

Ensuring Quality in AFRINEST and SATT

Clinical Standardization and Monitoring

Stephen N. Wall, MD,* Corinne I. Mazzeo, MHS,* Ebuloluwa A. Adejuyigbe, MB ChB, FMCPaed, † Adejumoke I. Ayede, MB BS, MSc, FMCPaed, ‡ Rajiv Bahl, MB BS, MD, PhD, § Abdullah H. Baqui, MB BS, DrPH, ¶ William C. Blackwelder, PhD, || Neal Brandes, MHS, ** Gary L. Darmstadt, MD, MS, †† Fabian Esamai, MB ChB, MMed, MPH, PhD, ‡‡ Patricia L. Hibberd, MD, PhD, §§ Marian Jacobs, MB ChB, ¶¶ Jerome O. Klein, MD, ||| Kasonde Mwinga, MBChB, MMed (Paed), MPH, *** Nigel Campbell Rollins, MB BS, PhD, § Haroon Saloojee, FCPaed, ††† Antoinette Kitoto Tshefu, MD, PhD, MPH, ‡‡‡ Robinson D. Wammanda, MB BS, FWACP Paed, §§§ Anita K. M. Zaidi, MB BS, SM, ¶¶¶ and Shamim Ahmad Qazi, MB BS, DCH, MSc, MD§

Background: Three randomized open-label clinical trials [Simplified Antibiotic Therapy Trial (SATT) Bangladesh, SATT Pakistan and African Neonatal Sepsis Trial (AFRINEST)] were developed to test the equivalence of simplified antibiotic regimens compared with the standard regimen of 7 days of parenteral antibiotics. These trials were originally conceived and designed separately; subsequently, significant efforts were made to develop and implement a common protocol and approach. Previous articles in this supplement briefly describe the specific quality control methods used in the individual trials; this article presents additional information about the systematic approaches used to minimize threats to validity and ensure quality across the trials.

Methods: A critical component of quality control for AFRINEST and SATT

was striving to eliminate variation in clinical assessments and decisions regarding eligibility, enrollment and treatment outcomes. Ensuring appropriate and consistent clinical judgment was accomplished through standardized approaches applied across the trials, including training, assessment of clinical skills and refresher training. Standardized monitoring procedures were also applied across the trials, including routine (day-to-day) internal monitoring of performance and adherence to protocols, systematic external monitoring by funding agencies and external monitoring by experienced, independent trial monitors. A group of independent experts (Technical Steering Committee/Technical Advisory Group) provided regular monitoring and technical oversight for the trials.

Conclusions: Harmonization of AFRINEST and SATT have helped to ensure consistency and quality of implementation, both internally and across the trials as a whole, thereby minimizing potential threats to the validity of the trials' results.

Key Words: community-based research, quality assurance, trial monitoring

(*Pediatr Infect Dis J* 2013;32:S39–S45)

Accepted for publication June 5, 2013.

From the *Saving Newborn Lives, Save the Children, Washington, DC; †Department of Paediatrics and Child Health, Obafemi Awolowo University, Ile-Ife, Nigeria; ‡Department of Pediatrics, College of Medicine, University of Ibadan, Ibadan, Nigeria; §Department of Maternal, Newborn, Child and Adolescent Health, World Health Organization, Geneva, Switzerland; ¶Johns Hopkins Bloomberg School of Public Health; ||University of Maryland, Baltimore, Maryland; **United States Agency for International Development, Washington, DC; ††Bill and Melinda Gates Foundation, Seattle, Washington; ‡‡Department of Child Health and Paediatrics, School of Medicine, Moi University, Eldoret, Kenya; §§Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts; ¶¶University of Cape Town, Cape Town, South Africa; |||Boston University School of Medicine, Boston, Massachusetts; ***Health Promotion Cluster, World Health Organization Regional Office for Africa, Democratic Republic of Congo; †††University of the Witwatersrand, Johannesburg, South Africa; ‡‡‡University of Kinshasa School of Public Health, Kinshasa, Democratic Republic of Congo; §§§Department of Paediatrics, Ahmadu Bello University Teaching Hospital, Ahmadu Bello University, Zaria, Nigeria; and ¶¶¶Aga Khan University, Karachi, Pakistan.

The views and opinions expressed in this article are those of the authors and not necessarily the views and opinions of the World Health Organization, Save the Children Federation, Inc, or the United States Agency for International Development.

