COMMENTARY

Open Access

Standardized study performance, quality assurance, and quality control in a cluster-randomized trial: the Pneumococcal Vaccine Schedules trial

Isaac Osei[1,2]* , Benjamin Young[1], Golam Sarwar[1], Yekini A. Olatunji[1], Ilias Hossain[1], Babila G. Lobga[1], Baleng M. Wutor[1], Williams Adefila[1], Emmanuel Mendy[1], Banjo Adeshola[1], Yasir Shitu Isa[1], Yusuf A. Olawale[1], Keita M. Lamin[1], Ebrimah Nyimanta[3], Bubacarr Baldeh[4], Abdoullah Nyassi[1], Momodou M. Drammeh[1], Barjo Ousman[1], Minteh Molfa[1], Rasheed Salaudeen[1] and Grant A. Mackenzie[1,2,5,6]

Abstract

Randomized controlled trials are considered the "gold standard" for evaluating the effectiveness of an intervention. However, large-scale, cluster-randomized trials are complex and costly to implement. The generation of accurate, reliable, and high-quality data is essential to ensure the validity and generalizability of findings. Robust quality assurance and quality control procedures are important to optimize and validate the quality, accuracy, and reliability of trial data. To date, few studies have reported on study procedures to assess and optimize data integrity during the implementation of large cluster-randomized trials. The dearth of literature on these methods of trial implementation may contribute to questions about the quality of data collected in clinical trials. Trial protocols should consider the inclusion of quality assurance indicators and targets for implementation. Publishing quality assurance and control measures implemented in clinical trials should increase public trust in the findings from such studies. In this manuscript, we describe the development and implementation of internal and external quality assurance and control procedures and metrics in the Pneumococcal Vaccine Schedules trial currently ongoing in rural Gambia. This manuscript focuses on procedures and metrics to optimize trial implementation and validate clinical, laboratory, and field data. We used a mixture of procedure repetition, supervisory visits, checklists, data cleaning and verification methods and used the metrics to drive process improvement in all domains.

Keywords Quality assurance, Quality control, Cluster-randomized trial, Monitoring, Supervisory visits

*Correspondence:
Isaac Osei
Isaac.Osei@lshtm.ac.uk
[1] Medical Research Council Unit, The Gambia at London School of Hygiene & Tropical Medicine, Fajara, The Gambia
[2] Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK
[3] Regional Health Directorate, Upper River Region, Ministry of Health, Basse, The Gambia
[4] Regional Health Directorate, Central River Region, Ministry of Health, Bansang, The Gambia
[5] Murdoch Children's Research Institute, Melbourne, Australia
[6] Department of Paediatrics, University of Melbourne, Melbourne, Australia

Introduction and background

The importance of generating high-quality data in clinical trials has been well-documented [1, 2]. According to the World Health Organization (WHO), data quality as applied to clinical research is "...the ability to achieve desirable objectives using legitimate means. Quality data represents what was intended or defined by their official source, are objective, unbiased, and comply with known standards"[3]. Implementation of data quality monitoring methods, such as quality assurance (QA) and quality control (QC) measures are essential to ensure that data generated within clinical trials are measured and reported in adherence with protocols, good clinical practice recommendations, and regulatory compliance [4]. Meinert defines QA as applied to clinical trials as "any method or procedure for collecting, processing, or analyzing study data that is aimed at maintaining or enhancing their reliability or validity" [2]. QA activities include assessing and updating the design of case report forms to accurately request the required data, optimizing data collection methods to minimize human error, and routine retraining of data entry personnel to ensure adherence to trial protocols. It involves any study-related activity from design to publication that enhances data accuracy and integrity [5, 6]. Gutierrez et al. defined QC as "a set of processes intended to 'mitigate' or 'eliminate' the impact of errors that have occurred during data capture and/or processing and that were not prevented by QA procedures" [7]. QC activities include performing regular monitoring of data collection procedures. This may be automated within data collection tools, such as in electronic medical records (EMRs), implementing QC updates with automated value range checks, or cross-field and cross-form validation. Thus, QC is the process of QA that is used to ensure that procedures related to clinical trials adhere to predetermined standards [8].

To date, few studies have reported on data QA and QC procedures applied to ensure data integrity and consistency during clinical trial implementation [4, 9–12]. To ensure high-quality, reliable data and to promote public awareness of the rigorous measures that go into reviewing data for quality and thereby increase public confidence in the evidence generated from clinical trials, researchers must be encouraged to develop, implement, and publish their data QA and QC methods [1, 4].

