BMJ Open

A protocol for examining effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in routine settings in Kenya

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016784
Article Type:	Protocol
Date Submitted by the Author:	13-Mar-2017
Complete List of Authors:	Malla, Lucas; Kenya Medical Research Institute - Wellcome Trust; University of Oxford Perera, Rafael; University of Oxford, Primary Health Care McFadden, Emily; University of Oxford, Primary Health Care English, Mike; University of Oxford, Nuffield Department of Medicine and Department of Paediatrics
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology
Keywords:	propensity score, multiple imputation, observational research, comparative effectiveness

A protocol for examining effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in routine settings in Kenya

Lucas Malla[1*] (lucas.malla@some.ox.ac.uk)

Rafael Perera-Salazar[2] (rafael.perera@phc.ox.ac.uk)

Emily McFadden[2] (emily.mcfadden@phc.ox.ac.uk)

Mike English[1, 3] (menglish@kemri-wellcome.org)

[1]Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom.

[3] Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya.

***Corresponding author:**

Lucas Malla

Introduction

WHO treatment guidelines are widely recommended for guiding treatment for millions of children with pneumonia every year across multiple low and middle income countries. Guidelines are based on synthesis of available evidence that provides moderate certainty in evidence of effects for forms of pneumonia that can result in hospitalisation. However, trials have included fewer children from Africa than other settings and it is suggested that African children with pneumonia have higher mortality. Thus despite improving access to recommended treatments and deployment with high coverage of childhood vaccines, pneumonia remains one of the top causes of mortality for children in Kenya. Establishing whether there are benefits of alternative treatment regimens to help reduce mortality would utilize pragmatic clinical trials. However, these remain relatively expensive and time consuming. This protocol describes an approach to using a new, large routine dataset as a potentially cheaper and quicker way to examine the comparative effectiveness of penicillin versus penicillin plus gentamicin in treatment of indrawing pneumonia. Addressing this question is important as although it is now recommended that this form of pneumonia is treated with oral medication as an outpatient it remains associated with non-trivial mortality that may be higher outside trial populations.

Methods and analysis

We will use a large routine dataset that captures data on all admissions to 13 Kenyan county hospitals. These data represent the findings of clinicians in practice and, because the system was developed for large observational research, pose challenges of non-random treatment allocation and missing data. To overcome these challenges this analysis will use a rigorous approach to study design, propensity score methods and multiple imputation to minimize bias.

Ethics and dissemination

This analysis will be conducted as part of the Kenyan Clinical Information Network project which has received ethical clearance from the Kenya Medical Research Institute.

Strength

- This study will be used as a platform to explore effectiveness of alternative treatments in routine care in a low income setting to improve health outcomes for children.

Limitation

- The analysis will be limited to the variables in the routine dataset – and therefore risk bias due to unmeasured key variables.
- The influence of any resulting bias, to alter results, will however be assessed through the use of alternative methods as instrumental variables.

Introduction

Kenya has developed and disseminated national treatment guidelines largely drawing on those of WHO for a number of childhood diseases including pneumonia (1, 2). These pneumonia guideline recommendations are based on synthesis of available evidence that provides moderate certainty in evidence of effects of treatments for forms of pneumonia that can result in hospitalisation (3). Such guidelines have been shown to be effective in reducing pneumonia related mortality and thus Kenyan clinicians are supposed to use them in routine practice to treat pneumonia (and other diseases) (4, 5). However, although the guidelines are based on the best available evidence, the evidence available from trials conducted in Africa remains limited (6). There has also been little thorough investigation of the effectiveness of treatments in non-trial populations in routine settings that may often differ from those enrolled in formal clinical trials. For example many children admitted with pneumonia may have co-morbidity that might exclude them from trials (7). These issues can prove problematic when making national guidelines where study generalisability can be contested (8).

The WHO and Kenyan pneumonia treatment guidelines are implicitly based on risk stratification of illness with children deemed at higher risk of severe illness and mortality offered broad spectrum antibiotic regimens and those at lower risk narrow spectrum antibiotics (2, 9-11). This risk stratification approach is operationalized by requiring clinicians to look for specific features in the clinical history and examination that are used to define illness severity and therefore recommended treatment (Panel 1). Previous studies conducted in Kenya have, however, indicated that clinicians do not always follow guideline recommendations in treating pneumonia (5). Variation from the guideline recommended approach can occur at the point of pneumonia severity assignment (clinicians do not follow the rules linking clinical signs and severity category) and at the point of treatment assignment

(clinicians do not follow the rules linking treatment and severity). This variability in treatment assignment provides the opportunity for comparative effectiveness evaluation if similar populations of children with pneumonia are prescribed different treatments. Clinicians may create such a situation by not following recommendations because they have inadequate

knowledge or if they believe (potentially contrary to the evidence) that certain treatments result in better health outcomes.