This journal article was supported with funds provided by Save the Children Federation, Inc's Saving Newborn Lives Program and the World Health Organization through separate grants from the Bill & Melinda Gates Foundation. Its contents are solely the responsibility of the authors and do not necessarily reflect the views of Save the Children, the World Health Organization, or the Bill & Melinda Gates Foundation.

Address for correspondence: Stephen N. Wall, MD, Saving Newborn Lives, Save the Children, 2000 L Street NW, Suite 500, Washington, DC 20036. E-mail: swall@savechildren.org.

Copyright © 2013 by World Health Organization, United States Agency for International Development and Bill and Melinda Gates Foundation (WHO, USAID, and BMGF). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License, where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

ISSN: 0891-3668/13/3209-0S39

DOI: 10.1097/INF.0b013e31829ff801

To address the priority research needed to identify effective alternative, simplified antibiotic regimens for community-based management of newborn sepsis, 3 clinical trials were conceived and implemented, as previously described.¹ Each of these trials was designed as a randomized open-label trial to test the equivalence of simplified antibiotic regimens compared with the more complex standard regimen of 7 days of parenteral antibiotics. Trials known as the Simplified Antibiotic Therapy Trials (SATTs) were designed separately; through a 2008 consultation convened by the Bill and Melinda Gates Foundation (BMGF) and subsequent meetings, the separate study protocols were harmonized as a single, common protocol to be implemented in each country. Implementation of SATT started first in Bangladesh in July 2009 followed by Pakistan in January 2010. The African Neonatal Sepsis Trial (AFRINEST) was subsequently initiated in 2010 and was designed to use the same protocol (with additional arms) and outcomes as SATT in sub-Saharan Africa [the Democratic Republic of the Congo (DRC), Kenya and Nigeria].

The main objective of SATT and AFRINEST was to evaluate simplified alternative antibiotic regimens of antibiotics that could subsequently be recommended for community-based management of serious infections in newborns and young infants. Eligible participants include newborns and young infants (0–59 days old) with suspected serious infection, whose families do not accept or comply with referral for hospital-based treatment. The trial methods have been described in detail elsewhere.^{2–5} Although these research trials were originally conceived and funded separately, subsequent efforts

were undertaken to harmonize the trials to ensure consistency and quality, both internally and across the trials as a whole.¹ The Bangladesh SATT, developed by the Johns Hopkins University and Dhaka Shishu Hospital, was originally funded by United States Agency for International Development (USAID). The Pakistan SATT was developed by Aga Khan University (Karachi) and was funded by the Saving Newborn Lives initiative (SNL) of Save the Children with support from BMGF, with additional funding provided by World Health Organization (WHO) and USAID. AFRINEST was originally conceived and developed by WHO in consultation with SNL and USAID with funding from BMGF. Through a systematic proposal development process, principal investigators (PIs) of AFRINEST were selected from Nigeria (University of Ibadan, Obafemi Awolowo University and Ahmadu Bello University), DRC (Kinshasa School of Public Health) and Kenya (Moi University).

Previous articles in this supplement¹⁻⁵ have included information on the unique internal quality control processes that were used within each separate trial, as well as the development of a common protocol (including standardized definitions and terminologies) for the trials as a whole. This article focuses on the quality assurance: it presents the methods used for quality control across SATT and AFRINEST and additional information regarding efforts to harmonize the trial-specific internal monitoring procedures.

QUALITY CONTROL METHODS IN AFRINEST AND SATT

Potential Threats to Validity

During the planning and design stages of AFRINEST and SATT, study teams identified inherent design-related issues that had the potential to threaten the quality of the studies and validity of results. Significant efforts were made to develop systematic approaches to minimize these challenges and ensure quality of implementation. A primary concern was that the enrollment and treatment outcome criteria rely on clinical judgment, as opposed to objective measures, such as laboratory confirmation of severe infection or death as the primary outcome. Reliance on subjective determination of clinical signs of severe infection presents the potential for misidentification of clinical signs of severe infection and of enrolling patients with different types of illnesses and infections. Were this threat not effectively mitigated by quality control efforts, otherwise strong evidence of antibiotic regimen equivalence might be considered inconclusive. The determination of treatment failure, with the exception of death, also relies on clinical assessment; this poses a similar threat to conclusive evidence of equivalence. In addition, the trials are open-label, meaning that study personnel and participants are not masked to treatment assignment and thus know which therapeutic regimen an infant is receiving. The nature of these trials did not allow for masking the treatment regimen. Placebo injections would not be ethical, and sham injections would be potentially unethical and logistically problematic. The PIs, donors (BMGF and USAID) and trial technical steering committee (TSC) and technical advisory group (TAG) therefore undertook rigorous quality control methods to mitigate these and other potential threats, to internal validity within and across these trials.

