ (possible range 1–1000), and a threshold of 100 was set, at or above which an operation was considered critical and requiring priority action. Corrective and preventive actions were proposed for every operation and implemented by the pharmacy team (still in progress for one operation at the time of analysis); the timing and fidelity of implementation of each action were not systematically documented. After implementation, operations were re-scored on the same scales, within the study period (March–June 2025), to obtain a new criticality index (C'). Scoring was performed by staff involved in the improvement process, without independent or blinded assessment.

Failure modes were identified and grouped into 17 operation-level items (reception 8, storage 3, ordering 6), which constituted the unit of analysis. Conformity of the evaluation criteria (conforming or non-conforming) was also recorded for each process.

Statistical analysis

Continuous variables are described by median [interquartile range, IQR], mean $\pm$ standard deviation (SD) and range, and categorical variables by number (percentage). The primary outcome was the criticality index before and after preventive actions, compared with the Wilcoxon signed-rank test [11]; D, O and S were compared in the same way. The proportion of operations at or above the threshold was compared with the exact McNemar test [12]. The relative reduction in criticality (RR, %) was defined as $(C - C')/C \times 100$, and 95% confidence intervals were obtained by percentile bootstrap with 10,000 resamples [13]. Because $\ln(C/C')$ equals the sum of $\ln(D/D')$, $\ln(O/O')$ and $\ln(S/S')$, the contribution of each parameter to the overall log-reduction in criticality was computed exactly.

Univariate analysis explored the factors associated with RR: process (reception, storage, ordering) with the Kruskal–Wallis test [14], and baseline detectability, occurrence, severity and criticality with Spearman's rank correlation; crude linear regression coefficients (β , in percentage points) with 95% confidence intervals were also estimated. Conformity across processes was compared with the exact Fisher–Freeman–Halton test [15].

Variables with $p < 0.20$ in the univariate analysis (process, baseline detectability, baseline severity and baseline criticality) were considered for multivariable linear regression with RR as the dependent variable. Because the criticality index is by construction the product of D, O and S, it was never entered together with its components, and because only 17 operations were available, models were limited to four or five parameters. Model 1 included process and

baseline criticality (log₂-transformed and centered); Model 2 included process, baseline detectability and baseline severity. Coefficients are reported with 95% confidence intervals and p-values based on heteroscedasticity-consistent (HC3) standard errors [16]. Multicollinearity (variance inflation factor), normality of residuals (Shapiro–Wilk), heteroscedasticity (Breusch–Pagan) and influential observations (Cook’s distance) were checked, and a bootstrap with 5,000 resamples was used to verify the confidence interval of the main coefficient. A model on the logarithmic scale, $\ln(C'/C)$, was fitted as a complementary analysis. The sample size is well below that usually recommended for regression analysis [17]. In addition, operations within a process share staff, organizational conditions and corrective actions, so observations are not independent; with only three processes, cluster-robust variance estimators are unreliable, and process was therefore entered as a fixed effect. The resulting p-values and confidence intervals are probably too optimistic, and the multivariable analyses are presented as descriptive and hypothesis-generating.

To ensure internal consistency, criticality indices were recalculated from the component scores for all operations in the primary analysis. Three sensitivity analyses were performed after verification of the source tables: (i) exclusion of the packaging-integrity item, whose re-scoring was atypical and for which no supporting evidence was available (severity decreased from 9 to 4 and the detectability score increased from 3 to 4, although preventive actions are expected to act mainly on occurrence and detectability); (ii) use of the post-action criticality of handling staff recorded in the original evaluation table ($C' = 63$) instead of the value recalculated from its scores ($4 \times 3 \times 7 = 84$) used in the primary analysis; and (iii) exclusion of both items. All tests were two-sided with $\alpha = 0.05$; no adjustment for multiplicity was applied and p-values should be interpreted as exploratory. Analyses were performed in Python 3 with SciPy [18] and statsmodels [19].