With an average of 2000 data entries per day in the Pneumococcal Vaccine Schedules (PVS) trial in a rural and low-resourced setting, the consequences of lapses in data QA could swiftly compromise the results of the trial. Given the importance of maintaining a high standard of data quality, and internal validity in randomized controlled trials, the PVS trial developed and implemented standardized prospective study performance metrics and standards to assess quality assurance and data quality. In this paper, we describe the development and implementation of the trial's data quality and assurance methods.

The Pneumococcal Vaccine Schedules (PVS) study

The PVS protocol has been previously described [13]. PVS (trial number: ISRCTN15056916), is a parallel-group, unmasked, non-inferiority, cluster-randomized trial of the community-level impact of two different schedules of pneumococcal conjugate vaccine (PCV13). The primary endpoint of the trial is nasopharyngeal (NP) carriage of vaccine-type pneumococci in children aged 2 weeks–59 months with clinical pneumonia presenting to health facilities in the study area. Approximately 35,000 infants will receive the intervention (PCV13) delivered in two schedules, one with doses scheduled at ages 6, 10, and 14 weeks (3+0 schedule) and the other with doses scheduled at ages 6 weeks and 9 months (1+1 schedule) in 68 randomized clusters. Participant enrolment occurs at 68 geographically separate Reproductive and Child Health (RCH) clinics. Surveillance for clinical endpoints is conducted across 11 health facilities in the study area. The priority indicators for quality assurance in the trial are shown in Table 1. The quality management system, i.e., the QA and QC systems, focused on trial activities essential to human subject protection and the reliability of trial results. The priority list of indicators for QA in the PVS trial was developed based on the consensus of a meeting of the trial investigators. Decisions regarding thresholds of priority indicators were supported by evidence from the literature and available local data. The PI and the trial staff had just completed a 10-year longitudinal pneumococcal surveillance study before the start of the PVS trial and the experiences gained from the implementation of the surveillance study as well as current literature informed the development of the QC procedures.

Clinical

Development of standard operating procedure documents

To ensure standardization of activities across all 11 health facilities and 68 RCH clinics, study-specific procedures (SSPs) were developed. To date, 37 SSPs have been developed covering all relevant study-related activities (see Additional file 1). These SSPs are revised every 2 years or as and when necessary. Staff are trained on SSPs before conducting any study-related activity.

Staff training and certification

A training program for staff has been developed to ensure that every staff undergoes the requisite training on their job requirements. The International Conference

Table 1 Priority indicators for quality assurance in the PVS trial

Activity	Indicator
Recruitment	All enrolments have signed consent forms
EPI PCV13	No stock-outs
	> 97% compliance with PCV13 storage at the appropriate temperature
	Utilization and accountability follow EPI procedure
Vaccination	> 97% of children receive correct schedule
	> 95% receive complete schedules
	> 90% complete schedule by age 7 months (3 + 0) and 12 months (1 + 1)
Demographic surveillance	> 95% follow-up of internal migrations out of 1 + 1 clusters before receipt of a booster dose
Clinical	> 97% of children investigated with NP according to SOP
	> 95% of NP swabs performed to adequate quality as per SOP
	Nurse counts respiratory rate within ± 3 breaths per minute of clinicians count > 95% of the time
	Nurse correctly classifies clinical pneumonia in > 95% of instances
Laboratory	< 7% blood cultures contaminated
	< 5% discordance of pneumococcal isolate serotyping in internal and external QC
Radiology	> 95% of radiographs classified as "satisfactory" according to WHO standard

NP nasopharyngeal, *WHO* World Health Organization, *QC* quality control, *SOP* standard operating procedures, *EPI* Expanded Program on Immunization, *PCV* pneumococcal conjugate vaccine

on Harmonisation – Good Clinical Practice (ICH-GCP) training with certification, either in person or online is organized for all staff. Refresher training with certifications is also conducted yearly for staff whose GCP certificates have expired. Study protocol training is conducted by the Principal Investigator (PI) or the Trial Epidemiologist for all staff at the beginning of the trial and all new staff before they commence any study-related activities. At least two refresher trainings on SSPs are conducted yearly for staff, usually at each half of the year. All training is documented and signed by all attendees and filed in their respective training files in the Trial Master File.

