



RESEARCH ARTICLE

A NON-PARAMETRIC FRAMEWORK FOR ANALYZING SPATIAL HETEROGENEITY AND CONTAMINATION PATHWAYS IN HEALTHCARE ENVIRONMENTS

Mostafa Essam Eissa 

Independent Researcher and Consultant, Cairo, Egypt.

Article Info:



Article History:

Received: 5 June 2025
 Reviewed: 12 July 2025
 Accepted: 24 August 2025
 Published: 15 September 2025

Cite this article:

Eissa ME. A non-parametric framework for analyzing spatial heterogeneity and contamination pathways in healthcare environments. Universal Journal of Pharmaceutical Research 2025; 10(4): 26-31. <http://doi.org/10.22270/ujpr.v10i4.1390>

*Address for Correspondence:

Dr. Mostafa Essam Eissa, Independent Researcher and Consultant, Cairo, Egypt. Tel: +201006154853; E-mail: mostafaessameissa@yahoo.com

Abstract

Background: The systematic management of microbial bioburden in Class C healthcare cleanrooms is a critical factor in patient safety. Standard environmental monitoring often overlooks the complex spatial and statistical relationships of contamination. This study applies a rigorous statistical framework to a comprehensive environmental monitoring dataset to accurately map contamination risk.

Methods: A cross-sectional analysis was performed on 318 microbial surface samples from 28 distinct operational locations in a Class C facility. Colony Forming Unit (CFU) data were analyzed using non-parametric statistics due to non-normal distribution, confirmed by Shapiro-Wilk tests on all locations with sufficient sample size (n=12). The Kruskal-Wallis test with Dunn's post-hoc analysis was used for group comparisons. Spearman's correlation was used to assess inter-location relationships.

Results: Significant spatial heterogeneity in microbial contamination was confirmed ($p < 0.0001$). Dunn's test identified CP C 11 W as the location with the highest contamination burden (mean CFU=12.17). The most statistically robust contrasts were observed when comparing high-burden sites against the cleanest location, CP C 32 WNme (mean CFU=0.67), which serves as a control benchmark. Multiple high-burden locations, including CP C 11 W and CP C 30 NCu, were found to be significantly more contaminated than this benchmark. No Spearman correlations survived the strict Bonferroni correction; however, the relationship between CP C 11 W and CP C 45 Wif ($r=0.882$, $p < 0.05$) approached the significance threshold, suggesting a potential pathway requiring further investigation.

Conclusions: Microbial contamination within the facility is spatially patterned, not random. The analysis provides a definitive hierarchy of risk, highlighting CP C 11 W as the primary target for enhanced sanitation. While correlational pathways could not be statistically confirmed, near-significant results provide a clear direction for future, more targeted sampling to validate operational links between zones.

Keywords: Cleanroom, contamination control, environmental monitoring, hotspot analysis, non-parametric statistics, spatial heterogeneity.

INTRODUCTION

In modern healthcare, the use of microbially controlled environments is indispensable for the safe preparation of sterile products and the execution of sensitive medical procedures¹. Cleanrooms are engineered spaces designed to limit airborne particulates and microbial contamination to rigorously defined levels, thereby protecting both patients and products²⁻⁴. The ISO 14698 standard provides a specific framework for bio-contamination control, outlining the principles for monitoring and managing microbial risk in these

environments⁵. Class C cleanrooms (often corresponding to ISO Class 7 or 8) are critical support areas where the threat of contamination transfer into more sterile zones must be meticulously controlled.

Contamination sources are well-established, with personnel being the most significant contributor, alongside materials, equipment, and HVAC systems^{6,7}. Microorganisms deposited on surfaces can create persistent reservoirs, posing a continuous risk of cross-contamination⁸. An effective environmental monitoring (EM) program is therefore the cornerstone of cleanroom quality assurance. While surface monitoring via contact

plates is standard practice, the subsequent data analysis is frequently limited to checking compliance against static action levels. Accordingly, this approach is inherently reactive and often fails to uncover underlying spatial patterns or systematic risks⁹.

