

When Incident Rates Mislead: An Integrated Safety and Quality Analysis of FP/PAC Services in Sierra Leone

Onunkwor F.I.,^{1*}, Abiodun, O.P.¹, Aguolu, R.²

¹Department of Public Health, Texila American University, Georgetown, Guyana

²National Agency for the Control of AIDS (NACA), Abuja, Nigeria

***Correspondence to be addressed:** E-mail - felixikennaonunkwor@gmail.com

Cite this article as: Onunkwor FI, Abiodun OP, Aguolu R. When incident rates mislead: an integrated safety and quality analysis of FP/PAC services in Sierra Leone. Texila Int J Public Health. 2026;14(3).
<https://doi.org/10.21522/TIJPH.2013.14.03.Art022>

Received: 06.07.2026

Accepted: 10.08.2026

Published: 30.09.2026

Article Type: Original Research

Abstract Safety monitoring in family planning (FP) and post-abortion care (PAC) services in low- and middle-income countries (LMICs) relies on safety flags and incident rates, yet these indicators are rarely interpreted together. This creates a risk that incident rate differences between service delivery channels may be misread as quality deficits when they reflect differences in service scope and case complexity. This study examined safety flags, incident rates, and quality assessment performance in an integrated framework across centre-based facilities and outreach teams in a Sierra Leone FP/PAC programme. A longitudinal observational analysis used 108 Quality Technical Assessment (QTA) records from 23 paired service delivery points across 14 districts (2019–2022). Safety flags and safety domain scores were compared between remote and in-person assessment modalities using chi-square and Mann-Whitney U tests. Clinical and product incident rates per 1,000 couple-years of protection (CYP) were compared between channels. Domain-specific QTA scores were compared across eight quality domains. Safety flags did not differ significantly between modalities ($p = .740$). Centre-based facilities showed significantly higher incident rates than outreach services ($U = 0.00$, $Z = -2.309$, $p = .021$, $r = -0.82$). Despite this, no significant channel differences were observed in composite QTA final scores ($p = .448$) or in any of the eight quality domains. The divergence between higher incident rates at centres and equivalent quality scores indicates the difference reflects structural differences in service scope and case complexity rather than quality deficits. These findings support integrated safety frameworks in reproductive health programme monitoring.

Keywords Incident Rates, Integrated Analysis, Post-Abortion Care, Quality Assessment, Safety Flags, Sierra Leone.

Introduction

Safety monitoring is a fundamental component of quality assurance in family planning (FP) and post-abortion care (PAC) services [6, 15]. In LMICs, service delivery spans heterogeneous channels—including hospital-level facilities, primary health centres, and community-based outreach teams—and safety data are routinely collected alongside structured quality assessments to provide a composite picture of service performance [1, 2].

Two categories of safety data are commonly collected in FP/PAC quality assurance systems: safety flags, which are assessor-identified observations of protocol deviations or risk conditions at the time of a supervisory visit; and incident rates, which are retrospective counts of adverse clinical or product events per unit of reproductive activity, expressed per 1,000 couple-years of protection (CYP) [3]. These indicators capture different dimensions of safety performance: safety flags reflect conditions for safe practice at the point of assessment, while incident rates reflect adverse outcome frequency across a service delivery period. Both are informative, but their interpretation can be misleading when analysed independently rather than in relation to each other and to the structural characteristics of the channels generating the data.

A particularly important interpretive challenge arises when incident rates differ across service delivery channels in programmes that simultaneously assess quality performance. If incident rates are higher at centre-based facilities compared with outreach teams, a naive interpretation might conclude that quality is worse at centres. However, this fails to account for systematic differences in clinical complexity, patient acuity, referral burden, and service scope [4, 5]. Centre-based facilities typically provide a broader range of clinical services, including surgical procedures, management of complex and referred cases, and care for clients with complicating factors that preclude community-based delivery. These

structural differences generate higher incident counts that reflect service complexity rather than quality deficits [12].

The risk of misinterpreting incident rate differentials as quality differentials is amplified where these indicators are reported separately without analytical integration. Programme managers reviewing incident data without quality scores may draw incorrect conclusions that misdirect supervisory resources or trigger inappropriate corrective action at high-performing but clinically complex facilities [16].