RESULTS

Failure modes and process conformity

The failure modes were grouped into 17 operation-level items scored by FMECA (Table 2). Conformity of the evaluation criteria was 6/10 for reception, 7/10 for storage and 4/10 for ordering (Table 1); ordering therefore showed the lowest conformity, although the difference between processes was not statistically significant (exact $p = 0.531$).

Criticality before and after preventive actions

Before preventive actions, criticality ranged from 48 to 294 (median 168 [IQR 120–210]) and 14 of 17 operations (82.4%) were at or above the threshold of 100: 6/8 in reception, 2/3 in storage and 6/6 in ordering (Table 1, Figure 1). The highest criticality indices concerned the order-launch schedule (294), reporting of anomalies during order control (280), reception documentation (240) and the preparation/control schedule (240) (Table 2). Baseline criticality differed by process (Kruskal–Wallis $p = 0.019$), being highest in ordering (median 225 [210–270]) compared with reception (130 [105–162]) and storage (147 [98–161]).

After preventive actions, no operation remained at or above the threshold (14 to 0; exact McNemar $p < 0.001$): the maximum criticality was 96 (documentation) and the median was 48 [36–63] (Table 2, Figure 1). Detectability and occurrence scores decreased, whereas severity scores were essentially unchanged (Table 3). The median paired reduction in criticality was 112 points (Wilcoxon $p < 0.001$; effect size $r \approx 0.88$). The estimated relative reduction averaged 69.1% (SD 10.3; 95% CI 64.3–73.9; median 70.0% [IQR 60–79]) and ranged from 50.0% (order correspondence) to 85.7% (order-form signature). Summed over the 17 operations, criticality fell from 2,885 to 846 (–70.7%), and the geometric mean ratio C'/C was 0.29 (95% CI 0.24–0.35).

Occurrence accounted for 50.2% of the total log-reduction in criticality, detectability for 46.0% and severity for only 3.9% (Figure 2B). Occurrence decreased in all 17 operations, whereas severity was modified in a single operation (packaging integrity, 9 to 4).

Univariate analysis

The relative reduction in criticality differed across processes (Kruskal–Wallis $H = 7.29$, $p = 0.026$): median 62.4% [58.8–67.5] in reception, 71.4% [67.7–73.2] in storage and 80.0% [78.9–82.1] in ordering (Table 4, Figure 2A). In unadjusted pairwise comparisons (exact Mann–Whitney tests), the difference between ordering and reception was $p = 0.020$, whereas reception versus storage ($p = 0.194$) and storage versus ordering ($p = 0.167$) did not differ. In crude regression, ordering was associated with a 15.6-point greater reduction than reception (95% CI 5.0–26.2; $p = 0.007$). Baseline detectability scores were higher (that is, detection was poorer) in ordering (Kruskal–Wallis $p = 0.016$). None of the baseline parameters was significantly associated with the relative reduction (detectability: $\rho = 0.37$, $p = 0.142$; severity: $\rho = -0.37$, $p = 0.149$; criticality: $\rho = 0.34$, $p = 0.177$; occurrence: $\rho = 0.22$, $p = 0.402$). As expected arithmetically, the relative reduction was positively correlated with the reductions in

occurrence ($p = 0.56$, $p = 0.020$) and detectability scores ($p = 0.62$, $p = 0.008$); these variables were not entered in the multivariable models.

Multivariable analysis (exploratory)

Given the small number of observations and the dependence of operations within processes, the following models are exploratory. In Model 1 (process and baseline criticality; $R^2 = 0.50$, adjusted $R^2 = 0.39$), the estimated difference between ordering and reception was +17.1 points (95% CI 4.1–30.2; $p = 0.014$; bootstrap 95% CI 2.7–25.6), storage did not differ from reception (+7.4 points; 95% CI –3.6 to 18.3; $p = 0.172$) and baseline criticality was not associated with the relative reduction (–1.9 points per doubling; $p = 0.584$) (Table 5). The joint test for process gave $p = 0.032$. On the logarithmic scale, the ratio C'/C in ordering was 0.53 times that in reception (95% CI 0.33–0.84; $p = 0.011$). Because observations within a process are not independent, these p -values and intervals are probably too optimistic and should not be read as evidence of an independent effect of process.