Panel 1: Pneumonia treatment algorithm

The pneumonia severity classification that was recommended by Kenyan guidelines up to March 2016 (10) (and previously by WHO guidelines (1)) defined the following three severity classes:

1. ***Very severe pneumonia:*** *If a child had either oxygen saturation less than 90% or central cyanosis or was grunting or unable to drink or not alert, then s/he was classified as having very severe pneumonia, put on oxygen and treated with a combination of gentamicin and penicillin.*

   {The new WHO (2) and Kenyan guidelines (10) renamed this class as "severe pneumonia" – and currently recommend treatment with a combination of ampicillin (or penicillin) with gentamicin plus oxygen}.
2. ***Severe pneumonia:*** *If a child had lower chest wall indrawing (but did not have any of qualifying signs for very severe pneumonia above) and was alert then s/he was to be classified as having severe pneumonia and be treated with benzyl penicillin only.*

   Note: *The term indrawing pneumonia is hereafter used in this protocol to define this category of children to avoid confusion.*
3. ***(Non – severe) Pneumonia:*** *If a child had none of the mentioned signs but had cough or difficulty breathing and a respiratory rate greater than or equal to 50 breaths/minute (for age between 2 and 11 months) or respiratory rate greater than or equal to 40 breaths/minute (for age above 12 months) then s/he was classified as having non severe pneumonia and treated with cotrimoxazole or amoxicillin if previously treated with cotrimoxazole.*

   {The current WHO and Kenyan guidelines collapsed severity classes 2 and 3 into one category referred to as "non –severe pneumonia". This group of patients are currently treated with oral amoxicillin – partly informed by a local trial (18)}.

In particular, a previous study showed that clinicians over-prescribed gentamicin, adding this to penicillin for the treatment of pneumonia characterized by lower chest wall indrawing but no other signs of severe illness instead of penicillin alone as was recommended (5)[1]. Therefore, this protocol is for a study that seeks to explore whether there is any benefit from adding gentamicin to penicillin in treating children with indrawing pneumonia. Such a benefit could accrue if bacterial causes of pneumonia that were previously (prior to introduction of new vaccines) proportionately less common (eg. *S. aureus* and gram negative bacteria) are now accounting for an increased proportion of pneumonia deaths – as in such cases, the addition of gentamicin might provide effective treatment for a broader spectrum of pathogens. Tackling this question is of importance as WHO have recently changed indrawing pneumonia treatment guidance based on trials that suggest equivalence of oral amoxicillin and injectable penicillin (13-16). New guidance recommends outpatient oral treatment for a population of children previously admitted to hospital (11). However, mortality from pneumonia has been reported to be higher in African settings (17) despite the increasing use of multiple vaccines spanning: measles, pertussis, HiB and pneumococcal conjugate vaccines. It remains possible therefore that for a small number of children a broader spectrum antibiotic regimen might be of benefit. This study addresses this question that has not been the subject of prior clinical trials.

Objectives

Primary

1) Experiment 1: To compare the effectiveness of injectable penicillin versus penicillin plus gentamicin (both injectable) in treatment of indrawing pneumonia; where severity level is imputed using data recorded on each child's clinical signs (hospitals

[1] The fact that inadequate knowledge in handling childhood pneumonia may result in inconsistent treatment allocation is supported by a survey conducted in seven developing countries showing that 56% of nurses and doctors had inadequate knowledge in managing pneumonia in children (12).

use a structured record form that supports recording of signs highlighted in guidelines) such that severity classification is consistent with guideline recommendations. This scenario will provide an evaluation of alternative therapies within a guideline class (where children have very similar clinical signs) and thus is the best mimic of a prospectively designed comparative evaluation in which clinicians stick to the rules of severity classification (see (18) for an example of a RCT in Kenya that this would be similar to – where classification is based on clinical signs). Recommended treatment for this disease classification was penicillin alone, treatment with combination therapy may therefore represent over-treatment. Alternatively the combination treatment that provides broader antimicrobial cover could provide an advantage in a small proportion of cases that would only be detected in moderately large studies – where the addition of gentamicin offers improved treatment for specific organisms not susceptible to penicillin alone.

Secondary

2) Experiment 2: To compare effectiveness of injectable penicillin versus penicillin plus gentamicin in treatment of indrawing pneumonia; where we use clinician assigned severity level. This experiment will provide a test of alternative therapies amongst those where clinicians used their own judgement (possibly including gut feeling) to classify and treat (19) and have on occasions (potentially) over-ridden or ignored the guideline recommendations. In this case although the same label of indrawing pneumonia is given to all, the treatment selected may be an indicator of perceived severity and there may be a potential bias as a result – and the propensity score distributions (see below) may help demonstrate this and in theory may overcome this potential bias. Here if there is no clinically relevant difference between treatments within a group of patients that reflects clinicians' actual classification decisions this

could reassure them that monotherapy with penicillin (or amoxicillin) would be acceptable.