Sierra Leone presents an important context for examining these questions. The country has a geographically dispersed health system across 16 districts, with FP and PAC services delivered through both static facilities and community-based outreach teams [13]. FP and PAC quality assurance in Sierra Leone has relied on structured Quality Technical Assessments (QTAs) conducted through both remote and in-person modalities [17].

This study addresses three specific questions using data from a Sierra Leone FP/PAC quality assurance programme. First, do safety flags or safety domain scores differ between remote and in-person assessment modalities? Second, do clinical and product incident rates differ between centre-based and outreach channels? Third, if incident rates differ by channel, are those differences accompanied by corresponding differences in quality assessment performance?

Materials and Methods

Study Design and Setting

This study used a longitudinal observational design, drawing on routine programme data from a quality assurance programme for FP and PAC services in Sierra Leone [13]. The programme covered 23 service delivery points over 2019 to 2022: 9 static health facilities (centre-based) and 14 community-based outreach teams across 14 active programme districts. Centre-based facilities provide the full

clinical scope of FP and PAC services, including surgical PAC procedures, implant and intrauterine device (IUD) insertion, and management of complex or referred cases. Outreach teams provide medical PAC, counselling, and non-surgical FP methods, and refer clients requiring surgical management to centre-based facilities or public-sector hospitals.

Data Sources and Variables

Three programme data sources were linked. The Quality Technical Assessment (QTA) dataset provided assessment-level safety flag records and domain-specific quality scores, including a composite QTA final score, across 108 facility-year-channel records. Safety flags were defined as assessor-identified deviations from protocol or risk conditions observed during the structured quality assessment visit. Incident data were derived from programme safety records capturing clinical incidents — defined as adverse events during or attributable to a clinical procedure — and product incidents, expressed as rates per 1,000 CYP by service delivery point, channel, and year. Assessment modality (remote versus in-person) was recorded for each QTA assessment cycle.

Statistical Analysis

Safety flags were compared between remote and in-person assessment modalities using Pearson chi-square analysis for presence or absence, and the Mann-Whitney U test for continuous safety domain scores and percentage of high-risk services observed. Incident rates were compared between centre-based and outreach channels using the Mann-Whitney U test, with effect size calculated as

$r = Z/\sqrt{N}$. Domain-specific QTA scores were compared by channel across all eight quality domains using the Mann-Whitney U test. Statistical significance was set at $p < .05$. An integrated interpretation was applied by examining incident rate differences alongside quality performance simultaneously to assess whether higher incident rates were associated with lower quality scores in the same channel. All analyses were conducted using IBM SPSS Statistics.

Ethical Considerations

Ethical approval was granted by the Sierra Leone Ethics and Scientific Review Committee (SLESRC), Ministry of Health and Sanitation, Government of Sierra Leone (Reference No: 008/09/2022, extended to 20 July 2023). All data were de-identified secondary programme records. No individual-level client, provider, or assessor data are reported.

Results

Safety Flags by Assessment Modality

Safety flags were present in 28 of 72 in-person assessments (38.9%) and 12 of 36 remote assessments (33.3%). The difference was not statistically significant ($\chi^2 = 0.354$, $p = .740$), indicating that supervisory modality did not systematically influence whether safety issues were identified during assessment. Safety domain scores were also comparable between modalities ($M = 91.4$ for in-person versus $M = 92.8$ for remote; $U = 1,182.5$, $p = .312$), as were the percentage of high-risk services observed (14.2% versus 13.7%; $U = 1,254.0$, $p = .929$). Table 1 presents the full comparison.

Table 1. Safety Flags and Safety Domain Performance by Assessment Modality (N = 108)

Variable	In-Person (n = 72)	Remote (n = 36)	Statistic	p
Safety flags present, n (%)	28 (38.9%)	12 (33.3%)	$\chi^2 = 0.354$	.740
Safety domain score, M (SD)	91.4 (6.2)	92.8 (5.4)	U = 1,182.5	.312
% High-risk services observed, M (SD)	14.2 (8.9)	13.7 (9.1)	U = 1,254.0	.929

Note: U = Mann-Whitney U statistic; χ^2 = chi-square. No statistically significant differences between modalities across any safety indicator.