In Model 2 (process, baseline detectability and baseline severity; $R^2 = 0.55$, adjusted $R^2 = 0.41$), the estimate for ordering was similar in size but imprecise (+15.6 points; 95% CI –2.7 to 33.8; $p = 0.088$); neither baseline detectability ($p = 0.837$) nor baseline severity ($p = 0.255$) was associated with the relative reduction, and the joint test for process gave $p = 0.134$. Variance inflation factors were ≤ 1.8 (Model 1) and ≤ 2.3 (Model 2); residuals did not depart markedly from normality (Shapiro–Wilk $p = 0.066$) or from constant variance (Breusch–Pagan $p = 0.758$), and the largest Cook’s distance was 0.31 (handling staff).

Sensitivity analyses

Results were consistent (Wilcoxon $p < 0.001$ in all scenarios). After excluding the packaging-integrity item ($n = 16$), the median relative reduction was 70.7% and the adjusted difference between ordering and reception was +17.2 points (95% CI 4.2–30.1; $p = 0.014$). Using the originally recorded C' for handling staff (63), the mean relative reduction was 69.7% and the difference was +19.0 points (95% CI 9.2–28.8; $p = 0.001$). After excluding both items ($n = 15$), the median relative reduction was 71.4% and the difference was +21.3 points (95% CI 13.5–29.2; $p < 0.001$).

DISCUSSION

In this uncontrolled before–after FMECA of the stock circuit of a pediatric hospital pharmacy, more than four in five operations (14/17) were at or above the criticality threshold at baseline, and none was after the implementation of a set of largely low-cost organizational actions; estimated criticality was on average about 70% lower. This reduction reflected lower occurrence and detectability scores rather than severity, and it was larger in the ordering process than in reception, a difference that we regard as hypothesis-generating.

The ordering process combined the highest baseline criticality, the lowest conformity of audit criteria (4/10) and the poorest detectability. Its dominant failure modes (late launch of orders, delayed preparation and control, unreported anomalies, missing signatures and insufficient handling staff) are organizational rather than technical and point to the absence of a forecast calendar, undefined priorities and a workload exceeding capacity. In reception, the leading failure modes (delivery delays, incomplete delivery documentation and labeling errors) depend largely on suppliers, and in storage the main contributors were insufficient space and inadequate shelving, which are structural constraints. Logistics-related failures also had the highest criticality indices in an FMECA of a hospital pharmacy relocation [7].

The magnitude of the change (−70.7% in summed criticality) lies between the reductions reported after mitigation measures in an FMECA of hospital relocation (−76% to −80%) and of emergency evacuation (−39% to −52%) [8]. Because settings, scales and processes differ, this comparison is indicative only. These figures concern estimated FMECA criticality, not measured medication-safety outcomes. That severity remained virtually unchanged is expected: preventive actions such as order calendars with reminders, double-checking, written procedures and training act on how often a failure occurs and how likely it is to be caught, not on the harm it would cause if it reached the patient. Residual criticality is therefore bounded by severity. Operations with high severity scores (8–9) retained the highest post-action indices (for example, documentation, 96, and computer data entry, 72), which supports continuous monitoring of these items and, where feasible, designs that make failures less likely to occur or easier to detect, such as electronic checks or contractual supplier requirements.