A proactive, risk-based approach requires a more sophisticated application of statistical tools to transform EM data into actionable intelligence¹⁰. A critical aspect of this is recognizing that microbiological data are rarely normally distributed, instead displaying skewed profiles with frequent low counts and occasional high excursions¹¹. Therefore, this statistical reality invalidates the use of parametric tests and demands robust, non-parametric methods for accurate analysis¹²⁻¹⁴. Due to the aforementioned challenges, this study employs such a framework to analyze a large surface contamination dataset from a Class C facility selected as a model example from Bangladesh, Pakistan and India region. Thus, the objectives are to accurately characterize the bioburden distribution, statistically validate contamination hotspots based on the most robust contrasts, and critically assess the significance of potential contamination pathways, providing a precise, data-driven foundation for advanced contamination control.

MATERIALS AND METHODS

Study Design

A retrospective, cross-sectional analysis was performed on a dataset of 318 environmental monitoring results from a single Class C healthcare facility based on Asian country¹⁵. The data were collected from 28 functionally distinct operational zones as part of a routine monitoring schedule.

Data Collection

Surface microbial bioburden was quantified using the contact plate method, with results reported in Colony Forming Units (CFU). The methodology is presumed to follow ISO 14698-1 standards, utilizing a general nutritive agar (e.g., Tryptic Soy Agar) with incorporated disinfectant neutralizers (e.g., lecithin, polysorbate 80)⁵. Standard incubation protocols would typically involve a dual-temperature regimen (e.g., 20-25°C and 30-35°C) to facilitate the recovery of both environmental bacteria and fungi¹⁶. Twenty-five locations were sampled 12 times, and three locations (CP C 44, CP C 45, CP C 46) were sampled 6 times.

Spatial Analysis

The spatial interpretation of data reflects the actual operational layout of the facility, which consists of functionally clustered zones rather than a uniform geometric grid^{16,17}. Visualizations and interpretations are based on this organic, process-driven layout.

Statistical Analysis

All analyses used a significance level of $\alpha = 0.05$, unless otherwise specified.

- **Descriptive Statistics:** Standard metrics (mean, median, standard deviation (SD), minimum and maximum) were calculated for each location¹⁸.
- **Normality Testing:** Normality was assessed for the 25 locations with sufficient sample sizes ($n=12$) using the Shapiro-Wilk test. The remaining 3

locations ($n=6$) were excluded from normality testing due to insufficient statistical power¹⁹.

- **Comparative Analysis:** The non-parametric Kruskal-Wallis test was used to compare median CFU counts across all 28 locations¹³. A significant result was followed by Dunn's post-hoc test to identify significantly different location pairs²⁰.
- **Correlation Analysis:** A Spearman's rank correlation matrix was generated¹⁴. A strict Bonferroni correction was applied to the significance threshold to account for 378 multiple comparisons, resulting in a corrected alpha of $p < 0.00013$ ²¹.

RESULTS AND DISCUSSION

The statistical analysis provided a revised and highly accurate understanding of the facility's microbial contamination patterns²². A summary of descriptive statistics for all 28 locations is provided in Table 1. This table details parameters such as mean, median, standard deviation, minimum, and maximum Colony Forming Unit (CFU) counts for each site. On the other hand, Figure 1 demonstrates dispersion of microbiological count data using box plot diagram.

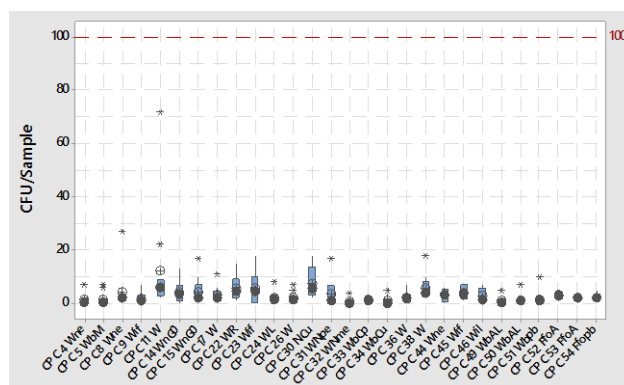


Figure 1: Distribution of microbial surface contamination across 28 sampling locations. A box-and-whisker plot illustrates the median, interquartile range, and full range of CFU counts for each operational zone.

