BJ O&PM Brazilian Journal of Operations & Production Management

ISSN 2237-8960 (Online)

ABEPRO ASSOCIAÇÃO BRASILEIRA DE ENGENHARIA DE PRODUÇÃO

RESEARCH PAPER

An AHP–FCE framework for pharmaceutical quality risk prioritization: application to GMP manufacturing processes

Songil Jo[1], Unsong Ri[1], Jiyon Kye[1], Jusong Kim[1], Unhui Yun[1], Cholman Jo[1], Hyongwang Ri[1], Choljin Pak[2], Kyungmin Hwang[3], Choljin Pak[3], Gangsong Ri[4]

[1]KIM IL SUNG University, Pyongyang, Democratic People's Republic of Korea, Korea.
[2]Pyongyang JangCholGu University of Commerce, Pyongyang, Democratic People's Republic of Korea, Korea.
[3]University of Science, Pyongyang, Democratic People's Republic of Korea, Korea
[4]KIM IL SUNG University, Pyongyang, Democratic People's Republic of Korea, Korea.

How to cite: Jo, S. et al. (2026), "An AHP–FCE framework for pharmaceutical quality risk prioritization: application to GMP manufacturing processes", *Brazilian Journal of Operations and Production Management*, Vol. 23, No. 2, e20263231. https://doi.org/10.14488/BJOPM.3231.2026

ABSTRACT

Goal: This study develops and validates an AHP-FCE risk-prioritization model for GMP manufacturing to improve discrimination of critical failure modes beyond conventional RPN-based FMEA.

Design / Methodology / Approach: Five cross-functional QRM experts (production, QA, QC) scored occurrence, severity, and detection for identified failure modes. AHP pairwise comparisons generated criterion weights, and FCE (weighted-score defuzzification) produced final risk grades. Outputs were compared with traditional FMEA rankings; differences were tested using a Wilcoxon signed-rank test and examined against retrospective batch-performance data (2023-2024).

Results: The proposed model significantly reclassified risks upward for critical items (p=0.0043). In routine production, batch-failure rates decreased from 2.8% to 1.1% for albendazole (n=254) and from 2.7% to 1.4% for tetracycline (n=426).

Limitations of the investigation: Evidence comes from a single-site observational before/after implementation without randomization or concurrent controls, covers only two product lines, and depends partly on expert elicitation, which may limit external validity.

Practical implications: The framework can be embedded in regular GMP risk-review and CAPA meetings to prioritize high-impact risks, allocate QA/QC resources, and strengthen traceability of decisions for inspection readiness.

Originality / Value: The study integrates auditable AHP weighting with uncertainty-aware FCE in one operational protocol and links risk-ranking changes to observed manufacturing-quality outcomes rather than rank differences alone.

Keywords: Operations and production management, Risk prioritization, FMEA, AHP, FCE, GMP.

1 INTRODUCTION

Risk prioritization is a central task in operations and production management because it determines how organizations allocate preventive resources, schedule control activities, and stabilize process performance. In GMP manufacturing, these decisions affect batch release reliability, rework burden, and cost of poor quality. Although ICH Q9 and Q9(R1) establish risk-based management principles, they do not prescribe a single quantitative prioritization model for heterogeneous production environments (EMA, 2023; FDA, 2023; ICH, 2005).

Financial support: none.
Conflict of interest: The authors have no conflict of interest to declare.
Corresponding author: cioc1@ryongnamsan.edu.kp
Received: 17 March 2026.
Accepted: 06 July 2026.
Editor: Osvaldo Luiz Gonsalves Quelhas.

Failure mode and effects analysis (FMEA) is a widely adopted tool for process-level risk analysis in quality management. It decomposes complex operations into failure modes and supports structured ranking of potential losses (Bhattacharya, 2015; Cavaignac and Uchoa, 2018; Monforte et al., 2015； Sahu et al., 2016). In standard practice, risks are ordered using RPN=O×S×D (Scipioni et al., 2002). However, equal weighting of occurrence, severity, and detection can understate high-impact events, while ties and subjective ratings may weaken managerial discrimination among competing improvement priorities. These limitations are critical in regulated production systems where low-frequency failures can generate major operational and compliance consequences.

To improve prioritization performance, recent studies have combined FMEA with fuzzy, grey, and other uncertainty-aware decision methods (Li and Chen, 2019; Mangeli et al., 2019; Pillay and Wang, 2003; Xu and Tang, 2002; Zhang and Chu, 2011; Zhou and Thai, 2016). Hybrid applications have expanded across energy, steel, digital-learning, maintenance, and related contexts (Al-Refaie and Aljundi, 2024; Alrifaey et al., 2019; Başaran and Ighagbon, 2024; Cardiel-Ortega and Baeza-Serrato, 2023; Cardiel-Ortega and Baeza-Serrato, 2024; Padash et al., 2024； Pereira and Reis, 2026), showing a broader shift from single-index ranking to multi-criteria risk assessment.The analytic hierarchy process (AHP), originally proposed by Saaty, supports multi-criteria structuring of complex decisions through pairwise comparisons and consistency checking (Saaty, 1987). Its use in risk analysis is well established because it translates expert judgment into transparent weights that can be audited and communicated in managerial settings (Al-Bahar, 1991; Bhatta and Doppler, 2010; Fabjanowicz et al., 2018; Hussein et al., 2010; Li et al., 2018; Maleki-Ghelichi and Sharifi, 2017; Seejata et al., 2018). In FMEA-related applications, AHP has been used to improve scoring criteria and weighting logic in automotive, medical, and manufacturing processes (Aguiar et al., 2010; Chang et al., 2022; Ekmekçioğlu and Kutlu, 2012).

Fuzzy comprehensive evaluation (FCE) complements AHP by aggregating linguistic and uncertain expert assessments through membership functions and weighted synthesis. Prior studies applying AHP–FCE in vulnerability and ecosystem assessments (Cheng and Tao, 2010; Sun et al., 2019) indicate strong potential for decision contexts with imprecise information, yet evidence remains limited for risk governance for routine production quality.