The multivariable models suggest that the larger relative reduction observed in ordering was not simply a function of its higher baseline criticality, since baseline criticality was not associated with the relative reduction once process was accounted for. This does not exclude regression to the mean in a design without a control group. A possible but untested explanation is that ordering actions (calendars, reminders, fixed time slots, mandatory signature, staffing

adjustments) are under the direct control of the pharmacy, whereas reception depends on external suppliers whose behavior can be influenced only indirectly. Given 17 observations nested within three processes, the estimates are imprecise, and the adjusted association should not be interpreted as evidence of an independent predictor: after additional adjustment for baseline detectability and severity, the estimate was similar in size but no longer distinguishable from zero.

Several limitations should be acknowledged. First, the uncontrolled before–after design cannot establish that the implemented actions caused the observed change. Baseline and post-action scores were agreed by consensus among staff involved in the improvement process, without blinded or independent assessment, so expectation bias is possible; the timing, fidelity and individual effectiveness of the actions were not documented in detail, one action was still in progress, and no objective outcomes (stock-outs, expired or non-conforming products, documented non-conformities, errors reaching patients) were collected. The reported reductions therefore reflect changes in estimated criticality and not demonstrated improvements in medication safety. Second, FMECA scores are semi-quantitative and subjective; the multiplicative index treats ordinal scores as if they were ratio-scaled and is sensitive to small scoring differences, and the validity of FMEA outputs has been questioned [9,10]. The 2023 revision of ICH Q9 explicitly addresses the management of subjectivity in risk assessment [3]. We used structured grids, documentary sources and consensus scoring, but inter-rater reliability was not measured. Third, the unit of analysis was the operation ($n = 17$, nested within three processes), which limits power, creates a risk of overfitting and unstable coefficients, and violates independence between observations; the multivariable models are exploratory, p -values are not adjusted for multiplicity, and the results should not be regarded as confirmatory [17]. A regression of C on D, O and S was deliberately not performed, since C is their product. Fourth, two items of the source tables required verification: the post-action criticality recorded for handling staff (63) was inconsistent with its scores (84, used in the primary analysis), and the decrease in severity for packaging integrity (9 to 4) is not supported by evidence available to us, since preventive actions are expected to act on occurrence and detectability. Both were examined in sensitivity analyses without change in the overall conclusions. Finally, this was a single-center study with a short follow-up.

From a practical standpoint, these findings support prioritizing the ordering process in the quality plan of the pharmacy, maintaining a cycle of FMECA re-evaluation, and pairing FMECA scores with objective indicators such as compliance with delivery lead times, rate of

documentation non-conformity, number of stock-outs and inventory discrepancies. Future multicenter studies with independent, blinded scorers and objectively measured operational outcomes are needed to confirm these results and to model the determinants of improvement more robustly.

CONCLUSION

FMECA applied to the reception, storage and ordering of medicines in a pediatric hospital pharmacy identified a high proportion of critical operations (82%), concentrated in the ordering process, and showed an average 69% lower estimated criticality after the implementation of organizational actions, mainly through lower occurrence and better detectability scores. This uncontrolled, consensus-based before–after study cannot demonstrate that the actions caused the change or that medication safety improved, and the multivariable findings are exploratory. Independent scoring, objective indicators and multicenter confirmation are warranted.

FIGURES:

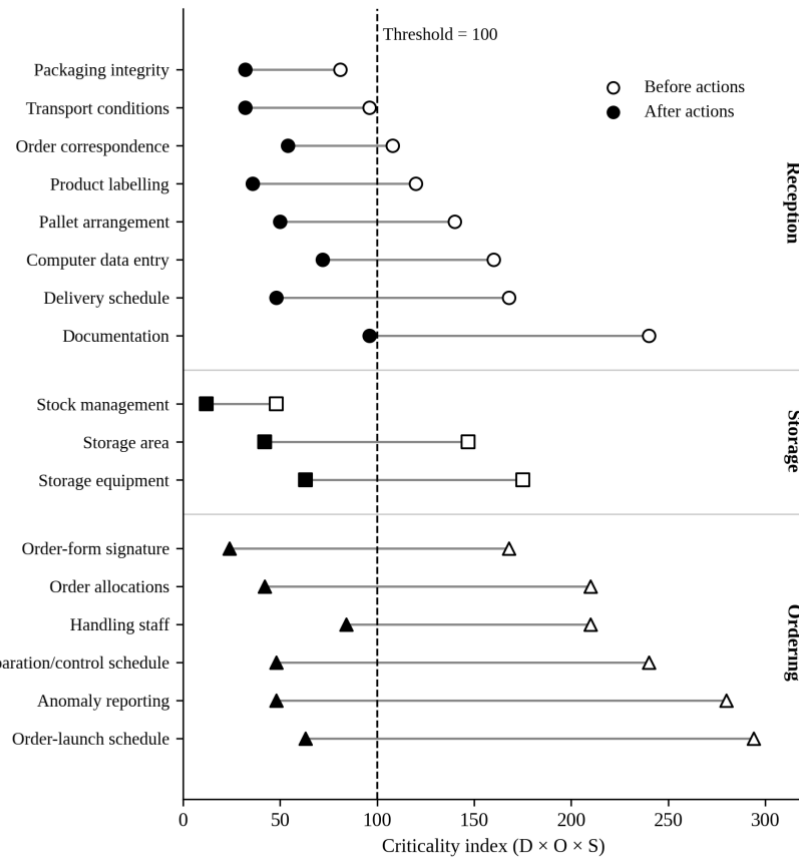


Figure 1: Criticality index ($D \times O \times S$) of the 17 evaluated operations before (open symbols) and after (filled symbols) preventive actions, by process. The dashed line indicates the criticality threshold (100). D, detectability; O, occurrence; S, severity.

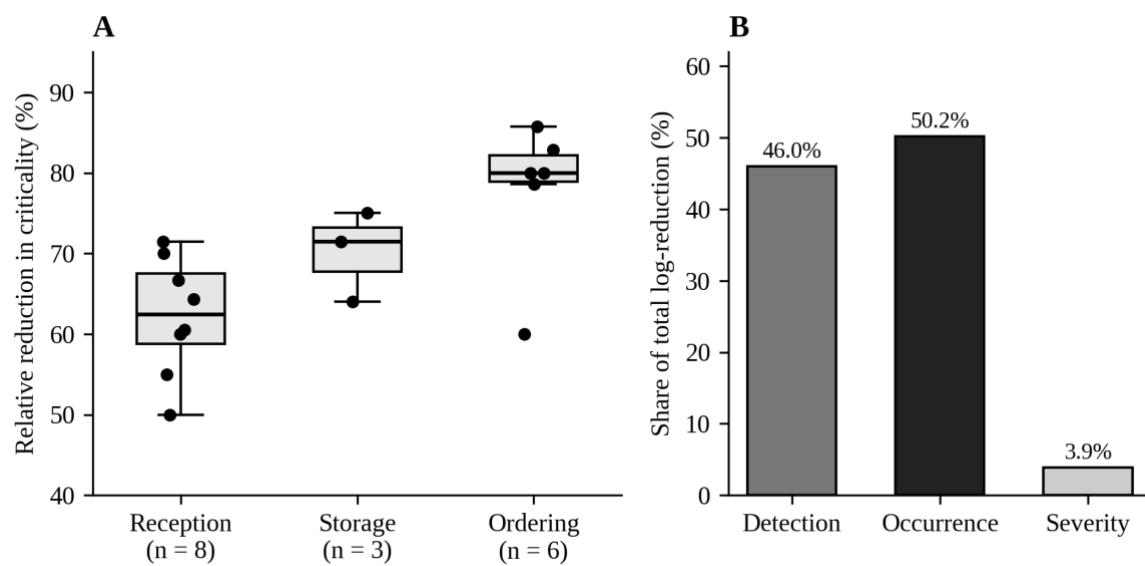


Figure 2: (A) Relative reduction in criticality by process (box: median and interquartile range; dots: individual operations). (B) Share of the total log-reduction in criticality attributable to each FMECA parameter (detectability, occurrence, severity).