3) Experiment 3: To compare effectiveness of injectable penicillin versus penicillin plus gentamicin in treatment of all cases of pneumonia admitted to hospital. This is an extension of the logic of experiment two. To date there have been no trials of penicillin alone compared with alternative combination therapies for all forms of inpatient pneumonia, and addressing this question may be relevant for two reasons. First, the population of children admitted with severe forms of pneumonia is now largely one that has received *H. influenzae* Type B and pneumococcal conjugate vaccines that have likely changed the aetiology of this illness. Second, if clinicians are poorly trained and unable to classify illness severity – resulting in non-adherence to guidelines - it would be useful to explore the potential impact of this across all levels of severity of pneumonia. This analysis has the largest numbers of subjects.

Experiment 1 will be primary as it most approximates a typical randomised trial where recruitment would be based on specified clinical signs. Experiment 2 will demonstrate what effects would be achieved if clinicians assign a category of indrawing pneumonia but are not necessarily adherent to guidelines[2] recommending use of specific clinical signs to assign illness severity and experiment 3 will explore effects of treatments across all potential severity classes of pneumonia in children admitted to hospital (even though the primary focus will be on indrawing pneumonia).

Methods and analysis

[2] Includes patients where clinicians have been non-adherent to clinical classifications recommended in guidelines.

To answer these three questions we will use the Kenyan Clinical Information Network (CIN)

Panel 2: Clinical Information Network

CIN was initiated to improve data availability from secondary care in paediatrics and as a model for demonstrating the value of routine data in improving quality of care in the county (formerly district) hospitals. These hospitals typically have a single paediatrician leading services predominantly provided by junior clinical teams. Data in these hospitals are collected prospectively post discharge by trained data clerks, guided by well-defined standard operating procedures, under close supervision by the hospital medical records department and the research team. It is worth noting that the research team has no personnel checking quality of clinical process and whether clinicians correctly document what they do. However, the patient record is the formal (and legal) document describing the clinical condition and management. These documents are used for data abstraction and they include patient files with standardized Paediatric Admission Record (PAR) forms, treatment sheets, discharge summary forms, laboratory reports and clinician notes. The collected data are used to assess documentation of history, physical examination, diagnosis, laboratory investigations, treatment and discharge plans. Feedback to hospitals as part of the CIN activities has helped improve the quality of clinical data (38). The description of hospital selection and their populations of patients is detailed in Ayieko (2015) (7).

dataset that provides routine data on all admissions to 13 Kenyan County hospitals (panel 2).

The analysis will proceed in two stages – design and outcome analysis as suggested by Rubin (2008) (20) as an objective way for analysing observational datasets.

Study Design

This will be an observational study conducting analyses of data routinely collected from hospital paediatric wards in Kenya's CIN. The design process for the three experimental scenarios will be similar and broadly consists of the following steps suggested in Rubin (2008) (20):

a) Definition of inclusion and exclusion criteria.

b) Understanding the pneumonia diagnosis and treatment assignment processes. This is to help understand key and auxiliary variables required for analysis.

c) Verification of sample size if sufficient for any meaningful analyses.

d) Creation of comparable treatment arms – which will be addressed analytically aiming to overcome non – random treatment assignment and deal with missing data.

e) Outcome analysis follows after conceptualisation of design in steps a – d.

a) Inclusion and exclusion

This analysis will include all children aged 2 – 59 months and will exclude children with any co-morbidity of HIV, meningitis, tuberculosis and or acute malnutrition as there are specific antibiotic treatment rules for these children that supersede those for pneumonia. Importantly therefore children with other co-morbidities such as mild anaemia, diarrhoea and malaria are not necessarily excluded from the analysis.

b) Understanding the diagnosis and treatment assignment rules for pneumonia paediatric patients

Clinicians are supposed to use guidelines widely disseminated as the 'Basic Paediatric Protocols' in Kenya (10) that are adapted from WHO guidance, based on available evidence and developed by consensus by a national guideline panel (see (21-23)). In standard practice, the process of treatment assignment happens in three steps; first, there is assessment and documentation of each clinical sign. Step two involves integration of clinical information into severity classification, and in step three severity classification is translated into a treatment assignment (see panel 1). In Kenya, as in many low and middle income countries these

recommendations reflect the absence of access to further diagnostic tests. Thus pulse oximetry, blood culture or tests for inflammatory markers are not routinely available (7). As indicated above clinicians may fail to adhere to guideline recommendations by making errors or over-riding recommendations at any of the three steps of assessment, severity classification and treatment assignment. However, based on the clinical symptoms and signs recorded it is possible to assign a severity classification (and thus expected treatment) based on the data. It is a data informed and investigator assigned classification as indrawing pneumonia that is used in the primary analysis (experiment 1).

c) Analysis Variables

Outcome variable

Mortality will be used as the outcome variable in all the three experiments.