Recent 2025-2026 literature further shows that pharmaceutical operations research is moving toward Pharma 4.0 and digitally enabled quality systems. Systematised reviews of emerging digital innovations in pharmaceutical manufacturing identify PAT, real-time release testing, process modelling, digital twins, AI/ML, IoT, and data-management technologies as major themes for improving product quality, batch release, deviation analysis, and real-time monitoring (Fitzgerald et al., 2026). AI-IoT integration and AI-based manufacturing reviews also emphasize process optimization, predictive monitoring, automated inspection, and quality-control support as important directions in smart pharmaceutical manufacturing (Kodumuru et al., 2025; Nakka et al., 2026). Regulatory reviews of AI/ML in pharmaceutical GMP settings similarly emphasize that implementation must address validation, data integrity, risk management, regulatory oversight, and lifecycle control (Niazi, 2025). Recent Q9(R1)-focused implementation reviews also highlight proportional formality, evidence-based risk assessment, lifecycle integration, and persistent challenges in linking risk outputs to pharmaceutical quality systems (Salve et al., 2026). Therefore, recent AI and Pharma 4.0 oriented methods do not remove the need for transparent, auditable, and low-burden QRM tools; rather, they increase the need for decision frameworks that can connect expert judgment, uncertainty handling, documentation, and operational action.

Against this background, this study proposes an integrated AHP–FCE framework for pharmaceutical process-risk prioritization and positions it as an operations decision-support model. The study addresses three questions: whether AHP–FCE improves discrimination of critical risks relative to conventional FMEA, whether it better aligns with cross-functional expert judgment used in process oversight, and whether its implementation is associated with measurable reduction of quality failures at the batch level in routine manufacturing.

The theoretical contribution of the study is to reframe pharmaceutical FMEA from a product of three equally weighted scores into an integrated multi-criteria decision-support problem for regulated production systems. The proposed model does not simply place two existing methods side by side. Instead, it assigns complementary decision roles to AHP and FCE: AHP makes the relative importance of occurrence, severity, and detection explicit and auditable, whereas FCE aggregates uncertain expert ratings into operational risk grades. This structure advances knowledge by linking criterion weighting, uncertainty handling, and GMP risk-review decisions within a reproducible protocol that can be examined against quality outcomes in manufacturing.

2 MATERIALS AND METHODS

To improve managerial reproducibility, the method was implemented as a predefined decision protocol with four stages: (1) definition of risk criteria and rating rules, (2) AHP-based weight elicitation with consistency control, (3) FCE-based aggregation of expert judgments, and (4) decision-grade assignment for operational prioritization. The same protocol, scoring sheets, and acceptance rules were used across all evaluated failure modes.

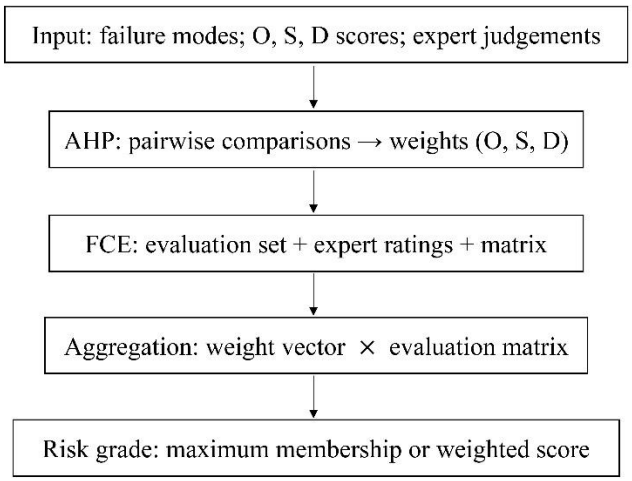

Figure 1 - Procedure for AHP–FCE-based risk grading, outlining the stepwise workflow from inputs to final risk grade

Table 1 - O, S, and D rating scale (1–5) used in this study, defining how occurrence, severity, and detection are interpreted

Score	Occurrence (O)	Severity (S)	Detection (D)
1	Very low (rare)	Negligible impact	Very high
2	Low (infrequent)	Minor impact	High
3	Moderate (occasional)	Moderate impact	Moderate
4	High (frequent)	Major impact	Low
5	Very high (very frequent)	Critical impact	Very low

FCE setting

The evaluation set is defined as $V = \{v_1, v_2, v_3, v_4, v_5\}$, corresponding to five risk grades from very low to very high. Membership degrees are obtained from expert rating frequencies, and the fuzzy evaluation matrix is $R = [r_{ij}]$, where $r_{ij} = p_{ij}/p$ is the proportion of experts assigning factor i to grade v_j and $\sum_{j=1}^{5} r_{ij} = 1$. The comprehensive evaluation vector is computed as $B = W \cdot R$, and the final grade is determined by the maximum membership (or weighted score) rule.

Each factor score $x_i \in [1,5]$ is assigned by experts using the rating scale in Table 1. The fuzzy evaluation matrix is constructed as $R = [r_{ij}]$ with $r_{ij} = p_{ij}/p$, i.e., each row corresponds to a risk factor and each column corresponds to a grade in V. Defuzzification is performed either by the maximum membership principle (selecting the grade with the largest b_j) or by a weighted score $G = \sum_{j=1}^{5} b_j \; v_j$ with $v_j = \{1,2,3,4,5\}$, and the final grade is reported according to the interval containing G. In this study, the weighted-score rule was used for all reported grades to preserve information across adjacent membership levels. Appendix A provides a worked example showing the numerical path from expert ratings to the fuzzy evaluation matrix, AHP-weighted membership vector, weighted-score defuzzification, and final risk grade.

Expert panel, evaluation criteria, and data sources

The expert panel consisted of five QRM practitioners with at least five years of experience in pharmaceutical manufacturing and quality assurance. Experts represented production, quality control, and quality assurance functions to ensure multidisciplinary GMP input. Evaluation criteria

followed the O, S, and D definitions in Table 1. Occurrence scores were primarily quantitative, anchored to internal batch records, deviation/CAPA logs, and validated process parameters when measurable frequencies were available, with expert judgment used only to map observed rates onto the 1–5 scale. Severity and detection scores were primarily judgment-based, grounded in SOPs, control strategy documents, and historical inspection performance. For transparency and reproducibility, each score was labeled in the scoring sheet as "data-anchored" (quantitative frequency available in batch/deviation records within the study period) or "judgment-based" (no quantitative frequency; assigned by structured consensus). Each judgment-based score included a brief justification referencing the relevant SOP/control step or inspection history. To reduce expectancy bias, experts completed O/S/D scoring before final AHP-FCE grades and outcome comparisons were calculated. Batch-performance data were not used to revise expert ratings.