TABLES :

Table 1: Conformity of audit criteria and criticality by process before and after preventive actions

Process	Operations, n	Conforming criteria, n/N (%)	Baseline criticality, median [IQR]	C ≥ 100 before, n (%)	C ≥ 100 after, n (%)
Reception	8	6/10 (60.0)	130 [105–162]	6 (75.0)	0 (0)
Storage	3	7/10 (70.0)	147 [98–161]	2 (66.7)	0 (0)
Ordering	6	4/10 (40.0)	225 [210–270]	6 (100)	0 (0)
All processes	17	17/30 (56.7)	168 [120–210]	14 (82.4)	0 (0)

C, criticality index ($D \times O \times S$); IQR, interquartile range; operations, operation-level items. Conformity compared across processes with the Fisher–Freeman–Halton exact test ($p = 0.53$); baseline criticality compared with the Kruskal–Wallis test ($p = 0.019$). Threshold of criticality = 100.

Table 2: FMECA scores of the 17 operations before and after preventive actions

Operation	D/O/S before	C	D/O/S after	C'	RR (%)
Reception process					
Delivery schedule	3/7/8	168	2/3/8	48	71.4
Order correspondence	4/3/9	108	3/2/9	54	50.0
Pallet arrangement	4/7/5	140	2/5/5	50	64.3
Packaging integrity	3/3/9	81	4/2/4	32 ‡	60.5
Product labelling	4/5/6	120	2/3/6	36	70.0
Transport conditions	3/4/8	96	2/2/8	32	66.7
Documentation	6/5/8	240	3/4/8	96	60.0
Computer data entry	5/4/8	160	3/3/8	72	55.0
Storage process					
Stock management	4/2/6	48	2/1/6	12	75.0
Storage area	3/7/7	147	2/3/7	42	71.4
Storage equipment	5/5/7	175	3/3/7	63	64.0
Ordering process					
Order-launch schedule	6/7/7	294	3/3/7	63	78.6
Preparation/control schedule	5/6/8	240	2/3/8	48	80.0
Order-form signature	7/4/6	168	2/2/6	24	85.7
Order allocations	6/5/7	210	3/2/7	42	80.0
Anomaly reporting	7/5/8	280	3/2/8	48	82.9
Handling staff	5/6/7	210	4/3/7	84 †	60.0

D/O/S, detectability/occurrence/severity scores (1–10); C and C', criticality before and after actions; RR, relative reduction in criticality, $(C - C')/C \times 100$. † Recalculated from the scores ($4 \times 3 \times 7 = 84$); the original evaluation table recorded 63 (action still in progress). ‡ As recorded: severity 4 (before: 9) and detectability 4 (before: 3); no supporting evidence available (see sensitivity analyses).

Table 3: Comparison of FMECA parameters before and after preventive actions (n = 17 operations)

Parameter	Before, median [IQR]	After, median [IQR]	Before, mean $\pm$ SD	After, mean $\pm$ SD	p
Detectability (D)	5 [4–6]	3 [2–3]	4.7 $\pm$ 1.4	2.6 $\pm$ 0.7	<0.001
Occurrence (O)	5 [4–6]	3 [2–3]	5.0 $\pm$ 1.5	2.7 $\pm$ 0.9	<0.001
Severity (S)	7 [7–8]	7 [6–8]	7.3 $\pm$ 1.1	7.0 $\pm$ 1.3	0.317
Criticality (C)	168 [120–210]	48 [36–63]	169.7 $\pm$ 69.2	49.8 $\pm$ 21.2	<0.001

p-values from the two-sided Wilcoxon signed-rank test for paired data. SD, standard deviation; IQR, interquartile range. Severity changed in one operation only.