Incident Rates by Service Delivery Channel

Centre-based facilities showed significantly higher incident rates than outreach services across all four study years (U = 0.00, Z = -2.309, p = .021, r = -0.82, a large effect). Clinical incident rates at centre-based facilities

ranged from 0.19 to 0.28 per 1,000 CYP across study years, compared with 0.01 to 0.02 per 1,000 CYP at outreach services. Product incident rates were consistently lower than clinical incident rates in both channels and near-zero at outreach settings. Table 2 presents incident rates by channel and year.

Table 2. Clinical and Product Incident Rates per 1,000 CYP by Service Delivery Channel and Year (2019–2022)

Channel / Year	Clinical / 1,000 CYP	Product / 1,000 CYP	Total / 1,000 CYP	n
Centre-based — 2019	0.21	0.04	0.25	9
Centre-based — 2020	0.19	0.02	0.21	9
Centre-based — 2021	0.28	0.03	0.31	9
Centre-based — 2022	0.24	0.02	0.26	9
Outreach — 2019	0.02	0.01	0.03	14
Outreach — 2020	0.01	0.01	0.02	14
Outreach — 2021	0.01	0.00	0.01	14
Outreach — 2022	0.01	0.00	0.01	14
Channel comparison (U test)	U=0.00, Z=-2.309	—	p=.021*, r=-0.82	—

Note: CYP = couple-years of protection. *p < .05. Large effect size (r = -0.82) for channel comparison.

QTA Quality Performance by Service Delivery Channel

Despite significantly higher incident rates at centre-based facilities, no statistically significant differences were observed in composite QTA final scores between channels (p = .448). This pattern held consistently across

all eight quality domains. Clinical governance scores did not differ significantly by channel (p = .677), nor did product quality (p = .657), client-centred care (p = .275), total services section score (p = .890), medical PAC (p = .469), implant insertion (p = .988), or IUD insertion (p = .273). Table 3 presents quality scores by domain and channel.

Table 3. QTA Domain-Specific Performance Scores by Service Delivery Channel (N = 108)

QTA Domain	Centre M (SD)	Outreach M (SD)	U	p
QTA Final Score	92.1 (4.8)	92.5 (4.3)	1,276.5	.448
Clinical Governance	91.3 (5.1)	91.8 (4.7)	1,301.0	.677

Product Quality	93.2 (4.2)	92.9 (4.5)	1,290.0	.657
Client-Centred Care	90.8 (5.9)	91.9 (5.2)	1,188.0	.275
Total Services Section Score	92.4 (4.6)	92.7 (4.1)	1,256.0	.890
Medical PAC Score	91.9 (5.3)	92.3 (4.9)	1,241.0	.469
Implant Insertion Score	92.6 (4.4)	92.6 (4.3)	—	.988
IUD Insertion Score	91.4 (5.7)	92.1 (5.1)	—	.273

Note: *M* = mean; *SD* = standard deviation; *U* = Mann-Whitney *U* statistic. No domain showed a statistically significant difference by channel.

Integrated Interpretation

The findings across the three analytical components are internally consistent and interpretively important. Centre-based facilities generated significantly higher incident rates than outreach services, yet their assessed quality performance was statistically indistinguishable from outreach teams across every quality domain measured. This divergence is inconsistent with a quality deficit explanation, which would predict higher incident rates alongside lower quality scores. The observed pattern is instead consistent with a service scope and case complexity explanation: centre-based facilities provide surgical PAC, manage clinical complications and referrals from outreach settings, and care for a client population with systematically higher clinical risk profiles, generating structurally higher incident counts under equivalent quality conditions [18].

This interpretation is further supported by the safety flag analysis. If centre-based facilities were providing lower-quality care, one would expect higher rates of safety flag identification at those facilities. No such pattern was observed. Safety flags were distributed without significant difference between modalities, and safety domain scores were equivalent between assessment types. The absence of any quality signal accompanying the elevated incident rates at centres reinforces the conclusion that the incident rate differential is a structural caseload phenomenon rather than a quality performance deficit [19].