2.1 Determination of relative importance of risk parameters by AHP

AHP provides a hierarchical structure, pairwise comparisons, and an eigenvector-based weighting scheme with consistency checking. Decision makers can build a hierarchical model of a problem, including the goal, criteria, and sub-criteria (factors), to support structured judgments (Saaty, 1987). In this study, AHP was used to determine the relative importance of risk parameters.

The dimension of the comparison matrix should be $n \times n$ form since n criteria are used for assigning relative importance to the estimation criteria. If this matrix is defined as

$$A = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \cdots & \cdots & 1 & \cdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix}$$

Where a_{ij} is the element of i^{th} row and j^{th} column of the matrix, all of its entries are determined according to the relative importance of each one over another in the aspect of the goal. Thus, the entry a_{ij} denotes the relative importance of criterion i in comparison to criterion j. Pairwise comparison can be applied by giving an integer from 1 to 9 or the reciprocal of such an integer to each entry of the matrix so that the relative importance of the factors characterizing each cell can be quantified. In the matrix A, $a_{ij} = 1$ when $i = j$ and $a_{ji}=1/a_{ij}$.

The relative importance (weight) of the i^{th} criteria, w_i is calculated by determining the geometric mean (G_i) of the i^{th} row, and then normalizing as follows:

$$G_i = \sqrt[n]{\prod_{j=1}^{n} a_{ij}}, (i = \overline{1,n})$$

$$w_i = \frac{G_i}{\sum_{j=1}^{n} G_j}, (i = \overline{1,n})$$

Then, the weight matrix W is obtained. The eigenvector (λ_{max}) of the original pairwise comparison matrix A can be determined according to the following equation.

$$\lambda_{max} = \sum_{i=1}^{n} \frac{(AW^T)_i}{nw_i}$$

The consistency index (CI) and consistency ratio (CR) for examining the consistency in a pairwise comparison judgment are calculated from the following equations.

$$CI = \frac{\lambda_{max} - n}{n-1} \; and \; CR = \frac{CI}{RI}$$

If the CR is equal to or less than 0.1, then the judgment is considered to be consistent and acceptable. Otherwise the pairwise comparison matrix should be reconstructed with revised entries.

The random consistency index, RI, can be obtained from Table 2 (Cheng and Tao, 2010; Hussein et al., 2010).

2.2 Rating of quality risks by FCE

Assuming that there are *n* evaluation factors, the evaluation factors domain *U* is denoted as $U=(u_1, u_2, ..., u_n)$. The risk classification is determined by the evaluation rating domain. Assuming there are *m* evaluation ratings, and the evaluation rating domain *V* can be denoted as $V=(v_1, v_2, ..., v_m)$.

After prescribing evaluation factor set, the fuzzy evaluation matrix *R* is established according to the fuzzy relations between evaluation rating domain *V* and evaluation factor set *U*.

$$R=\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}, r_{ij}=\frac{p_{ij}}{p}, i=\overline{1,n}, j=\overline{1,m}$$

Where p_{ij} is the number of experts who assigned v_j value to u_i, *p* is the total number of experts.

Then, a comprehensive evaluation *B* will be obtained as a result of the fuzzy evaluation matrix *R* multiplied by the relative importance matrix *W* to estimate the evaluation factor set *U*.

$$B=W\bullet R=(w_1, w_2, \cdots, w_n)\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}=(b_1, b_2, \cdots, b_m)$$

Where $b_j=\bigvee_{i=1}^{n}(w_i \wedge r_{ij})$ or $b_j=max[\min(w_i, r_{ij})], j=\overline{1,m}$.

Finally, the corresponding evaluation factor is assigned to grade *j* based on the maximum value of the comprehensive evaluation vector *B*.

Table 2 - RI values for criteria numbers, showing the random consistency index by matrix size

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

3 RESULTS

3.1 Assignment of relative importance of risk evaluation factors

We determined the relative weights of occurrence, severity, and detection using AHP based on expert judgments. Specifically, five experts with at least five years of QRM experience completed the pairwise comparisons independently. Individual comparison matrices were aggregated using the geometric mean, and the group consistency ratio (CR) was required to be below 0.1 for acceptance; all reported results satisfied this criterion. The same five experts independently nominated the "most critical" failure modes used in Section 3.3.5.

The primary FMEA parameters (occurrence, severity, and detection) were confirmed as evaluation factors, and the rating domain was defined with five grades as follows.

U= {Occurrence, Severity, Detection}

V= {Very low, Low, Moderate, High, Very High}

For the fuzzy calculation, each evaluation factor was classified into five grades (see Tables 3–5).

Table 3 - Probability of occurrence rating scale classified by grade, defining likelihood bands for occurrence

Grade	Rating	Probability of occurrence
1	Remote	Failure is unlikely (<1 in 1 500 000)
2	Low	Relatively few failures (1 in 150 000 to 1 in 15 000)
3	Moderate	Occasional failures (1 in 200 to 1 in 81)
4	High	Repeated failures (1 in 80 to 1 in 20)
5	Very High	Failure is almost inevitable (>1 in 3)

Table 4 - Severity rating scale classified by grade, summarizing impact levels on product quality

Grade	Rating	Severity (Effect on product quality)
1	Remote	No negative effect
2	Low	Effect on technical procedures only, no effect on product quality
3	Moderate	Increased variation of product parameters or slight deviations from noncritical product specifications may occur that may require some moderate measures e.g. increased frequency of in-process controls, additional tests.
4	High	Significant deviations from product specifications may occur that would require immediate measures e.g. blocking of a batch, recall.
5	Very High	Product safety and efficacy may be impaired

Table 5 - Probability of detection rating scale classified by grade, indicating detection levels of failures

Grade	Rating	Probability of detection
1	Very high detectability	Almost certain: The implemented controls will detect potential cause or mechanism e.g. 100% automated check with alarm, or the failure will almost certainly be detected in the following work progress.
2	High detectability	High: High chance that implemented controls will detect potential cause or mechanism e.g. 100% check with several different measures, such as monitoring process parameters + alarm
3	Moderate detectability	Moderate: Moderate chance that implemented controls will detect potential cause or mechanism e.g. periodic in-process checks or easily recognizable failure and there is some sort of control.
4	Low detectability	Low: Low chance that implemented controls will detect potential cause or mechanism e.g. periodic checks or failure may not be easily detectable.
5	Very low detectability	Very low: The failure is not detectable or there is no routine check for it.

To determine the weight domain corresponding to the evaluation-factor domain *U*, a pairwise comparison table was prepared and the comparison matrix *A* was constructed.

$$A = \begin{pmatrix} 1 & 1/3 & 3 \\ 3 & 1 & 5 \\ 1/3 & 1/5 & 1 \end{pmatrix}$$

The relative-importance vector was calculated as W = (0.258 0.637 0.105).