Independent variables

These variables are grouped into key and auxiliary. Key variables are defined as those that should influence pneumonia severity classification and hence treatment based on the treatment protocol (10) (panel 1). Auxiliary variables are defined as those that might, *a priori,* be expected to influence treatment assignment based on clinical reasoning (for example they might make a clinician concerned for severe illness), although according to the formal rules (the guidelines) they are not considered reasons to alter treatment assignment. Such auxiliary variables were identified from those clinical symptoms and signs that are routinely collected within CIN. See table 1 for a summary of key and auxiliary variables that will be used in the analyses.

Table 1: Summary of key and auxiliary independent variables for experiments 1, 2 and 3[3]

Experiment 1 and 2 key variables	Experiment 3 key variables	Auxiliary variables for experiments 1, 2 and 3
Age (2 – 59 months)	Age (2 – 59 months)	Gender (male/female)
Indrawing (present/absent)	Indrawing (present/absent)	Cough duration (days)
History of cough (yes/no)	History of cough (yes/no)	Crackles (present/absent)
Difficulty breathing (present/absent)	Difficulty breathing (present/absent)	Weight (Kg)
Level of consciousness – AVPU (alert/verbal response/pain response/unresponsive)	Level of consciousness – AVPU (alert/verbal response/pain response/unresponsive)	Pallor (0, +, +++)
	Oxygen ordered (yes/no)	Capillary refill (immediate, 1 – 2 secs, 3 – 6 sec, > 6 secs)
	Cyanosis (present/absent)	Fever (present/absent)
	Inability to drink/breastfeed (yes/no)	Diarrhoea (present/absent)
	Grunting (present/absent)	Convulsions (present/absent)
	Respiratory rate (breaths/min)	Vomiting (yes/no)
		Referral (yes/no)
		Length of illness (days)
		Number of fits
		Thrush (present/absent)
		Quinine/artesunate (prescribed/not prescribed)
		Weight for age z – score
		Wheeze (present/absent)
		Comorbidities (Malaria and or diarrhoea)

d) Sample size verification

Here, sample size verification uses the formula cited in (24):

$$ns = \frac{k+1}{k}\frac{\bar{p}(1-\bar{p})(Z_{\beta}+Z_{1-\alpha/2})^2}{(p_1-p_2)^2}$$, where:

[3] Experiment 3 has more key variables than experiment 2 as it considers patient populations with "very severe, severe and non –severe pneumonia" – as classified in the previous WHO and Kenyan treatment guidelines. Therefore, in addition to variables used to classify severe pneumonia, other variables used to classify very severe and non-severe pneumonia are considered.

ns = size of smaller group.

k = ratio of larger group to smaller group.

$p_1 - p_2$ = clinical difference in proportions of the outcome.

Z_β = corresponds to power of 80%

$Z_{1-\alpha/2}$ = corresponds to two-tailed significance level (1.96 for $\alpha = .05$).

$\bar{p}$ = corresponds to average of outcome proportions in two groups.

The value for $\bar{p}$ is estimated from studies – two of which formed evidence for earlier WHO indrawing pneumonia treatment guidelines. See table 2 that shows the number of deaths per treatment arm reported in these studies.

Table 2: Summary of some of pneumonia studies that informed previous WHO guidelines

Study	Treatment arms	Mortality	$\bar{p}$
Shann et. al(1985)	Chloramphenicol alone	48/377	0.1470
	Chloraphenicol+Penicillin	62/371	
Addo – Yobo et. al (2002)	Injectable penicillin	7/845	0.0050
	Oral amoxicillin	2/857	
Agweyu (2015)	Injectable penicillin	3/264	0.008
	Oral amoxicillin	1/263	

For assessment of sample size for indrawing pneumonia experiments, a weighted[4] $\bar{p}$ of 0.041 from these studies is used. The ratio r is varied between 1 and 3. Figure 1 was generated by fixing power and significance level at 80% and 5% respectively. Estimates of $\bar{p}(1-\bar{p})$ derived from WHO studies were substituted in the sample size formula and data simulated in order to see what detectable differences would be achieved by different sample sizes. A total sample size of about 4000 would be sufficient to detect a minimum difference of 1.5% (absolute difference e.g. a reduction of mortality from X% to X – 1.5%) in any of these experiments[5].

[Insert **figure 1**]

[4] Weighting was done using the total sample sizes per experiment.

[5] A sample size of at least 4000 would be required for experiment 3 as this is the minimum sample for experiments 1 and 2 which are nested in experiment 3.

Statistical and outcome Analysis

Statistical analysis will proceed in the following four steps:

Step 1 – subset of patients of interest for the experiments will be obtained.