Supervisory checklists
Supervisory checklists were developed to standardize the internal monitoring of clinical and laboratory activities including the adherence to the protocol and SSPs. The checklists were developed prioritiing the monitoring of key trial-related indicators such as QC of nasopharyngeal swab collection technique (primary endpoint sample), QC of blood culture collection technique, anthropometry, respiratory signs, and other key study-related core activities. The checklist to monitor activities at the laboratory includes evaluation of functionality and calibration status of equipment and devices, storage conditions and expiration status of reagents/kits/media and reference strains, handling of specimens, documentation of internal and external QC activities, availability, and adherence to SOPs and staff training documentation. The checklists were finalized following 1-month piloting. The checklists enable a standard approach to internal monitoring of

study staff adherence to the protocol and SSPs (see Additional files 2 and 3).

Supportive supervisory visits
Well-structured bi-weekly supportive supervisory visits are conducted by research clinicians and nurse coordinators to all 11 health facilities. These visits are conducted using the supervisory checklist. The visits involve observation of study procedures implemented by staff and review of completed data collection forms. The core activities assessed during the supervisory visits to the clinics include participant registration, screening for the primary endpoint, disease surveillance indicators, primary endpoint sample collection technique, review of completed study informed consent forms, and sample handling and transport. Consistency in the assessment of respiratory indicators such as grunting, nasal flaring, lower chest wall in-drawing and respiratory rate which are key variables for primary endpoint determination was assessed. Trial nursing staff and nurse coordinators independently measured the respiratory indicators, and differences (respiratory rate within ± 3 breaths per minute) were resolved through paired assessment. The findings from this activity are reported during our monthly QA meetings. The visits enable the detection of recruitment and protocol adherence issues and prompt resolution. The PI and Trial Epidemiologist conduct laboratory supervisory visits at least once a month. At the end of each supervisory visit, the supervisor provides feedback on the performance of the staff and offers corrective actions to issues discovered during the visit. This

approach proactively improves the data quality through onsite verification of validity and completeness of data and promotes high standards and teamwork.

Weekly progress meetings

The trial management group, comprising the PI, Trial Epidemiologist, Trial Coordinator, Data Manager, and Higher Laboratory Scientific Officer meets weekly to review progress in recruitment, clinical endpoint surveillance, blood culture contamination rate, and data quality. Weekly clinical meetings are held to review the practical aspects of clinical activities. The meetings are also attended by research clinicians, nurse coordinators, laboratory scientific officers, data staff, and X-ray technicians. Data from the previous week are presented in electronic reports highlighting discrepancies within the disease surveillance data collection systems (Fig. 1). Reports also compare data from clinic and laboratory logbooks and the clinical and laboratory EMR for accuracy and completeness (Fig. 1). Any data discrepancies identified are tasked to data team members to resolve and follow up the following week. The meetings are held in person and virtually via Zoom to allow study staff in satellite locations to attend remotely.

Data cleaning meetings focused on the technical aspects of identifying and resolving data queries are held weekly. The PI, Trial Epidemiologist, and Project Manager hold weekly trial implementation meetings focusing on logistics to continue effective trial activities. The progress meetings also serve as a forum where discussions are held, and strategies are developed to address peculiar challenges arising from study procedure implementation.

Monthly quality assurance meetings

Monthly QA meetings bring members of all trial departments together to report aggregated monthly data reporting the performance of key indicators. Monthly QA meetings served as a platform to better observe and discuss trends in the indicators over time, with greater certainty relative to the weekly meetings. Monthly QA activities also involved departmental reports on QA activities such as (a) NPS performance and training, (b) accuracy of interpretation of clinical signs, (c) calibration of weighing scales, (d) anthropometry technique, (e) blood culture volume and contamination, (f) adherence to standardized investigation of patients, (g) lab staff competence evaluations, (h) lab QC of equipment and reagents, (i) internal and external lab QA on microbiological procedures, (j) total monthly enrolment by group allocation, (k) booster dose defaulter vaccine update, (l) linkage status of enrolled participants, and (m) data query resolution.