Discussion

This study provides integrated evidence that incident rate differences between centre-based and outreach FP/PAC service delivery channels in Sierra Leone reflect structural differences in service scope and case complexity rather than quality deficits. Centre-based facilities generated significantly higher incident rates than outreach teams across all four years examined yet showed statistically equivalent quality assessment performance across all eight quality domains. This pattern — divergent incident rates, equivalent quality scores — has important implications for how safety data are monitored and interpreted in reproductive health programmes in LMICs [12].

The core finding aligns with evidence from health systems research showing that case-mix complexity is a primary determinant of adverse event rates in clinical settings, independent of quality of care [4, 5]. In reproductive health, the distinction between facility and community service delivery creates predictable structural differences in the acuity and complexity of client populations each channel serves. Community-based outreach services are typically designed to reach lower-risk clients and to refer clinically complex cases to higher-level facilities [14]. The higher incident rates observed at centres—particularly for clinical incidents, which were 14 to 28 times higher per 1,000 CYP at centres than at outreach sites—are consistent with this structural model and corroborate findings from comparable sub-Saharan African reproductive health programmes [7, 20].

The finding that safety flags did not differ between remote and in-person assessment modalities adds an important complementary dimension. If supervisory modality had influenced safety detection — for example, if remote assessors were systematically less likely to identify safety concerns because they could not observe physical environments directly — this would undermine the validity of remote assessment for safety monitoring. The absence of any modality effect on safety flags, safety domain scores, or percentage of high-risk services observed confirms that remote quality assessment captures safety-relevant information with fidelity comparable to in-person assessment [8, 9]. This supports the use of remote assessment as a legitimate safety monitoring mechanism in LMIC FP/PAC programmes, which is particularly important as remote and hybrid supervisory models become more widely adopted globally [21].

The integrated analytical approach employed in this study addresses a methodological gap in reproductive health programme monitoring. Most published quality monitoring studies report safety indicators separately, leaving programme managers without guidance on how to reconcile divergent patterns [10]. The risk of misinterpretation is not trivial: a programme manager reviewing incident data without quality scores might conclude that centre-based facilities require intensive corrective action when they are in fact performing well under structurally demanding conditions [16]. The integrated framework demonstrated here provides a more analytically sound basis for safety-informed programme management and is consistent with calls for comprehensive quality monitoring frameworks in LMICs [15, 22].

These findings also have implications for the construction of channel-sensitive safety benchmarks in LMIC reproductive health monitoring systems. If centres and outreach teams are held to the same incident rate benchmarks, centres will systematically appear

to underperform not because they provide lower-quality care, but because they manage a more clinically complex population [11]. Channel-specific benchmarks that account for differences in service scope, clinical risk profile, and referral burden would provide a more equitable and defensible basis for performance monitoring. This is consistent with broader evidence that standardised performance benchmarks without case-mix adjustment can systematically penalise facilities managing higher-acuity populations [4, 23].

Strengths and Limitations

This study benefits from a longitudinal four-year dataset that allows incident rates to be examined for consistency across time and by channel, and from the simultaneous analysis of safety flags, incident rates, and quality scores within the same dataset—a methodological strength comparatively rare in reproductive health monitoring literature [10]. Coverage of all 23 service delivery points across 14 districts provides geographic representativeness within the programme context.

Several limitations should be acknowledged. The study relies on routinely collected programme data, which may contain reporting inconsistencies. Incident rates are facility-level aggregates and do not enable case-level analysis of specific clinical or procedural factors driving individual incidents. The interpretation of higher centre-level incident rates as a case complexity effect is plausible and internally consistent with the quality data but cannot be formally tested without individual patient-level clinical records documenting case acuity, referral status, and procedure type [24]. The study is conducted in a single programme context in Sierra Leone; whether similar patterns hold in other West African or LMIC FP/PAC programmes remains to be tested. Future research should incorporate case-level clinical data and formally decompose the structural and quality-related components of channel-level incident rate variation [25].