Data distribution and contamination heterogeneity

The Shapiro-Wilk test was rigorously applied to assess the normality of the data, confirming that microbial contamination data were predominantly non-normal²³. Out of the 25 locations with a sufficient sample size of $n=12$ for statistical testing, a significant majority 18 locations, representing 72%, failed to meet the assumption of normality ($p < 0.05$). The remaining three locations, which had smaller sample sizes ($n=6$), were consequently excluded from this specific normality analysis due to insufficient statistical power. Hence, this finding was critical as it firmly validated the necessity and appropriateness of employing non-parametric statistical methods throughout the entirety of this study. Furthermore, the non-parametric Kruskal-Wallis test provided clear statistical confirmation of significant spatial heterogeneity in microbial contamination across the facility (Table 2)²⁴.

Table 1: Summary of descriptive statistics for microbial contamination (CFU) at 28 sampling locations.

| Parameter ID | Mean | Median | SD | Minimum | Maximum |
|---------------|-------|--------|-------|---------|---------|
| CP C 4 Wne | 1.250 | 0.500 | 2.006 | 0.0 | 7.000 |
| CP C 5 WbM | 1.500 | 0.500 | 2.431 | 0.0 | 7.000 |
| CP C 8 Wne | 4.250 | 2.000 | 7.325 | 0.0 | 27.00 |
| CP C 9 Wif | 1.833 | 1.000 | 2.290 | 0.0 | 7.000 |
| CP C 11 W | 12.17 | 6.000 | 19.59 | 2.0 | 72.00 |
| CP C 14 WnG0 | 4.167 | 3.500 | 4.086 | 0.0 | 13.00 |
| CP C 15 WnG0 | 4.333 | 2.000 | 4.960 | 0.0 | 17.00 |
| CP C 17 W | 3.250 | 2.000 | 2.896 | 1.0 | 11.00 |
| CP C 22 WR | 5.500 | 4.500 | 4.380 | 1.0 | 15.00 |
| CP C 23 Wif | 5.500 | 4.500 | 5.808 | 0.0 | 18.00 |
| CP C 24 WL | 2.083 | 1.500 | 2.151 | 0.0 | 8.000 |
| CP C 26 W | 2.000 | 1.500 | 2.045 | 0.0 | 7.000 |
| CP C 30 NCu | 7.333 | 5.500 | 5.944 | 2.0 | 18.00 |
| CP C 31 WNpe | 3.667 | 1.000 | 5.123 | 0.0 | 17.00 |
| CP C 32 WNme | 0.667 | 0.000 | 1.155 | 0.0 | 4.000 |
| CP C 33 WbCp | 1.250 | 1.000 | 0.965 | 0.0 | 3.000 |
| CP C 34 WbCu | 1.000 | 0.000 | 1.758 | 0.0 | 5.000 |
| CP C 36 W | 2.167 | 2.000 | 2.038 | 0.0 | 7.000 |
| CP C 38 W | 5.750 | 4.000 | 4.634 | 2.0 | 18.00 |
| CP C 44 Wne | 3.000 | 3.000 | 2.366 | 0.0 | 6.000 |
| CP C 45 Wif | 4.000 | 3.500 | 2.530 | 1.0 | 7.000 |
| CP C 46 Wil | 2.667 | 1.500 | 2.733 | 0.0 | 7.000 |
| CP C 49 WbAL | 0.917 | 0.500 | 1.443 | 0.0 | 5.000 |
| CP C 50 WbAL | 1.083 | 1.000 | 1.929 | 0.0 | 7.000 |
| CP C 51 Wbpb | 1.417 | 1.000 | 2.778 | 0.0 | 10.00 |
| CP C 52 FfoA | 2.667 | 3.000 | 0.651 | 2.0 | 4.000 |
| CP C 53 FfoA | 2.250 | 2.000 | 1.055 | 1.0 | 4.000 |
| CP C 54 Ffopb | 2.250 | 2.000 | 1.215 | 1.0 | 5.000 |