	Report Date:	12-Jul-23	CLINICAL WEEKLY SUMMARY REPORT (URR)													
			03-Jul-2023 to 09-Jul-2023													
	Indicators		TOTAL			BANSANG		BRIKAMABA		JANJABUREH		JAHALLY		NOTES (for clinical team)		
			Clinic LB	EMR	Lab	Clinic LB	EMR	Clinic LB	EMR	Clinic LB	EMR	Clinic LB	EMR			
5	Total Registered															
6	Total Clinical Pneumonia Screened															
7	*Not screened (age<14 days or non-dss)*															
8	No. patients with clinical pneumonia	DSS														
9	Number of Cli pneu with NPS done, n (%)	DSS														
10	Total Susp. Pneumonia admission (DSS)															
11	Number of Total Susp Pneu admission with NPS															
12	Total Patients with Village Form															
13	Total patients with OPD form															
14	Susp. Pneumonia admission															
15	Susp. Sepsis admission															
16	Susp. Menin. Admission															
17	Focal Sepsis Admission															
18	Other Medical Problem Admission															
19	Burn/trauma/surgical															
20	Non-infectious neonatal admission															
21	Susp. Pneumonia and Sepsis admission															
22	Susp. Pneumonia and Menin admission															
23	Susp. Sepsis and Menin admission															
24	Total No. patients with suspected P/S/M admission															

Fig. 1 Weekly project progress report comparing data from clinical and laboratory logbooks and clinic and laboratory EMR. Clinic LB, Clinic Logbook; EMR, Electronic Medical Records; DSS, Demographic Surveillance System

Radiology

The detection of radiological pneumonia on X-rays of children admitted with suspected pneumonia, sepsis, or meningitis is a key safety outcome in PVS. In consideration of the fact that the performance, reading, and interpretation of X-rays remain a vital safety component of the study, we have established methods to perform and interpret X-rays according to WHO recommendations [14]. Two readings of each radiograph (masked to identity) are undertaken by two independent readers (clinicians). Readings discordant for radiological pneumonia are resolved by a third reader. All readers are calibrated to the WHO standard for the diagnosis of radiological pneumonia with consolidation. All readers are required to achieve high levels of agreement on radiological pneumonia with blinded samples of the WHO standard set of 222 radiographs (kappa statistic > 0.8) before they read radiographs.

Laboratory

External QC of conventional microbiology

The MRC laboratory in Basse, where PVS operates, participates in the One World Accuracy External Quality Assessment (EQA) program. The EQA program provides blinded microbiology specimens each quarter to assess the laboratories' accuracy in performing bacterial identification and antimicrobial susceptibility testing. The participation of the Basse laboratory in this EQA program ensures that microbiology test results are reliable and that conventional microbiology testing methods meet the highest international standards.

Inter-operator internal QC of sweep and isolate serotyping of nasopharyngeal samples

The primary endpoint specimen in this trial is the nasopharyngeal swab (NPS) taken from resident children aged 2 weeks–59 months with clinical pneumonia presenting to health facilities in the study area. For internal QC, a 10% random selection of isolate suspensions of positive culture for pneumococcal serotyping is processed by two independent operators, and the results are compared. The second operator uses the gold standard Quellung method which is compared to the initial isolate serotyping by the first operator. The data manager runs query scripts in the database to check for inconsistencies between the Quellung result and the initial serotype result. Investigations are conducted into the reasons for discrepancies and procedures are adjusted accordingly to avoid the identified errors. The final serotype result for two discordant results is based on the Quellung results.

External QC using spiked nasopharyngeal specimens

One hundred sixty-two spiked NP specimens from Murdoch Children's Research Institute (MCRI), Melbourne, Australia, with unknown combinations of pneumococcal serotypes were sent to Basse for external QC. The samples were stored in the laboratory freezer and selections of 10–20 specimens were periodically sent to the clinics and then submitted to the laboratory, mimicking samples collected at the clinic. PVS laboratory staff were blind to the nature of these samples. The results of an evaluation of these samples by the PVS laboratory team are sent to MCRI for comparison with the master list results at MRCI. Discrepancies are determined and appropriate corrective actions are undertaken to streamline the NP culture and serotyping process.

External QC using pneumococcal microarray serotyping

Aliquots of 100 PVS NP samples were sent to St George's University of London (SGUL), UK, for microarray assay. This was done by aliquoting 50 µL of NP specimen into cryo-vials and shipping on dry ice to the UK. The latex sweep serotyping results from the Basse laboratory are compared to the micro-array (gold-standard) results for concordance. Discordant results are assessed, and corrective actions are implemented.

Verification of laboratory logbook serotyping results compared to LEMR results

A non-laboratory project staff is trained and tasked to compare serotype results in the laboratory logbook to the results in the laboratory electronic medical records (LEMR). The 100% verification of the LEMR results with the laboratory logbook results ensures that discrepancies arising from data entry errors and transcription errors are identified and resolved prospectively.