Ensuring Quality

Development of SATT Quality Control Measures

Because the 2 Asia SATTs were separately funded and administered, special efforts were undertaken to coordinate the studies' implementation, technical oversight, data management and quality control. SNL was specifically funded by BMGF to organize a mechanism of technical oversight and monitoring for the 2 SATTs in Asia, using independent technical experts with experience

in design and implementation of clinical trials in developing countries, pediatrics, infectious diseases, ethics and biostatistics. This mechanism was mutually agreed to and planned by the PIs, funding agencies (SNL and USAID), WHO and BMGF; thus, a TSC was formed, consisting of independent experts, PIs, WHO, BMGF, USAID and SNL. A major initial concern was the need for standardized trial monitoring across the 2 SATTs to ensure uniform and high-quality implementation and internal monitoring procedures as much as possible. The TSC first considered whether an independent contract research organization (CRO) would be able to provide the necessary cross-site monitoring functions. The TSC held discussions with potential CROs and also specifically identified, together with the PIs, the particular issues that needed to be addressed to robustly mitigate the major threats to validity. The conclusion from these investigations and discussions was that CROs, in general, were not the most appropriate mechanism for monitoring in these trials. This was based on the limited capacity of CROs to specifically monitor the clinical standardization and quality control required for trials that rely on clinicians' skills and judgment to determine enrollment eligibility and treatment outcomes. The SATT TSC identified the following as key to quality control within and across sites:

1. Development and use of standardized protocols for clinical assessments for enrollment eligibility and treatment outcomes.
2. Development and use of standardized internal study team monitoring procedures to routinely assess and refresh clinical skills and judgments of staff.
3. Frequent and standardized monitoring visits by TSC members, using WHO site visit monitoring guidelines and checklist.

The SATT TSC used an external clinical monitoring expert, who was a trained pediatrician, to independently review the sites and make specific recommendations to standardize and harmonize all aspects of study implementation, clinical skills and judgments, and quality control procedures. These visits resulted in modification of standard operating procedures (SOPs) in both SATTs, particularly standardization across sites of clinical assessment and internal quality control methods.

Development of AFRINEST Quality Control Measures

AFRINEST is a single trial funded entirely by BMGF, with 5 sites and PIs in 3 sub-Saharan African countries, led and coordinated by a technical team at WHO. From the outset, AFRINEST used the common protocol that was developed initially for SATT. Modifications in implementation were necessary due to contextual factors, including differences in health delivery platforms between SATT Asia sites and those in the AFRINEST locations. As noted in the methods of previous articles,^{4,5} additional antibiotic regimens were tested in AFRINEST, including a simplified regimen for young infants with fast breathing only. The quality control measures instituted for AFRINEST were based on the development of SOPs, uniformly applied across all sites, including standardized training, WHO monitoring visits, assessment of clinical skills, refresher training and use of an external monitor for independent review.

SATT and AFRINEST Common Approaches to Quality Control

A critical component of quality control for AFRINEST and SATT was striving to eliminate variation in clinical assessments and decisions regarding eligibility, enrollment and treatment outcomes. While each of the trials adopted procedures specific for the setting, there was a uniform effort across the trials to assure that each "treatment failure" assignment was correct. In Bangladesh and Pakistan,

the outcome assessor may know the infant's treatment assignment. Infants who are identified as failing treatment (during routine follow-up) receive a second assessment by a different study physician; this second assessor is blinded to the infant's treatment assignment and history. If the 2 assessments disagree, the decision is referred to a supervising senior physician who makes the final determination. The AFRINEST sites use a separate "independent outcome assessor:" a nurse who is responsible for performing the outcome assessment and who is blinded to the infant's treatment assignment. Ensuring appropriate and consistent clinical judgment was accomplished through a combination of approaches, applied across the trials, including training standardization, routine (day-to-day) internal monitoring, systematic external monitoring by funding agencies and external monitoring by experienced, independent trial monitors.