The test yielded a statistic of $H=104.1$ with a highly significant p -value of $p<0.0001$. This strongly indicates that the risk of contamination is not uniformly distributed but rather varies significantly across different operational zones within the facility. An important observation during this analysis was the presence of an extreme outlier: a reading of 72 CFU recorded at location CP C 11 W. This specific value was intentionally retained within the dataset. Its inclusion underscores its representation of a genuine, high-risk event, emphasizing that effective monitoring programs must possess the capability to detect and respond to such occurrences, which, in a proactive scenario, would immediately trigger a root-cause investigation.

Hotspot hierarchy and significant differences

Dunn's post-hoc test provided a detailed and nuanced perspective on the hierarchy of contamination risk²⁴. While preliminary assessments might have indicated CP C 11 W as a singular primary concern, this rigorous analysis confirmed its status as the location with the highest absolute bioburden, exhibiting a mean CFU count of 12.17. Crucially, the analysis revealed that the most statistically robust contrasts were observed when comparing these high-burden sites against the facility's cleanest location, CP C 32 WNme. Importantly, the analysis clearly demonstrated that multiple locations, extending beyond just CP C 11 W, were significantly more contaminated than the facility's designated low-burden zones.

Table 2: Statistically significant pair wise comparisons from Dunn's Post-HOC Test.

| Higher-Burden Site | Lower-Burden Site | Significance Level |
|--------------------|-------------------|--------------------|
| CP C 11 W | CP C 4 Wne | * |
| CP C 11 W | CP C 5 WbM | * |
| CP C 11 W | CP C 32 WNme | *** |
| CP C 11 W | CP C 34 WbCu | ** |
| CP C 11 W | CP C 49 WbAL | ** |
| CP C 11 W | CP C 50 WbAL | ** |
| CP C 11 W | CP C 51 Wbpb | ** |
| CP C 30 NCu | CP C 32 WNme | *** |
| CP C 30 NCu | CP C 34 WbCu | ** |
| CP C 30 NCu | CP C 49 WbAL | ** |
| CP C 30 NCu | CP C 50 WbAL | ** |
| CP C 30 NCu | CP C 51 Wbpb | * |
| CP C 38 W | CP C 32 WNme | ** |
| CP C 38 W | CP C 34 WbCu | * |
| CP C 38 W | CP C 49 WbAL | * |
| CP C 38 W | CP C 50 WbAL | * |

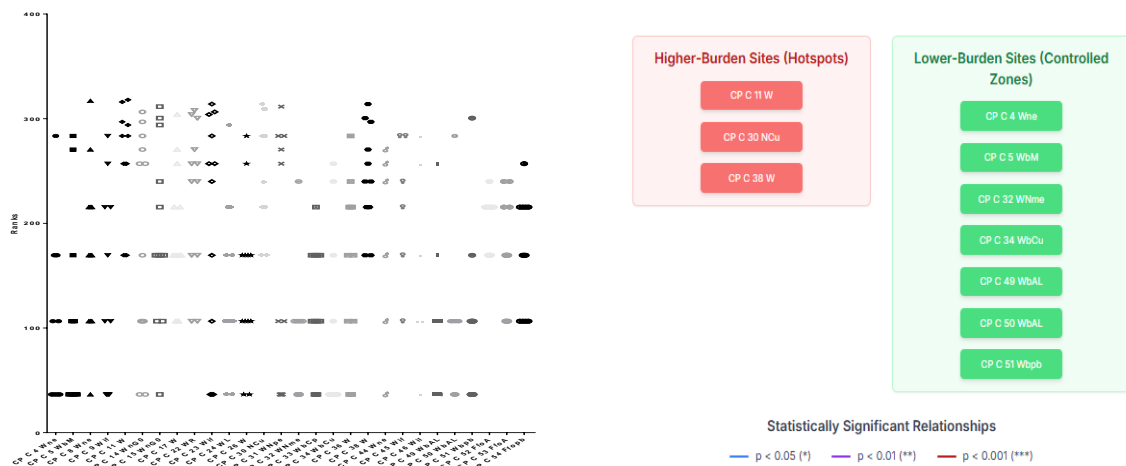


Figure 2: Microbial contamination hotspot cluster diagram and rank chart.