Field procedures

Field team meetings led by the trial coordinator and attended by the field coordinator, field supervisors, and senior field assistants are held every Friday. The meetings focus on the review of enrolment and vaccination weekly electronic data capturing records updates, vaccine storage fridge temperature monitoring charts, query resolution, and new data issues arising. Data reports are produced weekly to guide field activity and the field team reports weekly on; enrolment, informed consent documentation, verification of enrolment number with database record, vaccination of those due or defaulting the 3rd dose when allocated to 3 + 0 schedule, vaccination of those due or defaulting the booster dose when allocated

to 1 + 1 schedule, enrolment of registered but not consented infants, incorrect group allocations, and incorrect vaccine administration.

Enrolment using an optimal sampling frame

The sampling frame for participant selection is obtained from the health and demographic surveillance system (HDSS) in the trial area which is supplemented by real-time enumeration of births reported at RCH clinics. The sampling frame is constantly updated, verified, and distributed weekly for use by trial staff. The sampling frame included demographic information including the mother's name, infant's name, date of birth, father's name, name of household head, village name, compound number, household number, and individual ID number. Enrolment of eligible participants and group allocation is based on the cluster allocation of the village of residence.

Linkage of identity of individual participants

The HDSS assigns a unique 14-digit identity number (DSS ID) to everyone who has been a resident for more than 4 months as confirmed by DSS records in the study area. Children born to or cared for by a parent/guardian who has been resident for more than 4 months are also assigned a DSS ID. At enrolment, after linkage to the sampling frame, a trial sticker with the PVS ID number is fixed to the infant welfare card (IWC). A sticker with the DSS ID is also stuck on the cover of the IWC. If a child visits the health facility, the child's DSS ID on the IWC is used to search for the child's demographic details which are entered into the clinic EMR. This robust process allows the prompt linkage of individual participants across the data collection systems in our database.

Vaccination delivery and vaccination dates

The Expanded Program on Immunization (EPI) administrative units in all the 68 clusters are used for vaccine administration and this minimizes incorrect administration of the schedules. Enrolment and vaccination data are entered electronically in real-time using an offline application. Data from all sites are synchronized weekly. Field workers undertake real-time verification of vaccination records and signed consent forms at all the immunization clinics within the PVS study area.

Data management plan

A data management plan (DMP), which has been developed was approved by the Medical Research Council Unit The Gambia (MRCG) Head of Data Management. The DMP provides a detailed approach to managing data, both at the data collection centers and data management units. Data is electronically captured offline on electronic case report forms (eCRFs) in the RCH clinics, health facilities, and the laboratory using an encrypted laptop/windows tablet and then synchronized with a central server every week. The eCRFs have an in-built quality check process such as range checks, skip patterns, and validations that aims to minimize the rate of errors/missing fields and enhance the quality of the data. Reports are generated periodically by the data manager for data quality assurance [13]. The DMP is available on request.

Data quality reports

Data queries are generated weekly by the data manager for resolution by trial staff. A data query script is prepared for data quality checking. The query script is run to identify missing data, data inconsistencies, invalid form correspondence, invalid dates of birth, vaccinations administered and required data entries, hospital admissions and the presence of sequential forms, and a variety of criteria defined to highlight for implausible or invalid data. The data team provides periodic data support visits to the various trial sites. These visits enable the data team members to engage the trial staff on the types of queries generated and assist them in how to resolve the queries. At the end of every month, the data manager generates primary and secondary endpoint QA reports for review during the monthly QA meetings.