- Experiment 1: First, missing clinical signs data will be multiply imputed[6] (excluding outcome data) – and then key clinical signs data used to impute (construct) a pneumonia severity level for all patients based on the algorithms in the pneumonia treatment protocol (10). Thereafter, a subset of patients with guideline-defined indrawing pneumonia (for each of the imputed datasets) will be obtained for further analyses.
- Experiment 2: A subset of indrawing pneumonia patients (with severity as indicated by the clinicians) will be obtained from the raw dataset – and clinical signs data imputed using multiple imputation (without the outcome data).
- Experiment 3: The raw dataset containing all the patients with all forms of pneumonia severity will be used –and clinical signs data imputed using multiple imputation (without the outcome data).

Step 2 – patients in the alternative treatment arms will be matched using propensity score (PS) methods to overcome non – random treatment allocation. Standardised mean differences (and where necessary density plots) will be used as diagnostic checks for covariate balance and overlap (26, 27) between penicillin and penicillin plus gentamicin treatment groups. PS methods that utilise all the data (PS optimal full matching, weighting and sub classification) will be examined in experiments 1 and 2 (on each imputed dataset) and the method that results in the minimum average absolute standardised mean differences for the majority of the variables and retains the largest number of patients in the analysis will be considered appropriate (28). While only PS sub-classification will be used for experiment 3. As experiment 3 aims to investigate comparative effectiveness in all cases of pneumonia,

[6] For the three experiments, 20 datasets will be multiply imputed using chained equations (25).

propensity score will be used as a proxy for disease severity thus patients with lower propensity scores will be considered less ill while those with higher propensity scores will be considered more ill (grouped in propensity score subclasses for analysis).

Step 3: conducting outcome analysis.

For each imputed dataset (per experiment), outcome analysis will aim to investigate treatment causal effects across all the hospitals. Bayesian log binomial regression models (29) will be used to estimate overall treatment effects[7]. A hospital variable will be modelled as a fixed effect in the log binomial regression that measures treatment effects on pooled data. These models will be fitted on each imputed dataset (adjusting for other variables used in PS models) and results pooled using Rubin rules (30).

Step 4: sensitivity analysis will be conducted to investigate effects of unmeasured confounders and validity of estimates obtained through multiple imputation. Propensity score methods generate matched treated and (active) control patients whose distribution of measured covariates are as similar as possible. However, two patients with similar covariate distribution may differ in terms of unmeasured variables – and this may introduce bias in estimated treatment effects (31). On the other hand, if outcome and explanatory variables have missing data, then inclusion of outcome data in multiple imputation may contribute minor information in the substantive (outcome) model (32).

Exploring effects of unmeasured confounders

[7] Bayesian models will be used to overcome any bias due to sparsity of data as PS sub-classification in itself reduces the effective sample size.

Sensitivity analysis for unmeasured confounders will involve the use of an instrumental variable (IV) (33) – weekend admission and PS trimming (34). A few IV sources in health studies have been described in Baiocchi (2014) (35). These include: distance to specialty, genes, insurance plan, timing of admission, calendar time and preference based IVs. Of relevance to this analysis would be timing of admission IVs. A study conducted by Berkley (2004) (36) in a Kenyan hospital demonstrated that children who were admitted during the weekend experienced higher mortality compared to those admitted during the weekdays – which is an indication of poor quality of care and treatment during the weekend. In other words, it is anticipated that children admitted during the weekdays would have better health outcomes. This, in theory, implies that the type of treatment and care received depend on the day of admission – and which later determines the type of health outcome of the patient.

Examining validity of multiple imputation

The analysis steps 1 – 3 above will exclude outcome data in the imputation model – however sensitivity analysis will include models in which the outcome variable is included in the imputation approach. This will aim to investigate if including outcome data in the imputation model has an influence[8].

Ethics and dissemination

This analysis will be based on a larger project (CIN) which was cleared by the Kenya Medical Research Institute ethics and review board (Protocol number: 2465). The findings will be useful in understanding the external validity of current treatments – and will provide a platform on which to do more similar analyses for different (combinations of) treatments.

Competing Interests

[8] The primary interpretations will consider results of multiple imputations without outcome if results differ from those of MI with outcome – as is the standard recommendation to analysis of observational datasets in Rubin (2008) (37).

The authors declare they have no competing interests.

Authors Contributions

The contributions of the authors were as follows: LM did an initial draft of this manuscript with the support of RP, EM and ME. Thereafter, all authors edited subsequent versions and approved the final copy.

Funding

We are grateful for the funds from the Wellcome Trust (#097170) that support ME through a fellowship and additional funds from a Wellcome Trust core grant awarded to the KEMRI-Wellcome Trust Research Programme (#092654) that supported this work. LM is supported by a Nuffield Department of Medicine Prize DPhil Studentship and Clarendon Scholarship (Oxford University). The funders had no role in drafting or submitting this manuscript.

References

1. Hospital Care for Children [Internet].; 2005 []. Available from: http://apps.who.int/iris/bitstream/10665/43206/1/9241546700.pdf.