The maximum eigenvalue $\lambda_{\max}$ of the original pairwise comparison matrix was calculated from the comparison matrix *D* and vector *W*. The consistency index *CI* and consistency ratio *CR* were then computed.

$\lambda_{\max}$ = 3.039, *CI* = 0.019, *RI* = 0.58

CR = 0.033

Because *CR* < 0.1, the judgments were considered consistent and acceptable.

Accordingly, the relative importance of O, S, and D was 0.258, 0.637, and 0.105, respectively.

3.2 Case study of the risk assessment

To demonstrate the applicability of the AHP–FCE-based method, a solid dosage-form production process from a pharmaceutical plant was used as a case study.

Quality risk identification for the solid dosage-form process was performed by experts from the quality risk management team, and the results are presented below.

Table 6 - Quality risk identification of solid preparation production process, listing process units, risk events, and associated risk factors

Process unit	Risk events	Failure-mode factors
Weighing	Weighing errors, dust generation, cross-contamination	Calibration
		Cleanliness of area
Milling and Sieving	Uniformity of particle size, homogeneity of product	Milling time
		Sieving speed

Mixing	Uniformity of product	Mixing time
Stirring	Tablet breakage, influence on disintegration rate	Dosage of binding agents
		Adding speed of binding agents
		Stirring time
Granulation	Direct influence on disintegration time, hardness, and abrasion rate	Mesh size
		Perfection of sieve
Drying	Reduced granule flowability, inappropriate content of active ingredient, tablet breakage due to over-drying	Drying time
		Drying temperature
Compression	Tablet weight variance from the standard, tablet breakage	Adjustment of compressor
Covering with Aluminum foil	Moisture absorption or contamination due to not strict coverage	Pressure of compressed air
		Sealing temperature

To enable comparison of quality risk assessment results, the traditional FMEA RPN values were classified into five grades as shown below. The RPN interval thresholds follow commonly used FMEA risk acceptability guidance and prior process-industry applications (e.g., Pillay and Wang, 2003; Scipioni et al., 2002) and were calibrated to the study site's internal risk-acceptability practice. Consistent with industry practice, these thresholds are organization-specific rather than universal standards.

Table 7 - Quality risk grade according to RPN and risk acceptability, mapping RPN ranges to grades and actions

RPN	Quality risk grade	Rating	Risk acceptability
80 ~ 125	5	Very High	Unacceptable
60 ~ 75	4	High	Further risk reduction measures are required before production commences
24 ~ 50	3	Moderate	Consider investigating further
12 ~ 20	2	Low	Acceptable, but always look for continuous improvement
1 ~ 10	1	Remote	Broadly acceptable

Then, O, S, and D were evaluated for each failure mode factor, and risk grades were assigned using both the traditional FMEA and the proposed method (Table 8).

Table 8 - Solid preparation production process risk assessment, comparing traditional FMEA scores with proposed risk grades

Process unit	Failure mode factors	Traditional FMEA				Risk grade traditional FMA	Proposed method
		O	S	D	RPN		
Weight	Calibration	1	5	2	10	1	5
	Cleanliness of area	2	3	2	12	2	3
Milling and Sieving	Milling time	1	4	4	16	2	4
	Sieving speed	3	3	3	27	3	3
Mixing	Mixing time	3	5	4	60	4	5
Stirring	Dosage of binding agents	4	5	4	80	5	5
	Adding speed of binding agents	4	5	5	100	5	5
	Stirring time	1	4	4	16	2	4
Granulation	Mesh size	2	3	1	6	1	3
	Perfection of sieve	1	3	1	3	1	3
Drying	Drying time	3	5	4	60	4	5

	Drying temperature	4	5	4	80	5	5
Compression	Adjustment of compressor	5	5	4	100	5	5
Covering with aluminum foil	Pressure of compressed air	3	5	4	60	4	5
	Sealing temperature	3	5	3	45	3	5

Comparative analysis between traditional FMEA and the proposed method

Concordance of risk rankings

To quantify the agreement between the two approaches, the ordinal risk grades (1–5) assigned by the traditional FMEA and by the proposed AHP–FCE method were compared using Spearman's rank correlation coefficient. For the 15 failure modes considered, the correlation between the two sets of risk grades was moderate (ρ=0.69), indicating that the methods share a general trend in risk prioritization but differ substantially for several individual failure modes. The moderate Spearman correlation should not be interpreted as simple disagreement or random noise; it indicates that the two methods are not interchangeable and that the proposed model changes priorities for a specific subset of failure modes. When risk grades were dichotomized into "low-to-moderate" (grades 1–3) and "high" (grades 4–5), Cohen's kappa coefficient was 0.48, corresponding to fair-to-moderate agreement, which supports the view that the proposed method reclassifies a nontrivial subset of risks into higher priority levels.

Reclassification patterns and risk spectrum

Beyond the increase in the number of high-risk factors from 7 to 11, the overall risk spectrum shifted upward under the proposed method. The mean risk grade across all 15 failure modes increased from 3.13±1.55 (traditional FMEA) to 4.33±0.90 (proposed method), and given that the risk grades are ordinal, the paired comparison was evaluated using the Wilcoxon signed-rank test (two-sided), which showed a significant upward shift ($p = 0.0043$). Focusing only on failure modes with severity scores ≥ 4, the average grade increased from 3.64 to 4.82 (Wilcoxon signed-rank $p = 0.0156$), indicating that high-severity risks were systematically promoted into higher risk categories, while low-severity modes showed a smaller, non-significant change (1.75→3.00, Wilcoxon signed-rank $p = 0.25$). This selective upward reclassification suggests that the AHP weighting amplifies the influence of severity without uniformly inflating all risk estimates.

Impact on prioritization and resource allocation

To illustrate the practical implications of the different rankings, we compared the cumulative "risk load" covered when progressively addressing failure modes in order of priority, where risk load is defined as the sum of severity scores across all modes. If resources were allocated to the top five ranked failure modes, both methods covered 38.5% of the total severity load; with the top eight failure modes, coverage increased to 61.5% for both methods. This equality reflects the identical set of highest-severity items appearing at the top of both rankings in this case study. In settings where severe but infrequent failures are under-ranked by RPN, severity-weighted prioritization is expected to yield larger differences in coverage for a fixed number of items.