Conclusion

Centre-based FP/PAC facilities in Sierra Leone showed significantly higher incident rates than community-based outreach services over the four-year study period yet demonstrated equivalent quality assessment performance across all eight quality domains measured. This pattern indicates that the incident rate differential reflects structural differences in service scope and case complexity rather than quality deficits. Safety flags did not differ between remote and in-person assessment modalities, confirming that supervisory modality does not influence safety detection and supporting the validity of remote quality assessment as a safety monitoring tool.

For programme managers and monitoring system designers, these findings underscore the importance of interpreting safety indicators in an integrated framework that considers quality assessment performance alongside incident rates and accounts for channel-specific differences in clinical scope and patient complexity. Developing channel-sensitive safety benchmarks would reduce the risk of misclassifying complex, high-performing facilities as poor performers on the basis of elevated incident counts. For the broader literature on reproductive health quality monitoring, this study demonstrates the analytical value of integrated safety frameworks that treat incident data and quality assessment findings as complementary rather than independent sources of evidence.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1]. Sprockett A, 2017, Review of quality assessment tools for family planning programmes in low- and middle-income countries. Health Policy

Acknowledgements

The authors acknowledge the management and staff of the implementing reproductive health programme in Sierra Leone for institutional support and access to routine programme data.

Ethical Approval

Ethical approval was granted by the Sierra Leone Ethics and Scientific Review Committee (SLESRC), Ministry of Health and Sanitation, Government of Sierra Leone (Reference No: 008/09/2022, extended to 20 July 2023). This study used de-identified aggregated secondary programme data; no individually identifiable participant data were used or reported.

Author Contributions

- **Onunkwor, F.I.:** Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing – Original Draft.
- **Paul, A.O.:** Supervision, Writing – Review and Editing.
- **Aguolu, R.:** Formal Analysis, Writing – Review and Editing.

All authors approved the final manuscript and take responsibility for its content.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to the de-identified secondary programme data arrangement with the implementing organisation.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Plan. 32(2), 292-302. Available from: <https://doi.org/10.1093/heapol/czw144>

[2]. Donabedian A, 2003, An introduction to quality assurance in health care. Oxford University Press, New York.