Training and Standardization

AFRINEST and SATT use different cadres of health workers [including physicians, nurses and community health workers (CHWs)] and have implemented standard training protocols to ensure common understanding and clinical judgment among and across all cadres.

In AFRINEST, an initial training-of-trainers was conducted to train the PIs; the PIs then led trainings at each study site for local study staff. WHO technical monitors visited each site as soon as the training of health workers was underway to conduct a standardization exercise (Appendix 1) to check clinical assessment skills of the study staff. Those staff who did not meet preset criteria in the standardization exercise were immediately retrained; if they did not perform well after retraining, they were withdrawn from the study. Staff who met the performance criteria were assigned to roles within the study based on the knowledge and skills demonstrated during the standardization exercise (eg, the highest performing nurses were assigned to be independent outcome assessors). Regular refresher trainings are conducted by the PIs every 3–6 months; additional standardization exercises are held at 3- or 6-month intervals (depending on the type of study staff). Results of the standardization exercises are input into a spreadsheet (Microsoft Excel) and scores outside the standard range are marked; those staff who do not perform well (eg, not conducting the required number of follow-up visits or performing poorly on clinical assessments) are retrained on specific areas in which their performance is low and reassessed before they are allowed to continue.

In SATT, intensive training was also conducted at the start of the trials. Refresher training sessions and standardization exercises are held regularly to ensure consistency in eligibility and treatment outcome evaluation. In Bangladesh, refresher trainings and standardization exercises are conducted quarterly; weekly meetings of the study staff also provide opportunities for "continuing education" through review of randomly selected cases of severe infection and discussion of any discrepancies in the assessors' identification of clinical signs (Appendix 2). In Pakistan, refresher trainings and standardization exercises are conducted on a monthly basis; the primary goal of these activities was to reinforce knowledge and skills related to clinical assessment (Appendix 3). Video recordings of selected infants with and without clinical signs of severe infection are used to assess and refresh skills.

Internal Monitoring

SOPs were developed for AFRINEST and SATT to guide the operational and clinical activities of the trials; this included SOPs for internal monitoring. The main objective of internal monitoring was to verify that the study protocol and SOPs are being followed correctly. As with other aspects of the trials, internal monitoring approaches were harmonized across all sites to ensure similar high

levels of quality and safety. In all cases, feedback from the routine internal monitoring is presented and discussed with the appropriate members of the study team on an ongoing, real-time basis.

In AFRINEST, PIs conduct routine monitoring visits to study sites and use a standard checklist developed by WHO. These visits include checking study equipment and supplies (eg, weighing scales), as well as conducting supervision activities for study staff. Scheduled and unscheduled visits are performed by PIs and field supervisors; a standardized checklist is used to monitor the skills of CHWs, treatment nurses and outcome assessors regarding identification of danger signs, classification, eligibility criteria, randomization and consent procedures, treatment and injection technique. At a minimum, 5% of a nurse's/CHW's visits must be directly observed; outcome assessment nurses must be observed during at least 20% of their visits. If errors or gaps in skills are observed during a supervisory visit, immediate feedback is given to the study team member.

In the SATT Bangladesh trial, a SOP was developed to guide the operational and clinical activities of the trial; this included SOPs for internal monitoring. The main objective of internal monitoring was to verify that the study protocol and SOP were being followed correctly. As with other aspects of the trial, internal monitoring approaches were harmonized across all sites to ensure similar levels of quality and safety. In all cases, feedback from the routine internal monitoring was presented and discussed with the appropriate members of the study team on an ongoing, real-time basis.

Site visits were conducted by study coordinators and by investigators (and/or consultant experts) in the SATT Bangladesh trial. The study coordinator is allocated 2 days every month for each site for this purpose. One day is randomly selected for each site for the monitoring visit, and the visit is completed on the next day. The coordinator uses a standard monitoring checklist (developed by WHO) to guide the visit, which includes evaluating study management, reviewing subject records and observing key study procedures. Findings are shared with the local study team and with investigators. A trip report is submitted after each monitoring visit. One random day is selected per month for each site to be visited by study investigators/consultants for a monitoring and supervision visit. The first 5 cases on the selected day are selected for a study physician to assess (the clinical component only) and the investigators/consultants to observe. Only disagreements are recorded, and the reason for disagreements are immediately addressed and discussed. These random visits are planned to ensure that all study physicians are subjected to a validation exercise once in a 3-month period. For clinical signs about which the investigators disagreed with study physicians' assessment, a database is kept; these signs are also emphasized during refresher training.