Sensitivity to expert judgment and robustness of rankings

A sensitivity analysis was performed by perturbing the AHP weights within ±10% of their baseline values while maintaining normalization, and recomputing grades using the same expert-frequency evaluation matrix framework defined in Section 2. Across 1,000 random perturbation scenarios, the identity of the top five high-risk failure modes remained completely stable in 100% of cases, and at least four out of five remained unchanged in 100% of cases. By contrast, when the traditional FMEA scores were perturbed by adding ±1 random point to O, S, or D (bounded to 1–5) and grades were recomputed using Table 7 thresholds, the composition of the top five failure modes changed in 99.3% of simulations. These results indicate that the proposed method produces a more robust and stable prioritization against small variations in expert scoring, which is desirable in a subjectivity-prone QRM environment.

Alignment with expert qualitative judgment

This comparison requires an independent list of expert-nominated "most critical" failure modes collected outside the scoring exercise. The expert nominations and their agreement with ranking methods are summarized in Table 9.

Five experts independently nominated up to five "most critical" failure modes based on their experience (Table 9). Three of the expert-nominated modes (FM1, FM3, FM4) appeared in the top-five list of the traditional FMEA (60% overlap), whereas four modes (FM1, FM2, FM3, FM5) appeared in the top-five list of the proposed method (80% overlap). Notably, FM2, a drying-related failure mode unanimously judged as critical by all experts, was ranked seventh by the traditional RPN but second by the AHP–FCE method, illustrating the improved alignment of the proposed approach with expert judgment.

Table 9 - Expert judgment on critical failure modes and agreement with ranking methods

Failure mode ID	Process unit	Failure mode description	No. of experts nominating as "critical" ($n = 5$)	Included in top-5 (Traditional FMEA)	Rank (Traditional RPN)	Included in top-5 (Proposed method)	Rank (AHP–FCE)
FM1	Mixing	Inadequate mixing time	5	Yes	1	Yes	1
FM2	Drying	Excessive drying temperature	5	No	7	Yes	2
FM3	Compression	Incorrect compressor adjustment	4	Yes	2	Yes	3
FM4	Granulation	Improper mesh size	3	Yes	4	No	6
FM5	Sealing (Al foil)	Insufficient sealing temperature	3	No	6	Yes	4

Table 8 shows a clear upward shift in risk classification toward higher priority items under the proposed method.

Quantitatively, 11 of 15 factors (73.3%) shifted to a higher risk grade, 4 (26.7%) remained unchanged, and none decreased. The most frequent upward transitions were 4→5 (3 items) and 2→4 (2 items), with additional shifts of 1→3 (2 items), 1→5 (1 item), 2→3 (1 item), and 3→5 (1 item).

Most upward shifts involved high-severity modes with lower occurrence or detection, underscoring the limitation of equal weighting in the traditional RPN. The direction of these discordant shifts was driven mainly by the AHP criterion weights, which assigned greater importance to severity than the conventional multiplicative RPN structure. The FCE membership functions did not create a separate ranking logic; rather, they translated the weighted expert evaluations into final ordinal risk grades.

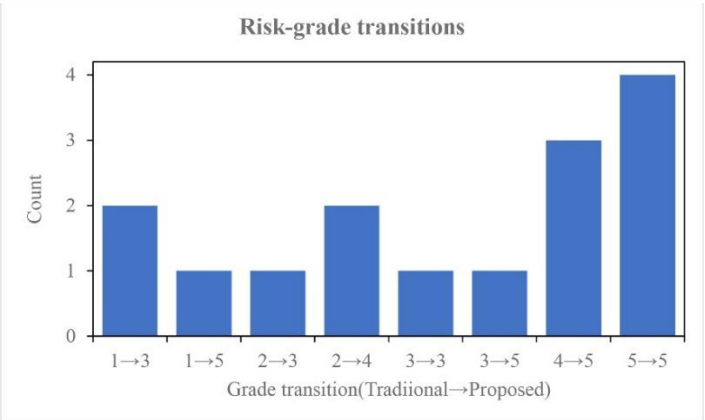

Figure 2 - Counts of risk grade transitions from traditional FMEA to the proposed method (Table 8)

Although some factors had relatively low RPN values, the proposed method still prioritized those with potentially serious quality impacts, thus enabling preventive control before production.

A summary comparison of the two approaches is provided in Table 10.

Table 10 - Summary comparison of traditional FMEA and the proposed method

Item	Traditional FMEA	Proposed AHP–FCE method
High-risk factors (grade 4–5)	7	11
Grade change tendency	Underestimates high-severity, low O/D risks	Elevates high-severity risks via weighting
Decision implication	Fewer items flagged	Broader preventive control

Method validation against GMP batch records from the Jinung Pharmaceutical Factory (Pyongyang, DPRK) during 2023–2024 showed practical effectiveness in routine production. Failure rate was defined as the proportion of batches with GMP deviations requiring rework or batch rejection, and the comparison was an observational before/after implementation review using the same production lines and acceptance criteria. For albendazole (n=254 batches) and tetracycline (n=426 batches), failure rates after implementation of the proposed AHP-FCE method were lower, decreasing from 2.8% and 2.7% (traditional FMEA) to 1.1% and 1.4%, respectively (average reduction 54%) indicating that the proposed risk assessment approach could improve manufacturing performance in practice.

4 DISCUSSION

From an operations and production management perspective, the results show that risk-prioritization structure functions as a managerial policy lever, not only as a technical scoring rule. When equal-weight RPN is replaced by weighted and fuzzy aggregation, the prioritization frontier shifts toward high-impact failure modes, which changes how preventive actions are sequenced, how inspection effort is distributed, and how cross-functional teams (production, QA, and QC) negotiate control trade-offs. In this sense, the AHP–FCE configuration supports both better risk discrimination and more transparent governance of decisions on process quality.

The study also contributes to the FMEA-improvement literature by linking methodological changes to operational outcomes. Rather than reporting only rank differences, the analysis connects reclassification behavior to routine manufacturing performance (batch-failure reduction), indicating practical value for resource-allocation decisions under capacity constraints. This strengthens the argument that multi-criteria, uncertainty-aware risk models can improve decision quality when production systems face heterogeneous failure profiles and limited preventive resources.