Impact and preliminary results

An overview of the PVS QA/QC activities is shown in [see Additional file 4]. Though the trial is ongoing, preliminary findings show the impact of the QC and QA measures on the primary and secondary endpoint measurements. The results of the last quarter of 2023 One World Accuracy EQA microbiology results showed a 100% pass. This result indicates the reliability of our microbiology testing methods and in effect, strengthens the safety monitoring procedures in our study. There was an approximately 8% discrepancy between the two primary X-ray readers. The 8% discrepancy is very low indicating high levels of agreement and high adherence to the WHO standard. The internal 10% QC on isolate serotyping results compared to Quellung (gold standard) results showed disagreement on 3.5% of isolate serotyping results. The 3.5% disagreement is within the acceptable range ($\sim$5%) indicating the reliability and validity of our primary endpoint measurement. The contamination rate of blood cultures is a key QA indicator. This indicator is significant because the detection of bacteremia from blood culture is an important safety monitoring endpoint. The cut-off of blood culture contamination rate is < 7% in our study (Table 1). The blood culture contamination rate for March/April 2022 is shown in Fig. 2. The chart shows higher than expected blood culture
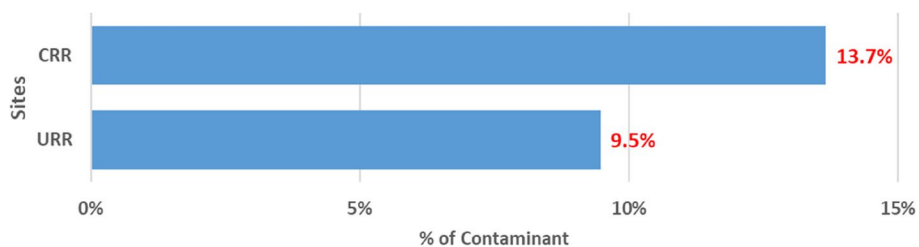
Fig. 2 Proportion of blood culture contamination in March/April 2022 in two PVS Clinical HDSS facilities. HDSS, health and demographic surveillance system; CRR, Central River Region; URR, Upper River Region

contamination rates in our two sites, the Central River Region (CRR) and the Upper River Region (URR). Following these findings, we enhanced our supervisory and monitoring visits, and in CRR, contamination rates decreased from 13.7% (38/277) in March/April to 7.4% (11/148) in May (Fig. 3) and 6.5% (12/184) in June (Fig. 4). In URR, contamination rates declined from 9.5% (26/273) in March/April to 3.8% (6/159) in May (Fig. 3) and 4.3% (7/160) in June (Fig. 4). Tables 2, 3, and 4 show form correspondence queries as reported during our monthly QA meetings over the same period, March to June 2022. The findings show a similar decline in form corresponding queries following enhanced supervision and provision of training and data support to the clinical staff. The implementation of enhanced QA and QC methods resulted in the decline of contamination rates and form correspondence queries, showing a good example of the positive impact of our continuous QA and QC measures on the improvement of quality data.

Summary and discussion

The PVS study developed and implemented QA and QC tools to evaluate study performance. Large cluster-randomized trials, such as the PVS trial which generate a high volume of data require a well-designed QA and QC program to ensure high-quality implementation. Our approach focussed on identifying errors and problems early to take prompt corrective actions. In contrast to PVS, which is a single-center study, the few studies that have reported on their QA and QC procedures were multi-country and multi-center studies, mostly coordinated by consortiums [4, 9–12]. The consortiums form Quality Assurance committees that develop performance measures based on predefined
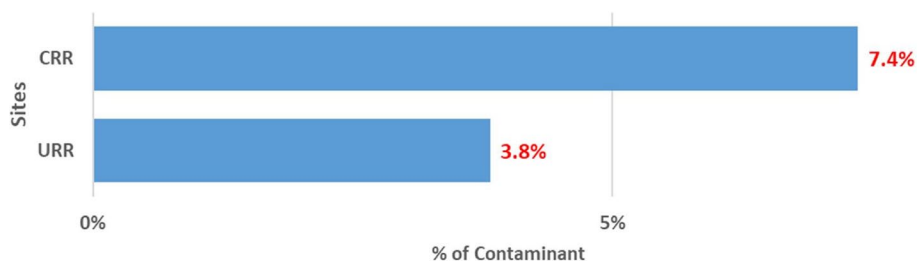

Fig. 3 Proportion of blood culture contamination in May 2022 in two PVS Clinical HDSS facilities. HDSS, health and demographic surveillance system; CRR, Central River Region; URR, Upper River Region
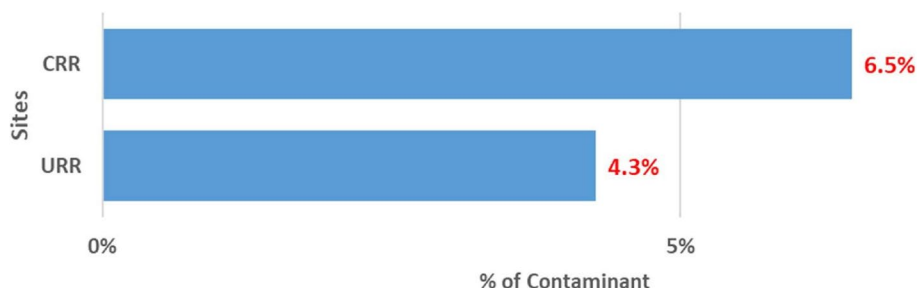