- [3]. Bruce J, 1990, Fundamental elements of the quality of care: a simple framework. *Stud Fam Plann.* 21(2), 61-91. Available from: <https://doi.org/10.2307/1966525>
- [4]. Rowe SY, Ross-Degnan D, Peters DH, Holloway KA, Rowe AK, 2022, The effectiveness of supervision strategies to improve health care provider practices in low- and middle-income countries: secondary analysis of a systematic review. *Hum Resour Health.* 20, 1. Available from: <https://doi.org/10.1186/s12960-021-00695-3>
- [5]. Oluwatola OA, Phelan A, Nellums LB, et al., 2025, Influence of staffing on quality of care in low- and middle-income countries: a systematic review. *Health Policy Plan.* 40(2), 143-56. Available from: <https://doi.org/10.1093/heapol/czae113>
- [6]. Ayelign A, Zerfw T, Liben ML, et al., 2025, Assessing health systems' capacities to provide post-abortion care: insights from seven low- and middle-income countries. *J Glob Health.* 15, 04020. Available from: <https://doi.org/10.7189/jogh.15.04020>
- [7]. Pasquier E, Moch M, Pinel A, et al., 2023, Abortion-related complications in fragile settings in sub-Saharan Africa: a retrospective population-based study. *Reprod Health.* 20(1), 68. Available from: <https://doi.org/10.1186/s12978-023-01613-3>
- [8]. Simmons M, Brown J, Willems J, 2022, Remote general practice supervision with video cameras. *Aust J Gen Pract.* 51(9), 696-702. Available from: <https://doi.org/10.31128/AJGP-10-21-6213>
- [9]. Hatef E, Wilson RF, Zhang A, et al., 2024, Effectiveness of telehealth versus in-person care during the COVID-19 pandemic: a systematic review. *npj Digit Med.* 7(1), 157. Available from: <https://doi.org/10.1038/s41746-024-01138-4>
- [10]. Hancock NL, Stuart GS, Tang JH, Chibweshia CJ, Stringer JSA, Chi BH, 2016, Renewing focus on family planning service quality globally. *Contracept Reprod Med.* 1, 1. Available from: <https://doi.org/10.1186/s40834-016-0010-9>
- [11]. Kruk ME, Gage AD, Joseph NT, Danaei G, García-Saisó S, Salomon JA, 2018, Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *Lancet.* 392(10160), 2203-2212. Available from: [https://doi.org/10.1016/S0140-6736\(18\)31668-4](https://doi.org/10.1016/S0140-6736(18)31668-4)
- [12]. World Health Organization, 2018, Delivering quality health services: a global imperative for universal health coverage. WHO, Geneva. Available from: <https://www.who.int/publications/i/item/9789241513906>
- [13]. Statistics Sierra Leone and ICF, 2020, Sierra Leone Demographic and Health Survey 2019. Stats SL and ICF, Freetown and Rockville. Available from: <https://dhsprogram.com/publications/publication-FR365-DHS-Final-Reports.cfm>
- [14]. Muga W, Owuor P, Njoroge A, et al., 2024, Barriers to post-abortion care service provision in Burkina Faso, Kenya and Nigeria. *BMC Health Serv Res.* 24, 112. Available from: <https://doi.org/10.1186/s12913-024-10574-3>
- [15]. Hobbs AJ, Mannava P, Hoope-Bender P, et al., 2019, Delivering quality health services: a global framework for action. *Int J Qual Health Care.* 31(Suppl 1), 1-3. Available from: <https://doi.org/10.1093/intqhc/mzy242>
- [16]. Semaan A, Sidney Annerstedt K, Beňová L, et al., 2023, Provision and utilization of maternal health services during the COVID-19 pandemic in 16 hospitals in sub-Saharan Africa. *Front Glob Womens Health.* 4, 1192473. Available from: <https://doi.org/10.3389/fgwh.2023.1192473>
- [17]. Odendaal WA, Anstey Watkins J, Leon N, et al., 2020, Health workers' perceptions and experiences of using mHealth technologies to deliver primary healthcare services: a qualitative evidence synthesis. *Cochrane Database Syst Rev.* 3, CD011942. Available from: <https://doi.org/10.1002/14651858.CD011942.pub2>
- [18]. Rowe AK, de Savigny D, Lanata CF, Victora CG, 2005, How can we achieve and maintain high-quality performance of health workers in low-resource settings? *Lancet.* 366(9490), 1026-35. Available from: [https://doi.org/10.1016/S0140-6736\(05\)67028-6](https://doi.org/10.1016/S0140-6736(05)67028-6)
- [19]. World Health Organization, 2019, WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening.

WHO, Geneva. Available from: <https://www.who.int/publications/i/item/9789241550505>

[20]. Bogale GG, Amde WK, Kasaye HK, Mwangi W, 2024, Factors associated with quality of family planning counselling in health facilities of sub-Saharan Africa: a systematic review. *Reprod Health*. 21(1), 14. Available from: <https://doi.org/10.1186/s12978-024-01741-8>

[21]. McCool J, Dobson R, Muinga N, et al., 2020, Factors influencing the sustainability of digital health interventions in low-resource settings: lessons from five countries. *J Glob Health*. 10(2), 020396. Available from: <https://doi.org/10.7189/jogh.10.020396>

[22]. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J, 2022, The updated Consolidated Framework for Implementation Research based on

user feedback. *Implement Sci*. 17(1), 75. Available from: <https://doi.org/10.1186/s13012-022-01245-0>

[23]. Bradley S, Kamwendo F, Masanja H, et al., 2013, District health managers' perceptions of supervision in Malawi and Tanzania. *Hum Resour Health*. 11, 43. Available from: <https://doi.org/10.1186/1478-4491-11-43>

[24]. Desta BF, Beshir IA, Tefera BB, et al., 2020, Does frequency of supportive supervisory visits influence health service delivery? A dose-response study. *PLoS ONE*. 15(6), e0234819. Available from: <https://doi.org/10.1371/journal.pone.0234819>

[25]. Sultan MA, Miller E, Tikkanen RS, et al., 2025, Competency-based education and training for community health workers: a scoping review. *BMC Health Serv Res*. 25(1), 263. Available from: <https://doi.org/10.1186/s12913-025-12403-7>