The theoretical contribution of this study is the formulation of pharmaceutical risk prioritization as a risk-governance architecture rather than as a revised scoring formula alone. In this architecture, three decision layers are made explicit: (1) criterion preference elicitation, where the relative importance of occurrence, severity, and detection is derived and checked for consistency; (2) uncertainty-aware synthesis, where dispersed expert ratings are converted into risk grades; and (3) operational feedback, where the resulting priorities are examined against routine quality outcomes in manufacturing. This framing extends FMEA-based risk assessment from a ranking exercise to a traceable decision mechanism for regulated production systems.

The contribution is therefore not the invention of AHP, FCE, or a new mathematical operator. Its theoretical value lies in specifying how weighting logic, judgment uncertainty, and manufacturing-performance feedback can be connected in one auditable decision chain. This is important for operations and production management because risk prioritization determines the sequence of preventive actions, the allocation of QA/QC resources, and the justification of CAPA decisions under limited operational capacity. In GMP manufacturing, this decision chain is especially important because risk priorities must remain compatible with batch release, deviation management, validation status, change control, and regulatory inspection.

This positioning also clarifies the boundary of the contribution in relation to prior FMEA extensions. The study does not claim that AHP, FCE, or their combination is new in itself. Traditional RPN-based FMEA has well-known limitations related to equal treatment of occurrence, severity, and detection, duplicate RPN values, and weak representation of expert uncertainty. However, many hybrid methods have already addressed parts of these limitations through fuzzy logic, grey theory, AHP, TOPSIS, evidential reasoning, and other multi-criteria decision models. Therefore, the contribution of the present study is not merely to identify limitations in conventional FMEA.

The specific limitation addressed here is that prior extensions often emphasize one part of the decision problem-for example, improved ranking discrimination, fuzzy representation of uncertain judgments, or criterion weighting-whereas fewer studies organize these elements as a GMP-oriented risk-governance protocol linked to operational quality performance. The proposed framework differs by integrating four elements in one workflow: auditable weighting of FMEA criteria, uncertainty-aware aggregation of expert judgments, translation into CAPA-relevant risk grades, and retrospective assessment against batch-level quality indicators. AHP is deliberately restricted to transparent and consistency-checked criterion weighting, while FCE is used to synthesize uncertain expert ratings into decision-ready grades. This separation of roles distinguishes the framework from approaches that mainly alter the ranking calculation and helps explain why the method is suitable for GMP risk review, CAPA prioritization, resource allocation, and inspection-ready justification. The fuzzy evaluation set also supports audit transparency because predefined categories, membership rules, expert ratings, AHP weights, aggregation matrices, and final grades can be retained as part of the QRM record, consistent with the documentation-oriented expectations of ICH Q9(R1).

This distinction can be seen in comparison with de Souza et al. (2022), who combined FMEA with composition of probabilistic preferences and weighting of FMEA criteria for fast decision-making in complex project scenarios. Both studies treat FMEA as a structured multi-criteria prioritization problem. However, de Souza et al. focus on project-risk prioritization and probabilistic preference composition, whereas the present study focuses on GMP quality risk, separates AHP-based criterion weighting from FCE-based judgment aggregation, and examines whether the resulting grades are consistent with routine batch-quality indicators. The computational burden remains manageable because AHP is applied to the three risk criteria rather than to every failure mode; when the failure-mode list expands, the additional work is mainly repeated FCE matrix calculation that can be prepared before CAPA meetings. In practice, the main expert time is spent on the initial O/S/D scoring and on resolving divergent ratings; the subsequent matrix aggregation and defuzzification are spreadsheet-level calculations. Once the AHP weights and rating template are fixed, additional failure modes can be processed by entering the five expert ratings and applying the same formulas.

For transfer to other organizations, the framework requires four basic inputs: a defined list of process failure modes, agreed O/S/D rating rules, cross-functional expert judgments, and routine quality-performance data such as deviations, rework, rejection, or batch-failure records. The required expertise should include production, QA, QC, and process or validation personnel, because the method depends on both shop-floor knowledge and quality-system interpretation. Implementation can be integrated into existing FMEA or QRM meetings by adding three activities: AHP-based weighting of O/S/D criteria, FCE-based aggregation of expert ratings, and periodic review of whether high-grade risks are reflected in CAPA priorities and subsequent quality indicators.

At the same time, interpretation should remain cautious. The empirical evidence is based on a single-site, before/after implementation setting with two product lines and no concurrent control line, so internal and external validity are not fully established. Potential confounders such as staff learning, equipment maintenance or upgrades, and seasonal raw-material variation cannot be fully excluded.. Future research should evaluate multi-plant datasets, additional dosage forms, and other regulated manufacturing sectors such as medical devices, food production, cosmetics, and chemical processing. Prospective controlled before/after studies, multi-period interrupted time-series designs, comparison with alternative FMEA extensions, and sensitivity testing for weight elicitation and membership-function design would help define when the observed performance gains generalize across industries, process architectures, and quality-system maturity levels.

5 CONCLUSION

This study develops and applies an AHP–FCE risk-prioritization model for quality management in production, extending conventional RPN-based FMEA with explicit weighting and fuzzy aggregation. In operational terms, the framework improves discrimination of high-impact failure modes while maintaining traceable decision logic for cross-functional GMP reviews. Empirically, the model produced a significant upward reclassification of critical risks (Wilcoxon signed-rank test, $p = 0.0043$) and was associated with lower batch-failure rates in routine production (albendazole: 2.8% to 1.1%; tetracycline: 2.7% to 1.4%).

For operations and production management research, the main contribution is to show that risk-scoring architecture is a managerial design lever, not only a compliance formality. By coupling auditable expert weighting with uncertainty-aware evaluation, the proposed approach supports better prioritization of improvement actions, more coherent resource allocation, and stronger

process-governance transparency. Future studies should test multi-plant settings, additional product families, and comparative performance against other multi-criteria prioritization models to define boundary conditions and external validity. Validation should particularly prioritize high-complexity pharmaceutical processes where uncertainty, interacting process parameters, and expert judgment are more prominent than in the solid dosage-form setting studied here, such as sterile injectable manufacturing, aseptic filling, lyophilization, biologics or vaccine production, continuous manufacturing, and personalized or small-batch production. These settings may derive greater value from fuzzy aggregation because risk evidence is often incomplete, linguistically expressed, or difficult to reduce to stable occurrence frequencies. To reduce model fragility, reliable implementation should use an expert panel of at least five members with cross-functional representation from production, quality assurance, quality control, validation, and engineering or process-development expertise when available. Pairwise-comparison matrices should continue to satisfy the standard CR<0.10 threshold, but this numerical criterion should be accompanied by documented expertise diversity and consensus review of strongly divergent pairwise judgments.