Fig. 4 Proportion of blood culture contamination in June 2022 in two PVS Clinical HDSS facilities. HDSS, health and demographic surveillance system; CRR, Central River Region; URR, Upper River Region

Table 2 Form corresponding queries (09 September 2019 to 30 April 2022)

Site	Clinic site	Total by site	Total
URR	Basse	49	**136**
	Demba Kunda	6	
	Fatoto	24	
	Garawol	12	
	Gambisara	27	
	Koina	13	
	Sabi	5	
CRR	Brikamaba	22	**307**
	Bansang	256	
	Jahaly	7	
	Janjanbureh	22	

Table 3 Form corresponding queries (09 September 2019 to 31 May 2022)

Site	Clinic site	Total by site	Total
URR	Basse	29	89
	Demba Kunda	2	
	Fatoto	15	
	Garawol	9	
	Gambisara	19	
	Koina	12	
	Sabi	3	
CRR	Brikamaba	6	231
	Bansang	192	
	Jahaly	4	
	Janjabureh	33	

Table 4 Form corresponding queries (09 September 2019 to 30 June 2022)

Site	Clinic site	Total by site	Total
URR	Basse	7	8
	Demba Kunda	0	
	Fatoto	0	
	Garawol	0	
	Gambisara	0	
	Koina	1	
	Sabi	0	
CRR	Brikamaba	1	56
	Bansang	52	
	Jahaly	1	
	Janjabureh	2	

metrics to standardize activities across countries and centers involved in the studies. Data were usually sent to a central coordinating center that remotely monitors the data. While our QA/QC system was pivoted on the use of trained senior study staff to undertake regular supervisory visits at data collection points, the aforementioned studies employed different methods such as the use of a Site Evaluation Report to assess data quality [4], use of data coordinating centers to generate weekly or monthly reports which is shared with the sites [9–11] or relied on clinical monitors [12] for their QA/QC programs. Similar to PVS, one study reported a real-time online review of data [4]. Investing in staff training on using electronic data collection tools, such as the EMR is a key step to automating QC and preventing errors. Our experience throughout the trial has led to the observation that adequately trained staff were less likely to generate queries than poorly trained staff. Strategies to ensure that on-field QC checks are conducted at the end of each day's data collection exercise help to detect errors in real-time and enable prompt resolution. Real-time review of data during our weekly progress meetings has been a key strategy in ensuring comparability of manually generated clinic and laboratory reports and data entered in the EMR. This ensures the accuracy and completeness of data entered in the EMR. Errors and discrepancies identified in these weekly meetings are documented as action points for prompt resolution. Resolution of these queries minimizes issues and data inconsistencies detected during our formal monthly QA meetings. The monthly QA meetings enable us to review the performance of the clinical, field, laboratory, and data teams. QA meetings provide a more in-depth view into data quality indicators requiring a larger number of samples to better observe trends in the QA. Feedback from site supervisory visits is provided during the QA meetings and this enables the team to consider performance improvement strategies to minimize future deficiencies and improve data quality. Meeting action points are communicated to staff via email and followed up by the data team for resolution. Frequent site supervisory visits have been an important component of our QA and QC program. It provides a useful means by which we can identify errors at the source, review individual staff members' performance and provide feedback, have an open and transparent discussion with staff concerning their project-related training needs, identify areas for further focussed training, and provide onsite training. The visits facilitate closer engagement between the supervisors and staff involved in data collection. This enhances communication and eventually contributes to the generation of high-quality data.

Considering the large size of our trial database, the large geographical catchment area covered, and different data collection applications, a robust and pragmatic internal QA and QC program pivoted on strong supervision at the data collection source was needed to ensure the generation of high-quality data, even during the Covid pandemic period [15]. No study has yet compared the different QA and QC procedures within a trial. Future studies should consider comparing the different QA and QC procedures in their designs to generate empirical evidence to guide researchers in decision-making regarding appropriate QA and QC procedures for trial design.

Abbreviations

CRR	Central River Region
eCRF	Electronic case report form
EPI	Expanded Program on Immunization
HDSS	Health and demographic surveillance system
PCV	Pneumococcal conjugate vaccine
PVS	Pneumococcal Vaccine Schedules study
SSP	Study-specific procedure
URR	Upper River Region
WHO	World Health Organization

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13063-024-08677-7.