Table 4: Univariate analysis of factors associated with the relative reduction in criticality (RR)

Factor	Median RR [IQR] or Spearman ρ	p (rank)	Crude β (95% CI)	p (reg.)
Process (Kruskal–Wallis H = 7.29)		0.026		
Reception (n = 8)	62.4 [58.8–67.5]		Reference	
Storage (n = 3)	71.4 [67.7–73.2]		7.9 (–2.5 to 18.3)	0.125
Ordering (n = 6)	80.0 [78.9–82.1]		15.6 (5.0 to 26.2)	0.007
Baseline detectability, per point	$\rho = 0.37$	0.142	3.5 (–0.1 to 7.0)	0.055
Baseline occurrence, per point	$\rho = 0.22$	0.402	1.5 (–2.5 to 5.5)	0.427
Baseline severity, per point	$\rho = -0.37$	0.149	–3.5 (–10.0 to 3.0)	0.271
Baseline criticality, per doubling	$\rho = 0.34$	0.177	4.4 (–8.4 to 17.3)	0.475

β , difference in RR in percentage points; CI, confidence interval; 95% CIs and p-values from simple linear regression with HC3 robust standard errors. ρ , Spearman’s rank correlation coefficient with RR. Median RR given with its interquartile range. Baseline criticality expressed per doubling (log₂ scale).

Table 5: Multivariable linear regression of the relative reduction in criticality (percentage points)

Variable	Model 1 β (95% CI)	p	Model 2 β (95% CI)	p
Process: reception	Reference		Reference	
Process: storage	7.4 (−3.6 to 18.3)	0.172	5.6 (−4.9 to 16.1)	0.271
Process: ordering	17.1 (4.1 to 30.2)	0.014	15.6 (−2.7 to 33.8)	0.088
Baseline criticality, per doubling	−1.9 (−9.1 to 5.3)	0.584	—	—
Baseline detectability, per point	—	—	−0.5 (−5.9 to 4.9)	0.837
Baseline severity, per point	—	—	−2.4 (−6.9 to 2.0)	0.255
Intercept	61.8 (55.5 to 68.1)	<0.001	82.9 (36.5 to 129.3)	0.002
R ² (adjusted R ²)	0.50 (0.39)		0.55 (0.41)	
Joint test for process, p	0.032		0.134	

Exploratory models, n = 17 operations; observations within a process are not independent, so CIs and p-values are probably too optimistic. RR, relative reduction in criticality; CI, confidence interval. 95% CIs and p-values with HC3 robust standard errors. Model 1: process and baseline criticality (log₂, centered); Model 2: process, baseline detectability and baseline severity. Intercept: expected RR (%) for reception at the mean baseline criticality (Model 1) or at zero baseline scores (Model 2, not interpretable). Joint test: HC3 Wald F-test for the two process terms.

ACKNOWLEDGEMENTS

The authors thank the staff of the pharmacy of the Children's Hospital of Rabat for their availability and valuable collaboration throughout the study.

DECLARATIONS

Funding: This study received no funding.

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

Ethics Approval: The data were extracted and provided by the officials of the pharmacy of the Children's Hospital of Rabat (CHU Ibn Sina) and analyzed anonymously.

Informed Consent: Individual informed consent was not required, as this was a quality improvement study on the organization of the pharmacy stock circuit, based on process data, with no patient data and no patient contact.

Author Contributions: EDDAHOUMI Y: data analysis, study management, writing of the original draft and revision of the manuscript. BENNANI I: critical reading and correction of the manuscript. BELACHHAB I: data collection. BOUATIA M: supervision of the study. All authors read and approved the final version of the manuscript.

Data Availability: All FMECA scores analyzed in this study are provided in Table 2.