In SATT Pakistan, routine internal monitoring is conducted on a daily basis by 2 dedicated clinician supervisors who observe the screening and enrollment process, randomization and the delivery of therapy. They also observe approximately 20% of interviews conducted by study physicians, perform spot checks and formally reinterview 10% of participants interviewed by study physicians. Finally, supervisors check case files for errors, inconsistencies and missing values. Additional monitoring visits are conducted by an "internal auditor:" a physician and monitoring expert who is not associated with the trials. The auditor performs a random visit once a month to observe and review clinical practices in the study; a standardized checklist is used and feedback is provided directly to the PI, research supervisor and field team.

External Monitoring

Both AFRINEST and SATT receive structured external monitoring visits from WHO monitors and other members of their

respective TAG (AFRINEST) or technical steering committee (SATT). WHO technical staff, following a standard WHO site monitoring guideline/checklist (Appendix 4), visit each AFRINEST site at least once every 6 months. A detailed structured review of study implementation is conducted at each visit, including recruitment rate, clinical practices (eg, identification of danger signs and critical signs, detection of treatment failure), data management and overall study procedures. All study documentation is checked, and a proportion of completed case report forms (including all reports of treatment failure and adverse events) is reviewed. The recommendations arising from the site visit are discussed with the PIs at the sites and followed up for actions taken and dates. At least once per year, the SATT Bangladesh and Pakistan sites receive structured external monitoring visits from a WHO technical expert, either alone or accompanied by SNL/SC technical staff and independent TSC member(s). Additional ad hoc monitoring visits by external TSC members are conducted for SATT Bangladesh, mostly in conjunction with TSC/Data Safety Monitoring Board (DSMB) meetings, which are frequently held in Dhaka. Additional ad hoc visits are intended for SATT Pakistan but are subject to security restrictions for non-Pakistanis. The format for SATT external monitoring visits is similar to AFRINEST: recommendations and written feedback are provided, and the PI subsequently updates the TSC on actions taken in response to each recommendation. The same WHO site monitoring checklist is used to structure the SATT monitoring process and to standardize across all sites.

In addition, all sites have been visited by an independent external trial monitor who examined study implementation and adherence to protocols and Good Clinical Practices standards. For SATT sites (Bangladesh and Pakistan), the external monitor conducted an independent assessment of clinical and quality control procedures and made recommendations on how to standardize or harmonize all such processes. PIs followed these recommendations to implement uniform internal monitoring and quality control procedures. In AFRINEST, an independent external monitor visited each site, provided an assessment using the same protocol as WHO monitors (Appendix 4), made recommendations for each separate sites that were implemented by PIs and then checked for by WHO during routine monitoring.

Technical Advisory Groups

Two external advisory groups were set up for the trials: AFRINEST formed a TAG and SATT formed a TSC to oversee the 2 trials in Bangladesh and Pakistan. These bodies are responsible for providing broad technical oversight for the trials. In the case of SATT, the TSC adds unique value by facilitating coordination and harmonization for the 2 separate studies. The SATT TSC has played an active role to provide coordinated oversight for the 2 separate trials, as explained above. This includes routine external monitoring visits, including visiting homes and clinics to observe enrollment and management of patients; the TSC then makes recommendations based on its observations and follow-up to ensure recommendations appropriately reviewed and implemented by each trial. The AFRINEST TAG is not needed to provide coordination of technical oversight, as this was done by WHO as the coordinating agency for the trial. The AFRINEST TAG has been convened for annual review meetings but has not engaged in direct field-level monitoring in the same way that it is done by the SATT TSC. Membership of the TAG and TSC includes experts from a variety of areas, including clinical pediatrics, neonatology and infectious diseases; global clinical trials (including ethics); epidemiology and biostatistics; and global health policy. There was an intentional overlap in the TAG and TSC membership; a number of individuals were selected to serve on both groups to facilitate greater harmonization between the trials. The SATT TSC holds quarterly and as-needed conference calls to review quarterly data

and discuss technical issues, as well as annual in-person meetings. The AFRINEST TAG holds teleconferences on an as-needed basis throughout the year, as well as annual in-person meetings.