2. Hospital Care for Children [Internet].; 2013 []. Available from: http://apps.who.int/iris/bitstream/10665/81170/1/9789241548373_eng.pdf.

3. WHO. Hospital Care for Children. 2013.

4. Sazawal S, Black R. Pneumonia Case Management Trials Group. Lancet Infect Dis. 2003;3(9):547-56.

5. Agweyu A, Kibore M, Digolo L, Kosgei C, Maina V, Mugane S, et al. Prevalence and correlates of treatment failure among Kenyan children hospitalised with severe community-acquired pneumonia: a prospective study of the clinical effectiveness of WHO pneumonia case management guidelines. Tropical Medicine & International Health. 2014.

6. Mulholland K, Carlin JB, Duke T, Weber M. Challenges of trials of antibiotics for pneumonia in low-income countries. The Lancet Respiratory Medicine. 2014;2(12):952-4.

7. Ayieko P, Ogero M, Makone B, Julius T, Mbevi G, Nyachiro W, et al. Characteristics of admissions and variations in the use of basic investigations, treatments and outcomes in

Kenyan hospitals within a new Clinical Information Network. Archives of Disease in Childhood. 2015.

8. Agweyu A, Opiyo N, English M. Experience developing national evidence-based clinical guidelines for childhood pneumonia in a low-income setting - making the GRADE? BMC Pediatric. 2012.

9. Ayieko P, English M. Case Management of Childhood Pneumonia in Developing Countries . Pediatr Infect Dis J. 2007;26:432-40.

10. Basic Paediatric Protocols: November 2013 Edition [Internet].: Ministry of Health; 2013 []. Available from: http://www.idoc-africa.org/index.php/86-clinical-guide/ken-guide/134-basic-paediatric-protocols-november-2013-edition.

11. Basic paediatric protocols [Internet].: Kenyan Ministry of Health; 2016 []. Available from: https://www.tropicalmedicine.ox.ac.uk/_asset/file/basic-paediatric-protocols-2016.pdf.

12. Graham SM, English M, Hazir T, Enarson P, Duke T. Challenges to improving case management of childhood pneumonia at health facilities in resource limited settings. Bulletin of the WHO. 2015.

13. Atkinson M, et al. A multicentre randomised controlled equivalence trial comparing oral amoxicillin and intravenous benzyl penicillin for community acquired pneumonia in children PIVOT Trial. Thorax. 2007;61:1102-6.

14. Hazir T, et al. Ambulatory short-course high-dose oral amoxicillin for treatment of severe pneumonia in children: a randomised equivalency trial. Lancet. 2008;371:49-56.

15. Addo-Yobo E, Chisaka N, Hassan M, Hibberd P, Lozano JM, Jeena P, et al. Oral amoxicillin versus injectable penicillin for severe pneumonia in children aged 3 to 59 months: a randomised multicentre equivalency study. Lancet. 2004;364:1141-8.

16. Addo-Yobo E, et al. Outpatient treatment of children with severe pneumonia with oral amoxicillin in four countries: the MASS study . Tropical Medicine and International Health. 2011;16:995-1006.

17. IVAC. Pneumonia and Diarrhoea Progress Report 2014. John Hopkins Bloomberg School of Public Health; 2014.

18. Agweyu A, Gathara D, Oliwa J, Muinga N, Edwards T, Allen E, et al. Oral amoxicillin versus benzyl penicillin for severe pneumonia among kenyan children: a pragmatic randomised controlled noninferiority trial. Clin Infect Dis. 2015.

19. Stopler E, Van Roeyn P, Van de Wiel M, Van Bokhoven M, Houben P, Van der Weijde T, et al. Consensus on gut feelings in general practice. BMC Family Practice. 2009.

20. Rubin DB. For Objective Causal Inference, Design Trumps Analysis. 2008;2(3):808-40.

21. Draft recommendation for management of children with severe febrile illness and impared circulation without signs of severly impaired circulation: Guideline Panel Meeting

[Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idoc-africa.org/images/documents/guidelines/Panel%20recommendation%20on%20fluid%20resuscitation%20I%20final.pdf.

22. Draft recommendations for care of umbilical cord in newborns: Guideline Panel Meeting. [Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idoc-africa.org/images/documents/guidelines/Panel%20recommendations%20on%20cord%20care%20final.pdf.

23. Draft recommendations for hospital management of children (less than 5 years) with sickle cell disease: Guideline Panel Meeting 2013. [Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idoc-africa.org/images/documents/guidelines/Panel%20recommendations%20on%20sickle%20cell%20anaemia%20final.pdf.

24. Wittes J. Sample size calculations for randomised controlled trials. Epidemiologic Reviews. 2002;24(1).

25. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? International Journal of Methods in Psychiatric Research. 2011;20(1):40-9.