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. Washington (DC): National Academies Press; 2000. doi:10.17226/9728
<https://doi.org/10.17226/9728>
2. World Health Organization. Medication without harm: global patient safety challenge on medication safety. Geneva: World Health Organization; 2017.
3. International Council for Harmonisation. ICH guideline Q9(R1) on quality risk management. Step 4. Geneva: ICH; 2023 Jan 18.
4. International Organization for Standardization. ISO 31000:2018 Risk management: guidelines. Geneva: ISO; 2018.
5. DeRosier J, Stalhandske E, Bagian JP, Nudell T. Using health care failure mode and effect analysis: the VA National Center for Patient Safety's prospective risk analysis system. Jt Comm J Qual Improv. 2002;28(5):248-67.
[https://doi.org/10.1016/S1070-3241\(02\)28025-6](https://doi.org/10.1016/S1070-3241(02)28025-6)
6. Anjalee JAL, Rutter V, Samaranayake NR. Application of failure mode and effect analysis (FMEA) to improve medication safety: a systematic review. Postgrad Med J. 2021;97:168-74.
<https://doi.org/10.1136/postgradmedj-2019-137484>
7. Schumacher L, Dobrinas M, Krähenbühl S, Pasteur C, Borlat ML, Rouiller F, et al. Using risk analysis to anticipate and mitigate failures during a hospital pharmacy relocation. Eur J Hosp Pharm. 2021;28(e1):e164-70. doi:10.1136/ejhpharm-2020-002525
<https://doi.org/10.1136/ejhpharm-2020-002525>
8. Schumacher L, Berthaudin F, Blanc AL, Blatrie C, Staines A, Bonnabry P, et al. Using risk analysis to ensure patients' medication safety during hospital relocations and evacuations. Eur J Hosp Pharm. 2021;28(e1):e171. doi:10.1136/ejhpharm-2020-002619
<https://doi.org/10.1136/ejhpharm-2020-002619>
9. Shebl NA, Franklin BD, Barber N. Failure mode and effects analysis outputs: are they valid? BMC Health Serv Res. 2012;12:150. doi:10.1186/1472-6963-12-150
<https://doi.org/10.1186/1472-6963-12-150>
10. Liu HC, Liu L, Liu N. Risk evaluation approaches in failure mode and effects analysis: a literature review. Expert Syst Appl. 2013;40(2):828-38. doi:10.1016/j.eswa.2012.08.010
<https://doi.org/10.1016/j.eswa.2012.08.010>
11. Wilcoxon F. Individual comparisons by ranking methods. Biometrics Bull. 1945;1(6):80-3. doi:10.2307/3001968
<https://doi.org/10.2307/3001968>
12. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947;12(2):153-7. doi:10.1007/BF02295996
<https://doi.org/10.1007/BF02295996>

13. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat.* 1979;7(1):1-26.
doi:10.1214/aos/1176344552
<https://doi.org/10.1214/aos/1176344552>
14. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc.* 1952;47(260):583-621. doi:10.1080/01621459.1952.10483441
<https://doi.org/10.1080/01621459.1952.10483441>
15. Freeman GH, Halton JH. Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika.* 1951;38(1-2):141-9. doi:10.1093/biomet/38.1-2.141
<https://doi.org/10.1093/biomet/38.1-2.141>
16. Long JS, Ervin LH. Using heteroscedasticity consistent standard errors in the linear regression model. *Am Stat.* 2000;54(3):217-24. doi:10.1080/00031305.2000.10474549
<https://doi.org/10.1080/00031305.2000.10474549>
17. Green SB. How many subjects does it take to do a regression analysis? *Multivariate Behav Res.* 1991;26(3):499-510. doi:10.1207/s15327906mbr2603_7
https://doi.org/10.1207/s15327906mbr2603_7
18. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* 2020;17(3):261-72. doi:10.1038/s41592-019-0686-2
<https://doi.org/10.1038/s41592-019-0686-2>
19. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. In: *Proceedings of the 9th Python in Science Conference*; 2010. doi:10.25080/Majora-92bf1922-011
<https://doi.org/10.25080/Majora-92bf1922-011>