Data Safety Monitoring Boards

Two DSMBs were formed, 1 for AFRINEST and 1 for SATT (Bangladesh and Pakistan combined), each with 5 members: 1 epidemiologist, 1 statistician and 3 clinicians/researchers. Members of the DSMB are not involved in the design or conduct of the trials. These groups are tasked with monitoring and assessing the safety of the trials (based on reports); responsibilities include ensuring patient safety within each site, setting up “stop rules” based on ethical and public health considerations, reviewing interim analysis data and recommending appropriate actions to be taken by the PIs. The DSMBs meet annually with the PIs and TSC/TAG and communicate among members via e-mail and teleconferences as needed to review data and make recommendations to PIs.

Data Management

There are 3 levels of data management and quality assurance in AFRINEST and SATT: local data management and quality control practices at study sites and within the individual trials; technical oversight, review and quality control for data management from the TSC/TAG; and external monitoring visits and high-level data management, review and quality checks performed by the data management team at the London School of Hygiene and Tropical Medicine (LSHTM). Methods for site-level and individual trial-level data management have been described elsewhere.²⁻⁵ The team at LSHTM performs a standard set of range and consistency checks for all data sent from the individual sites. Data are archived at LSHTM, and reports are generated as required by IRBs, the AFRINEST Coordinator (WHO) and TAG, and the SATT TSC.

Routine external monitoring visits include review of study data and data management procedures. Specifically, database and CRFs are reviewed for all death and treatment failure cases, and a random sample is taken of approximately 10–20% of eligible and enrolled cases. Finally, LSHTM provides data management and analysis support for both AFRINEST and SATT. AFRINEST submits data to LSHTM on a monthly basis; these data are shared with WHO monthly and with the TAG annually. The SATT trial sites submit their data to LSHTM on a quarterly basis, and the data are shared with the TSC quarterly for review. LSHTM also provides routine data summaries of deaths, treatment failures and severe adverse events for regular review by each DSMB.

Strategic Planning Committee

The Strategic Planning Committee is composed of AFRINEST and SATT funding agency representatives and chairs of the TSC and TAG. Its primary purpose was to review the trials’ progress and to provide strategic direction for planned study analyses to answer questions of global policy relevance. The Strategic Planning Committee provides general strategic direction to maximize policy impact, based on efforts to ensure sound evidence, relevant evidence from both trials.

CONCLUSIONS

To ensure quality of implementation and to minimize potential threats to validity, AFRINEST and SATT adopted a variety of purposeful quality control approaches; these were similar or the same across the trials. Key quality control processes have included standardization of training efforts and assessment of clinical skills, routine internal monitoring of study protocol adherence, and external monitoring and oversight of study implementation by funders and independent experts.

REFERENCES

1. Qazi S, Wall S, Brandes N, et al. An innovative multi-partner research programme to effectively address detection, assessment and treatment of neonatal infections in low-resource settings. *Pediatr Infect Dis J.* 2013;32:S3–S6.
2. Baqui A, Saha SK, Ahmed ASMNU, et al. Safety and efficacy of simplified antibiotic regimens for outpatient treatment of clinical serious infection among neonates and young infants aged 0–59 days in Bangladesh: Design of a randomized controlled trial. *Pediatr Infect Dis J.* 2013;32:S12–S18.
3. Zaidi AKM, Tikmani SS, Sultana S, et al. Simplified antibiotic regimens for the management of clinically diagnosed severe infections in newborns and young infants in first level facilities in Karachi, Pakistan: Study design for an out-patient randomized controlled equivalence trial. *Pediatr Infect Dis J.* 2013;32:S19–S25.
4. AFRINEST Study Group. Simplified regimens for management of neonates and young infants with clinical severe infection in situations when hospital admission is not possible: Study protocol for a randomized, open-label equivalence trial. *Pediatr Infect Dis J.* 2013;32:S26–S32.
5. AFRINEST Study Group. Treatment of fast breathing in neonates and young infants with oral amoxicillin compared to injectable penicillin-gentamicin combination: Study protocol for a randomized, open-label equivalence trial. *Pediatr Infect Dis J.* 2013;32:S33–S39.