26. Austin PC. Assessing balance in measured baseline covariates when using many-to-one matching on the propensity-score. Pharmacoepidemiol Drug Saf. 2008 2008;17(12):1218-25.

27. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009 10 Nov 2009;28(25):3083-107.

28. Stuart EA. Matching methods for causal inference:A review and a look forward. Statist. Sci. 2010;25(1):1-21.

29. Fine J, Gray R. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.

30. Rubin DB. An overview of Multiple Imputation. . 1988.

31. Rosenbaum PR. Design sensitivity and efficiency in observational studies. Journal of the American Statistical Association. 2010;105:692-702.

32. Roderick JAL. Regression With Missing X's: A Review. Journal of the American Statistical Association. 1992;87(420):1227-37.

33. Chiba Y. Bias analysis of the instrumental variable estimator as an estimator of the average causal effect. Contemporary Clinical Trials. 2010 January 2010;31(1):12-7.

406 34. Lee BK, Lessler J, Stuart EA. Weight Trimming and Propensity Score Weighting. PLoS
407 ONE [Electronic Resource]. 2011.

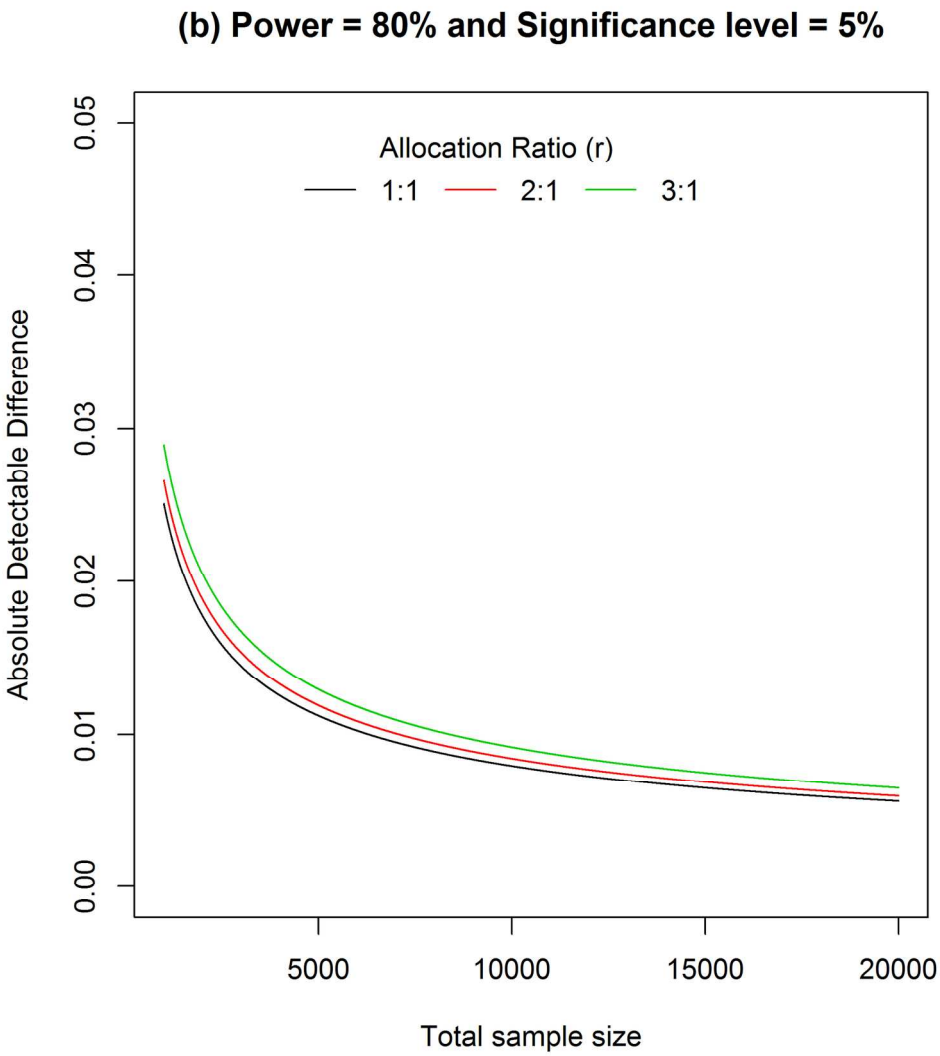
408 35. Baiocchi M, Cheng J, Small DS. Tutorial in Biostatistics: Instrumental Variable Methods
409 for Causal Inference. Statistics in Medicine. 2014;33(13):2297-340.

410 36. Berkley JA, Brent A, Mwangi I, English M, Maitland K, Marsh K, et al. Mortality among
411 Kenyan children admitted to a rural district hospital on weekends as compared with
412 weekdays . Pediatrics. 2004;114:1737-8.

413 37. Rubin D.B. For Objective Causal Inference, Design Trumps Analysis. The Annals of
414 Applied Statistics. 2008;2(3):808-40.