Table 2 provides a comprehensive summary of all statistically significant pair wise comparisons. Specifically, high burden locations such as CP C 11 W, CP C 30 NCu, and CP C 38 W consistently showed statistically significant differences when contrasted with low-risk areas, including CP C 32 WNme, CP C 34 WbCu, and CP C 49 WbAL. This pivotal finding significantly broadens the scope of necessary targeted interventions, shifting the focus from a single isolated hotspot to a broader cluster of interconnected high-risk operational zones. Figure 2 diagram illustrates the statistically significant pairwise comparisons between higher-burden (hotspot) locations and lower-burden (cleaner) locations, based on Dunn's post-hoc test results (extended from Table 2 from the statistical analysis).

Evaluation of correlational pathways

A critical re-evaluation of the Spearman correlation data was meticulously conducted with adjustment after initial extrapolation elucidated apparently significant

associations²⁵. Following the application of a stringent Bonferroni correction for the 378 multiple comparisons performed, utilizing a corrected alpha threshold of $p < 0.00013$, a crucial clarification emerged: no correlations remained statistically significant (Figure 3). This outcome is vital to understand, as it means that even though some correlation coefficients were numerically high (e.g., $r > 0.8$), the study, after rigorous correction for multiple testing, lacked the statistical power to confirm these observed relationships as definitively non-random events.

Nevertheless, it is noteworthy that the correlation observed between CP C 11 W and CP C 45 Wif yielded a strong Spearman correlation coefficient of $r = 0.882$ and an uncorrected p -value of 0.0004. While this particular p -value did not meet the very stringent Bonferroni threshold established for the study, it strongly suggests a potentially robust underlying relationship between these two locations.

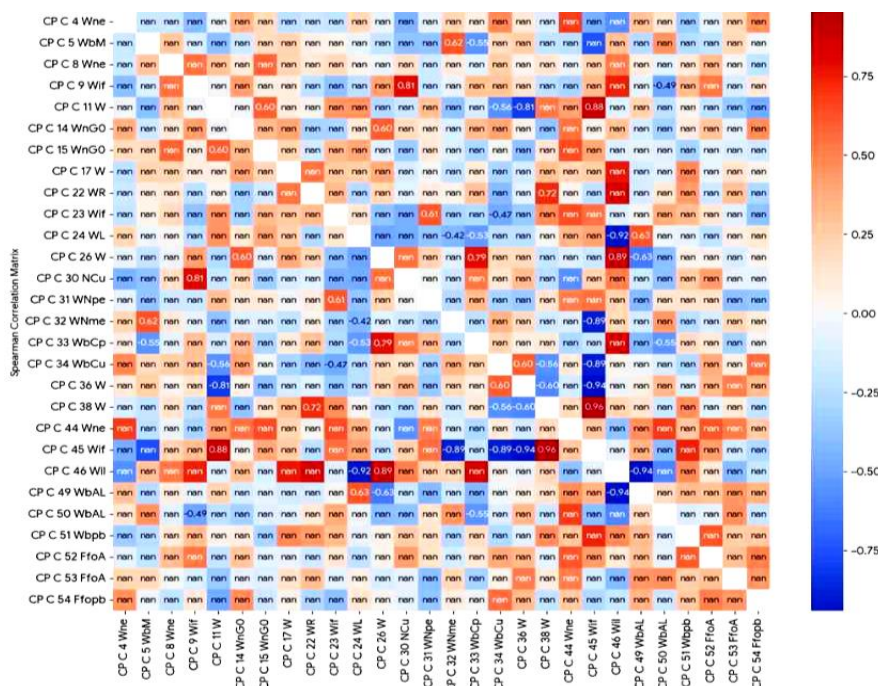


Figure 3: Heatmap displays the Spearman correlation values, but only for the correlations that are statistically significant (i.e., where the p -value is less than 0.05). Non-significant values are not annotated (nan) on the chart, which helps to focus on the most meaningful relationships.