APPENDIX 1. AFRINEST CLINICAL ASSESSMENT STANDARDIZATION PROCEDURE

1. Identify 5–10 eligible young infants in a community. Ask up to 5 CHWs (one by one) to assess 1 young infant being observed by the trained facilitator, both using Form 2 to record all findings. This process is repeated with each CHW by the facilitator.
2. Identify 5–10 eligible young infants in a health facility. Ask up to 5 enrollment or treatment or health center nurses/supervisors/coordinators (one by one) to assess 1 young infant being observed by the trained facilitator, both using Form 3 to record all findings. This process is repeated with each person by the facilitator.
3. Then the data are entered in a simple excel sheet. For the purpose of this exercise if the facilitator and the health worker differ in their assessment, it is categorized as an “error.” For respiratory rate ± 2 breaths per minute; for weight ± 0.1 kg and for temperature $\pm 0.1^\circ\text{C}$ is acceptable. If the difference between the facilitator and the HW is more than the preceding differences, then that is considered an error. At the bottom of the sheet, we total the errors over the observations (see example).
4. If the time is short and the young infant cannot be exposed for long due to sickness, the history can be taken by 1 CHW or nurse with other CHWs or nurses/supervisors recording the same information, but the examination has to be done separately at the same time with the trained supervisors.
5. Such standardization exercises were done with all health workers. Errors were improved by further practice and refresher training if necessary. A 5% error may be acceptable before the health worker starts enrolling patients. Workers who did not improve were replaced.

APPENDIX 2. SATT BANGLADESH CLINICAL ASSESSMENT STANDARDIZATION PROCEDURE

Eligibility Assessment Standardization at Hospital

In a calendar month, 2 days were randomly selected for eligibility assessment standardization in each of the 5 study sites (an estimated 6% of the total screening). Two study physicians independently assessed all young infants for eligibility on the randomly

selected days. The assessment of the first physician was entered in regular database. Second independent assessment was recorded on a separate data form. Analysis by each sign and overall determination of eligibility was conducted on weekly basis to find discrepancies and was discussed in weekly meeting as part of continuing training. Refresher training for standardization of assessments was conducted quarterly. In refresher training, special attention was given to the signs for which there were more discrepancies between the 2 independent physicians.

Follow-up Standardization at Home

Independent assessment of 5% of all home follow-ups was conducted by a second physician. An excel program was used to randomly select 5% of all follow-ups. Random numbers for follow-up IDs were generated every morning. The study coordinator assigned 2 study physicians to follow-up the randomly selected cases and organized the field follow-up schedule to ensure about an hour gap between the 2 assessments. Data of first physician's follow-up were entered into regular database, and second physician's assessment was entered into separate database. Data analysis was done weekly, and discrepancies were discussed in the weekly meeting. The discrepancies were taken into account during quarterly refresher training to standardize the assessments.

Ascertainment of Treatment Failure Standardization

All infants found to met clinical treatment failure criteria by study physicians on routine follow-up was designated as “provisional” treatment failures and were transported to hospital, accompanied by study personnel. At the hospital, the infant underwent a repeat examination by a second study physician. To the extent possible, the second physician assessor was blinded to the treatment allocation and prior history of the infant. If the second assessment supported the ascertainment of treatment failure, the case was considered a “confirmed” treatment failure. If the second medical assessment did not lead to the ascertainment of treatment failure (ie, disagreed with first assessment), the case was referred to a supervising senior physician. The decision of the senior physician was considered as final.

APPENDIX 3. SATT PAKISTAN CLINICAL ASSESSMENT STANDARDIZATION PROCEDURE

1. Initial training
 - 1.1 Initial training consisted of 2 days classroom lectures, 2 days in tertiary care hospital and 2 days at field sites (total 6 days training).
 - 1.2 In classroom setting trial protocol, standard operation procedures, case report form, drug and doses, outcome definition and ascertainment discussed and 20 videos of clinical signs are shown.
 - 1.3 In tertiary care hospital, each physician examines 15 infants for presence or absence of clinical signs, being observed by a trained facilitator (pediatric consultant).
 - 1.4 In field sites, physicians examine 5 infants under supervision of research supervisor, including anthropometry measurements and record their findings, which are compared with supervisor's findings.
2. Monthly standardization exercises
 - 2.1 Ten videos (including both sick and normal young infants) are shown to study physicians from our video database.