Pharmaceutical implications

For pharmaceutical production systems, the framework can be embedded in routine quality and operations meetings as a structured decision-support protocol for risk review, control planning, and CAPA prioritization. It helps teams move from equal-weight scoring to evidence-based prioritization aligned with process impact, detectability, and operational feasibility. At the implementation level, the approach improves consistency across quality assurance, quality control, and manufacturing functions, strengthens justification trails for inspections and batch-disposition decisions, and supports continuous improvement through clearer linkage between risk ranking and shop-floor performance outcomes.

ACKNOWLEDGMENTS

The authors thank the quality assurance and production staff at the study site for technical support in data compilation and process review.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Aguiar, D.C., de Souza, H.J.C. and Salomon, V. (2010), "An AHP application to evaluate scoring criteria for failure mode and effect analysis (FMEA)", *International Journal of the Analytic Hierarchy Process*, Vol. 2, No. 1, pp. 3–13.

Al-Bahar, J.F. (1991), "Project risk assessment using the analytic hierarchy process", *IEEE Transactions on Engineering Management*, Vol. 38, No. 1, pp. 46–52.

Al-Refaie, A. and Aljundi, H. (2024), "A fuzzy FMEA-resilience approach for maintenance planning in a plastics industry", *International Journal of Prognostics and Health Management*, Vol. 15, No. 2, pp. 1–16, https://doi.org/10.36001/ijphm.2024.v15i2.3851

Alrifaey, M., Hong, T.S., Supeni, E.E., As'arry, A. and Ang, C.K. (2019), "Identification and prioritization of risk factors in an electrical generator based on the hybrid FMEA framework", *Energies*, Vol. 12, pp. 649.

Başaran, S. and Ighagbon, O.A. (2024), "Enhanced FMEA methodology for evaluating mobile learning platforms using grey relational analysis and fuzzy AHP", *Applied Sciences*, Vol. 14, No. 19, pp. 8844, https://doi.org/10.3390/app14198844

Bhattacharya, J. (2015), "Quality risk management: Understanding and control the risk in pharmaceutical manufacturing industry", *Journal of Pharmaceutical Science and Innovation*, Vol. 4, No. 1, pp. 29–41.

Bhatta, G.D. and Doppler, W. (2010), "Farming differentiation in the rural-urban interface of the middle mountains, Nepal: Application of analytic hierarchy process modelling", *Journal of Agricultural Science*, Vol. 2, No. 4, pp. 37–51.

Cardiel-Ortega, J.J. and Baeza-Serrato, R. (2023), "Failure mode and effect analysis with a fuzzy logic

approach", *Systems*, Vol. 11, No. 7, pp. 348, https://doi.org/10.3390/systems11070348

Cardiel-Ortega, J.J. and Baeza-Serrato, R. (2024), "Probabilistic fuzzy system for evaluation and classification in failure mode and effect analysis", *Processes*, Vol. 12, No. 6, pp. 1197, https://doi.org/10.3390/pr12061197

Chang, J., Jang, S., Lalonde, R. and Huq, S.M. (2022), "Fuzzy analytic hierarchy process-based risk priority number for risk assessments of commissioning process of a ring gantry LINAC", *Journal of Applied Clinical Medical Physics*, Vol. 23, No. 11, pp. e13760, https://doi.org/10.1002/acm2.13760.

Cavaignac, A.L.O. and Uchoa, J.G.L. (2018), "Obtaining FMEA's indices for occupational safety in civil construction: a theoretical contribution", *Brazilian Journal of Operations & Production Management*, Vol. 15, No. 4, pp. 558-565, https://doi.org/10.14488/BJOPM.2018.v15.n4.a9

Cheng, J. and Tao, J.P. (2010), "Fuzzy comprehensive evaluation of drought vulnerability based on the analytic hierarchy process", *Agriculture and Agricultural Science Procedia*, Vol. 1, pp. 126–135.

De Souza, F.H., Gavião, L.O., Sant'Anna, A.P. and Lima, G.B.A. (2022), "Prioritizing risks with composition of probabilistic preferences and weighting of FMEA criteria for fast decision-making in complex scenarios", *International Journal of Managing Projects in Business*, Vol. 15, No. 4, pp. 572-594, https://doi.org/10.1108/IJMPB-01-2021-0007

Ekmekçioğlu, M. and Kutlu, A.C. (2012), "A fuzzy hybrid approach for fuzzy process FMEA: An application to a spindle manufacturing process", *International Journal of Computational Intelligence Systems*, Vol. 5, No. 4, pp. 611–626.

European Medicines Agency (EMA) (2023), "ICH guideline Q9 (R1) on quality risk management", Available at: https://www.ema.europa.eu/en/ich-q9-quality-risk-management-scientific-guideline (accessed 17 March 2026).

Fabjanowicz, M., Bystrzanowska, M., Namieśnik, J., Tobiszewski, M. and Płotka-Wasylka, J. (2018), "An analytical hierarchy process for selection of the optimal procedure for resveratrol determination in wine samples", *Microchemical Journal*, Vol. 142, pp. 126–134.

Fitzgerald, L., Niarchou, E., Jones, I. and Naughton, B. (2026), "Emerging digital innovations in pharmaceutical manufacturing quality: A systematised review", *Journal of Pharmaceutical Innovation*, Vol. 21, Article 46, https://doi.org/10.1007/s12247-025-10272-5

Hussein, K.M., Hassan, A.A. and Kamal, M.M. (2010), "The multiplicative analytic hierarchy process (MIAHP) as a quality criterion determining the technological value of the Egyptian cotton varieties", *American Journal of Plant Sciences*, Vol. 1, pp. 106–112.

Kodumuru, R., Sarkar, S., Parepally, V. and Chandarana, J. (2025), "Artificial intelligence and internet of things integration in pharmaceutical manufacturing: A smart synergy", *Pharmaceutics*, Vol. 17, No. 3, Article 290, https://doi.org/10.3390/pharmaceutics17030290

International Council for Harmonisation (ICH) (2005), "ICH guideline Q9 on quality risk management (Step 5, first version)", European Medicines Agency, Available at: https://www.ema.europa.eu/system/files/documents/scientific-guideline/wc500002873_en.pdf (accessed 17 March 2026).

Li, H.M., Ni, F.J., Dong, Q. and Zhu, Y.Q. (2018), "Application of analytic hierarchy process in network level pavement maintenance decision-making", *International Journal of Pavement Research and Technology*, Vol. 11, pp. 345–354.