This near-significant finding should not be disregarded. Instead, it offers a clear and valuable direction for future research. It strongly indicates that a more targeted investigation, involving an increased sampling frequency specifically at CP C 11 W and CP C 45 Wif, could provide the necessary statistical power to conclusively validate a true contamination pathway between them.

Limitations of the study

The study's conclusions are drawn from data collected at a single Class C healthcare facility located in the South Asia region. This specificity may limit the generalizability of the findings, as different facilities may exhibit unique contamination patterns due to variations in layout, operational protocols, personnel flow, and environmental conditions. Study needs futuristically to account for potential seasonal fluctuations, long-term trends, or the impact of specific operational changes over time.

CONCLUSIONS

Applying comprehensive non-parametric statistical analysis definitively confirms that surface microbial contamination within the studied Class C healthcare environment is spatially heterogeneous. The analysis has established a clear and statistically validated hierarchy of contamination risk, with CP C 11 W identified as the most contaminated site. Conversely, CP C 32 WNme serves as the most effectively controlled benchmark location within the facility. The findings further delineate a network of several high-risk zones that collectively warrant focused and targeted sanitation efforts. Crucially, this study powerfully underscores the paramount importance of applying appropriate statistical corrections to avoid spurious conclusions. After the application of these corrections, no definitive contamination pathways could be statistically confirmed. However, the identification of a near-significant link between two key sites provides a precise and actionable direction for future, more granular investigations. The primary recommendation derived from this analysis is the immediate implementation of a risk-based monitoring plan, strategically focused on the multiple statistically validated high-burden zones identified.

ACKNOWLEDGEMENTS

The authors acknowledge the facility's quality control personnel for their data collection efforts.

AUTHOR'S CONTRIBUTION

Eissa ME: designed the study, performed the statistical re-analysis, manuscript writing, microbiological interpretation, critically reviewed.

DATA AVAILABILITY

Data will be made available on request.

CONFLICT OF INTEREST

No conflict of interest is associated with this work.