- 2.2 Study physicians recording their findings on an answer sheet. Answer sheet checked by senior clinician and feedback provided.
- 2.3 Physicians who get 4 or more incorrect identifications on 3 consecutive tests are provided with full refresher training (see below).
3. Biannual refresher training and standardization:
 - 3.1 One week refresher training is conducted twice a year.
 - 3.2 Trial protocol and standard operation procedures reviewed.
 - 3.3 20 videos of clinical signs in young infants are shown for identification of signs.
 - 3.4 Individual assessment test includes clinical signs identification as well as protocol.

APPENDIX 4. WHO CHECKLIST FOR SITE MONITORING VISITS

Prerequisite

- Pre-agreed agenda and timing of the visit with the investigators.

Actual visit

- Meeting with PI and other study staff during the visit and discuss aspects of the study progress.

The following are carried out by review/observation/checking:

1. STUDY MANAGEMENT REVIEW

- Documentation:
 - Updated protocol with all the approved amendments
 - Procedures manual/standard operating procedures
 - Consent forms
 - Communications with the local IRB
 - IRB approval
- Storage of documents:
 - Designated data cabinets
 - Adequate storage of case report forms (CRF)
- Progress of the study/enrollment:
 - Expected recruitments
 - Actual recruitments
 - Other relevant aspects
 (include justification if discrepancy >10%)
- Screening of patients and enrollment records:
 - Completeness
 - Adequacy of storage
- Maintenance of screening calendars/forms:
 - Completeness
 - Adequacy of storage
- Maintenance of follow-up calendar
- Proper date of follow-up (window period) adequate implementation
- Place where patients report for follow-up (health center/home):
 - Adequacy of facility
 - Opening time
 - Staff
- CRF being sent regularly to data management centre in country:
 - Frequency of submission
 - Completeness, backlogs
- Data management (data entry progress, due, done, entered, etc):
 - No. of forms due
 - No. of forms completed

- No. of forms entered
- No of forms still to be entered
- Problems with data entry
- Regular transmission of electronic data to the central data coordination center:
 - Expected frequency
 - Actual frequency
- Security of data (password protection):
 - Sites of storage
 - Frequency of backups
 - Date of last backup
- Communication with central DCC and WHO:
 - Frequency
 - Date of last communication
- Inventory and storage and protection of study medicines:
 - Definition of responsible staff
 - Frequency of inventory
 - Date of last inventory
 - Adequacy of facility (security, temperature, stocktaking)
- Dispensing record for study medicines
- Collection of used needles, syringes, medicine vials and bottles and proper storage/disposal
- Fund utilization:
 - Expected expenditure
 - Actual expenditure
 - Explanation for differences >10%
- Situation of supplies/reordering of supplies

2. REVIEW OF SUBJECT RECORDS¹

- Check filled CRFs (10–20% new cases since the last visit)
- Check all (or at least 50% new cases since the last visit) treatment failure/outcome assessment and serious adverse events forms CRFs

3. OBSERVATION OF KEY STUDY PROCEDURES

- Screening, enrollment and randomization²
- Administering the consent form
- Management of patients (including preparation of study medicine and dispensing of medicines in accordance with the study protocol)
- Outcome assessment
- Consistency in measurements between various levels (CHW, study health worker, outcome assessment nurse and supervisor)
- Home visits to see some enrolled infants and interview mothers/caretakers⁴
- Follow-up of patients
- Completing follow-up information
- Communication between study staff and potential/enrolled patients' families
- Data entry, double data entry and cleaning
- Making backups
- Audit trail

4. STANDARDIZATION

- Review presence of training materials and training/retraining records as per standard operating procedures.
- Standardization records as per standard operating procedures.
- Interviews with a sample of data collectors to confirm that processes being used are according to the standard operating procedures.
- Organize a standardization session for key measurements and procedures during this visit.

5. SITE MONITORING REPORT

- Feedback and action for principal investigator
- Recommendations and actions for WHO
- Report prepared and submitted
 - Date and signature

NOTES

1. There are no omissions or discrepancies in the reports of specific data elements such as weight, age, clinical signs, correlation of

clinical signs with assessment and diagnosis, outcome assessment and treatment, randomization code, correlation of randomization code with treatment regimen, correlation of weight and treatment doses, adherence to treatment and serious adverse events.

2. During a site visit, there may not be any enrollment, randomization and treatment identification or administration of consent form. In that case, home visits are made to see some enrolled infants and interview mothers/caretakers.