Li, Z. and Chen, L.Y. (2019), "A novel evidential FMEA method by integrating fuzzy belief structure and grey relational projection method", *Engineering Applications of Artificial Intelligence*, Vol. 77, pp. 136–147.

Maleki-Ghelichi, E. and Sharifi, M. (2017), "Prioritize and choose the best process of anaerobic digestion to produce energy from biomass using analytic hierarchy process (AHP)", *Geology, Ecology, and Landscapes*, Vol. 1, No. 4, pp. 219–224.

Mangeli, M., Shahraki, A. and Saljooghi, F.H. (2019), "Improvement of risk assessment in the FMEA using nonlinear model, revised fuzzy TOPSIS, and support vector machine", *International Journal of Industrial Ergonomics*, Vol. 69, pp. 209–216.

Monforte, P.M., Oliveira, U.R. and Rocha, H.M. (2015), "Failure mapping process: an applied study in a shipyard facility", *Brazilian Journal of Operations & Production Management*, Vol. 12, No. 1, pp. 124--134, https://doi.org/10.14488/BJOPM.2015.v12.n1.a12

Nakka, G., Gupta, S. and Reddy, K.K. (2026), "Artificial intelligence in pharmaceutical manufacturing: Applications, case studies, and GxP implementation considerations", *Journal of Pharmaceutical Sciences*, Vol. 115, No. 5, Article 104240, https://doi.org/10.1016/j.xphs.2026.104240.

Niazi, S.K. (2025), "Regulatory perspectives for AI/ML implementation in pharmaceutical GMP environments", *Pharmaceuticals*, Vol. 18, No. 6, Article 901, https://doi.org/10.3390/ph18060901

Padash, A., Fattahi, R. and Vahidi, H. (2024), "A novel hybrid risk assessment approach under fuzzy environment in steel industry", *Engineering Applications of Artificial Intelligence*, Vol. 133, pp. 108655, https://doi.org/10.1016/j.engappai.2024.108655

Pereira, J.C. and Reis, J.C.G. dos (2026), "Method for risk assessment in aeroengine overhaul using a combination of Bayesian networks and fuzzy logic in the context of Industry 5.0", *Brazilian Journal of Operations & Production Management*, Vol. 23, No. 1, Article 2991, https://doi.org/10.14488/BJOPM.2991.2026

Pillay, A. and Wang, J. (2003), "Modified failure mode and effects analysis using approximate reasoning", *Reliability Engineering and System Safety*, Vol. 79, No. 1, pp. 69–85, https://doi.org/10.1016/S0951-8320(02)00179-5

Sahu, A., Yawatkar, P., Jatav, M.P., Jain, G. and Mishra, S. (2016), "Quality risk management of purified water generation system using failure mode and effect analysis as risk assessment tools", *World Journal of Pharmaceutical Sciences*, Vol. 4, No. 5, pp. 224–229.

Salve, N.S., Chordia, Y.K., Tikhe, R.P. and Bansode, A.S. (2026), "Quality risk management (QRM) under ICH Q9(R1): Advanced implementation strategies, regulatory perspectives, and lifecycle integration in pharmaceutical manufacturing", *Research Journal of Pharmacology and Pharmacodynamics*, Vol. 18, No. 2, https://doi.org/10.52711/2321-5836.2026.00027

Saaty, T.L. (1987), "A new macroeconomic forecasting and policy evaluation method using the analytic hierarchy process", *Mathematical Modelling*, Vol. 9, No. 3–5, pp. 219–231.

Scipioni, A., Saccarola, G., Centazzo, A. and Arena, F. (2002), "FMEA methodology design, implementation and integration with HACCP system in a food company", *Food Control*, Vol. 13, No. 8, pp. 495–501, https://doi.org/10.1016/S0956-7135(02)00029-4

Seejata, K., Yodying, A., Wongthadam, T., Mahavik, N. and Tantanee, S. (2018), "Assessment of flood hazard areas using analytical hierarchy process over the Lower Yom Basin, Sukhothai Province", *Procedia Engineering*, Vol. 212, pp. 340–347.

Sun, B.D., Tang, J.C., Yu, D.H., Song, Z.W. and Wang, P.G. (2019), "Ecosystem health assessment: A PSR analysis combining AHP and FCE methods for Jiaozhou Bay, China", *Ocean and Coastal Management*, Vol. 168, pp. 41–50.

U.S. Food and Drug Administration (FDA) (2023), "ICH Q9(R1) Quality Risk Management: Guidance for Industry", Available at: https://www.fda.gov/media/167721/download (accessed 17 March 2026).

Xu, K. and Tang, L.C. (2002), "Fuzzy assessment of FMEA for engine systems", *Reliability Engineering and System Safety*, Vol. 75, No. 1, pp. 17–29, https://doi.org/10.1016/S0951-8320(01)00101-6.

Zhang, Z.F. and Chu, Z.N. (2011), "Risk prioritization in failure mode and effects analysis under uncertainty", *Expert Systems with Applications*, Vol. 38, No. 1, pp. 206–214, https://doi.org/10.1016/j.eswa.2010.06.046.

Zhou, Q.J. and Thai, V.V. (2016), "Fuzzy and grey theories in failure mode and effect analysis for tanker equipment failure prediction", *Safety Science*, Vol. 83, pp. 74–79.

Author Contributions: SJ: Writing, original draft, Methodology. UR: Writing, review & editing. JK: Data curation, Investigation. JK: Formal analysis. UY: Project administration, Supervision. CJ: Resources, Supervision. HR: Conceptualization, Project administration. CP: Validation. KH: Visualization. CP: Data curation, Investigation. GR: Data curation, Investigation.

Abbreviations

AHP - Analytic Hierarchy Process

FCE - Fuzzy Comprehensive Evaluation

PAT - Process Analytical Technology

AI - Artificial Intelligence

ML - Machine Learning

GMP - Good Manufacturing Practice

RPN - Risk Priority Number

FMEA - Failure Mode and Effects Analysis
QRM - Quality Risk Management
QA - Quality Assurance
QC - Quality Control
CAPA - Corrective and Preventive Action
ICH - International Council for Harmonisation
EMA - European Medicines Agency
FDA - U.S. Food and Drug Administration
O - Occurrence
S - Severity
D - Detection
SOP - Standard Operating Procedure
CI - Consistency Index
CR - Consistency Ratio
RI - Random Consistency Index