REFERENCES

1. Fletcher SW. Clinical epidemiology: The essentials. Philadelphia: Lippincott Williams & Wilkins 2005;(4).
2. Weber DJ, Rutala WA, Miller MB, *et al.* Role of hospital surfaces in the transmission of emerging health care-associated pathogens: Norovirus, *Clostridium difficile*, and *Acinetobacter* species. *Am J Infect Control* 2010;38(5):S25-33. <https://doi.org/10.1016/j.ajic.2010.04.196>
3. International Organization for Standardization. ISO 14644-1:2015 Cleanrooms and associated controlled environments-Part 1: Classification of air cleanliness by particle concentration. Geneva: ISO 2015.
4. Ashour MS, Mansy MS, Eissa ME. Microbiological environmental monitoring in pharmaceutical facility. *Egypt Acad J Bio Sci, G. Micro* 2011;3(1):63-74. <https://doi.org/10.21608/eaajbsg.2011.16696>
5. International Organization for Standardization. ISO 14698-1:2003 Cleanrooms and associated controlled environments -Biocontamination control-Part 1: General principles and methods. Geneva: ISO 2003.
6. Whyte W. Cleanroom technology: Fundamentals of design, testing and operation. Chichester: John Wiley & Sons 2011;(2).
7. Rawlinson S, Ciric L, Cloutman-Green E. How to carry out microbiological sampling of healthcare environment surfaces? A review of current evidence. *J Hosp Infect* 2019;103(4):363-74. <https://doi.org/10.1016/j.jhin.2019.07.015>
8. Eissa ME, Nouby AS. Assessment of the risk in pharmaceutical facility to the human health based on the ecological surface quality of bacteria in the clean area. *J Innov Pharm Bio Sci* 2015;2(4):608-19.
9. Eissa ME, Mahmoud AM, Nouby AS. Statistical process control in the evaluation of microbiological surface cleanliness quality and spotting the defects in clean area of pharmaceutical manufacturing facility. *Haya: Saudi J Life Sci* 2016;1(1):1-17.
10. Sutton S. Microbial surface monitoring. In *environmental monitoring for cleanrooms and controlled environments* 2016 Apr 19: 93-114). CRC Press. <https://doi.org/10.1201/9781420014853-9>
11. Eissa M. Bioburden analysis and microbiological stability of municipal distribution system through examination of transformed total microbial count dataset. *Front Sci Res Tech* 2024;8(1):57-70. <https://doi.org/10.21608/fsrt.2023.251967.1115>
12. McKight PE, Najab J. Kruskal-wallis test. In: *The SAGE encyclopedia of communication research methods*. Thou Oaks, CA: SAGE Pub, Inc 2017:835-7.
13. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* 1952;47(260):583-621. <https://doi.org/10.1080/01621459.1952.10483441>
14. Spearman C. The proof and measurement of association between two things. *Studies in individual differences: The search for intelligence*. New York: App-Cent-Croft 1961:45-58. <https://doi.org/10.1037/11491-005>
15. Eissa ME. Novel rapid method in ecological risk assessment of air-borne bacteria in pharmaceutical facility. *Mahi Univ J Pharm Sci* 2016;43(3):115-26. <https://doi.org/10.14456/mujps.2016.14>
16. Eissa ME, Mahmoud AM, Nouby AS. Control chart in microbiological cleaning efficacy of pharmaceutical facility. *Dhaka Uni J Pharma Sci* 2015;14(2):133-8. <https://doi.org/10.3329/dujps.v14i2.28501>
17. Eissa M. Diversity of bacteria in pharmaceutical water: Significance and impact on quality. *EPR* 2015;3:54-57.
18. Eissa ME. Assessment of some inspection properties of commonly used medicinal excipients using statistical process control for monitoring of manufacturer quality. *J Nature Sci* 2024;5(1):19-30. <https://doi.org/10.61326/actanatsci.v5i1.3>

19. Eissa ME. Statistical analysis of the critical quality attributes of 1, 2-dihydroxypropane as a pharmaceutical excipient. *Germ J Pharma Bio* 2024 ;3(3):9-17.
<https://doi.org/10.5530/gjpb.2024.3.8>
20. Eissa ME. Study of parameters affecting infection risk from contaminated injectable products using multiple spot contamination model: A case study of insulin vials. *J Chin Pharm Sci* 2016;11.
<http://dx.doi.org/10.5246/jcps.2016.11.093>
21. García-Pérez MA. Use and misuse of corrections for multiple testing. *Method Psych* 2023;8:100120.
<https://doi.org/10.1016/j.metip.2023.100120>
22. Eissa M. Enhancing microbiological stability in municipal water distribution: A descriptive statistical analysis for public health assurance. *J Biomet Stud* 2024;4(1):11-30.
<https://doi.org/10.61326/jofbs.v4i1.02>
23. Eissa ME. Microbiological quality of purified water assessment using two different trending approaches: A case study. *Sume J Sci Res* 2018;1(3):75-9.
24. Eissa ME, Rashed ER, Eissa DE. Dendrogram analysis and statistical examination for total microbiological mesophilic aerobic count of municipal water distribution network system. *High Tech Inno J* 2022;3(1):28-36.
<http://dx.doi.org/10.28991/HIJ-2022-03-01-03>
25. Eissa ME, Rashed ER, Eissa DE. Covid-19 kinetics based on reported daily incidence in highly devastated geographical region: A unique analysis approach of epidemic. *Universal J Pharm Res* 2022. <https://doi.org/10.22270/ujpr.v7i6.870>