> Additional file 1. PVS SSP Index.
>
> Additional file 2. PVS Health Facility Supervisory Checklist.
>
> Additional file 3. PVS laboratory monitoring checklist.
>
> Additional file 4. Overview of PVS QA/QC activities.

Authors' contributions
GM is the Principal Investigator of PVS. IO conceived and wrote the first draft of the manuscript. All authors read, contributed to, and approved the final manuscript.

Data availability
The dataset(s) supporting the conclusions of this article is(are) included within the article (and its additional file(s)).

Ethics approval and consent to participate
The PVS trial was approved by the Gambia Government/MRC Joint Ethics Committee (ethics reference 1577) and by the LSHTM Ethics Committee (ethics reference 14515).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no interests.

Author details
[1]Medical Research Council Unit, The Gambia at London School of Hygiene & Tropical Medicine, Fajara, The Gambia. [2]Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK. [3]Regional Health Directorate, Upper River Region, Ministry of Health, Basse, The Gambia. [4]Regional Health Directorate, Central River Region, Ministry of Health, Bansang, The Gambia. [5]Murdoch Children's Research Institute, Melbourne, Australia. [6]Department of Paediatrics, University of Melbourne, Melbourne, Australia.

References
1. Knatterud GL, Rockhold FW, George SL, Barton FB, Davis C, Fairweather WR, et al. Guidelines for quality assurance in multicenter trials: a position paper. Control Clin Trials. 1998;19(5):477–93.
2. Meinert CL. ClinicalTrials: design, conduct and analysis. USA: OUP; 2012.
3. Organization WH. Improving data quality: a guide for developing countries. 2003.
4. Sandman L, Mosher A, Khan A, Tapy J, Condos R, Ferrell S, et al. Quality assurance in a large clinical trials consortium: The experience of the Tuberculosis Trials Consortium. Contemp Clin Trials. 2006;27(6):554–60.
5. Clarke DR, Breen LS, Jacobs ML, Franklin RC, Tobota Z, Maruszewski B, et al. Verification of data in congenital cardiac surgery. Cardiol Young. 2008;18(S2):177–87.
6. Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. Epidemiol Rev. 1998;20(1):71–80.
7. Gutierrez JB, Harb OS, Zheng J, Tisch DJ, Charlebois ED, Stoeckert CJ Jr, et al. A framework for global collaborative data management for malaria research. Am J Trop Med Hyg. 2015;93(3 Suppl):124.
8. Houston ML, Martin A, Probst Y. Defining and developing a generic framework for monitoring data quality in clinical research. InAMIA Annual Symposium Proceedings. American Medical Informatics Association. 2018;2018:1300.
9. Johnson MR, Raitt M, Asghar A, Condon DL, Beck D, Huang GD. Development and implementation of standardized study performance metrics for a VA healthcare system clinical research consortium. Contemp Clin Trials. 2021;108: 106505.
10. Malmstrom K, Peszek I, Botto A, Lu S, Enright PL, Reiss TF. Quality assurance of asthma clinical trials. Control Clin Trials. 2002;23(2):143–56.
11. Warden D, Rush AJ, Trivedi M, Ritz L, Stegman D, Wisniewski SR. Quality improvement methods as applied to a multicenter effectiveness trial—STAR* D. Contemp Clin Trials. 2005;26(1):95–112.
12. Richardson D, Chen S. Data quality assurance and quality control measures in large multicenter stroke trials: the African-American Antiplatelet Stroke Prevention Study experience. Trials. 2001;2:1–4.
13. Mackenzie GA, Osei I, Salaudeen R, Hossain I, Young B, Secka O, et al. A cluster-randomised, non-inferiority trial of the impact of a two-dose compared to three-dose schedule of pneumococcal conjugate vaccination in rural Gambia: the PVS trial. Trials. 2022;23(1):1–23.
14. Pneumonia Vaccine Trial Investigators. Group. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children. Geneva: Department of Vaccines and Biologicals. World Health Organization. 2001.
15. Hossain I, Osei I, Lobga G, Wutor BM, Olatunji Y, Adefila W, et al. Impact of the COVID-19 pandemic on a clinical trial of pneumococcal vaccine scheduling (PVS) in rural Gambia. Trials. 2023;24(1):271.