415 38. Tuti T, Bitok M, Malla L, Paton C, Muinga N, Gathara D, et al. Improving documentation
416 of clinical care within a clinical information network: an essential initial step in efforts to
417 understand and improve care in Kenyan hospitals. BMJ Global Health. 2016.

418



169x192mm (300 x 300 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/ item	ItemNo	Description	Line Number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1 – 3
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	Not Applicable
	2b	All items from the World Health Organization Trial Registration Data Set	
Protocol version	3	Date and version identifier	Not Applicable
Funding	4	Sources and types of financial, material, and other support	304 – 308
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	300 – 303
	5b	Name and contact information for the trial sponsor	Not Applicable
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	308
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	Not Applicable
Introduction			

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	57 – 103
	6b	Explanation for choice of comparators	57 – 103
Objectives	7	Specific objectives or hypotheses	104 – 152
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	158 – 170
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	155
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	172 – 177
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	56 – 103
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	Not applicable
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	Not applicable
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	Not applicable

Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	197
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Not applicable
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	214 – 230
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	Not applicable
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	Not applicable
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	Not applicable

Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	Not applicable
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	Not applicable
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	Not applicable
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	155
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	Not applicable
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	155
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	231 – 291

	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	231 – 291
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	Not applicable
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	Not applicable
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	Not applicable
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	Not applicable
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	292 – 298
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	Not applicable

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	Not applicable
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	Not applicable
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	Not applicable
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	298
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	Not applicable
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	Not applicable
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	295 – 296
	31b	Authorship eligibility guidelines and any intended use of professional writers	
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	

Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated.

BMJ Open

Comparative effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in Kenya: A protocol for an observational study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016784.R1
Article Type:	Protocol
Date Submitted by the Author:	15-May-2017
Complete List of Authors:	Malla, Lucas; Kenya Medical Research Institute - Wellcome Trust; University of Oxford Perera, Rafael; University of Oxford, Primary Health Care McFadden, Emily; University of Oxford, Primary Health Care English, Mike; University of Oxford, Nuffield Department of Medicine and Department of Paediatrics
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Health services research
Keywords:	comparative effectiveness, propensity scores, pneumonia, observational

Comparative effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in Kenya: A protocol for an observational study

Lucas Malla[1*] (lucas.malla@some.ox.ac.uk)

Rafael Perera-Salazar[2] (rafael.perera@phc.ox.ac.uk)

Emily McFadden[2] (emily.mcfadden@phc.ox.ac.uk)

Mike English[1, 3] (menglish@kemri-wellcome.org)

[1]Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom.

[3] Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya.

***Corresponding author:**

Lucas Malla

Abstract

Introduction

WHO treatment guidelines are widely recommended for guiding treatment for millions of children with pneumonia every year across multiple low and middle income countries. Guidelines are based on synthesis of available evidence that provides moderate certainty in evidence of effects for forms of pneumonia that can result in hospitalisation. However, trials have included fewer children from Africa than other settings and it is suggested that African children with pneumonia have higher mortality. Thus despite improving access to recommended treatments and deployment with high coverage of childhood vaccines, pneumonia remains one of the top causes of mortality for children in Kenya. Establishing whether there are benefits of alternative treatment regimens to help reduce mortality would utilize pragmatic clinical trials. However, these remain relatively expensive and time consuming. This protocol describes an approach to using secondary analysis of a new, large observational dataset as a potentially cheaper and quicker way to examine the comparative effectiveness of penicillin versus penicillin plus gentamicin in treatment of indrawing pneumonia. Addressing this question is important as although it is now recommended that this form of pneumonia is treated with oral medication as an outpatient it remains associated with non-trivial mortality that may be higher outside trial populations.

Methods and analysis

We will use a large observational dataset that captures data on all admissions to 13 Kenyan county hospitals. These data represent the findings of clinicians in practice and, because the system was developed for large observational research, pose challenges of non-random treatment allocation and missing data. To overcome these challenges this analysis will use a

rigorous approach to study design, propensity score methods and multiple imputation to minimize bias.

Ethics and dissemination

The primary data are held by hospitals participating in the Kenyan Clinical Information Network (CIN) project with de-identifed data shared with the KEMRI-Wellcome Trust Research Programme for agreed analyses. The use of data for the analysis described received ethical clearance from the Kenya Medical Research Institute Scientific and Ethical Review Committee. The findings of this analysis will be published.

Strength

- This study will be used as a platform to explore effectiveness of alternative treatments in routine care in a low income setting to improve health outcomes for children.

Limitation

- The analysis will be limited to the variables in the observational dataset – and therefore risk bias due to unmeasured key variables.
- The influence of any resulting bias, to alter results, will however be assessed through the use of alternative methods as instrumental variables.

Introduction

Kenya has developed and disseminated national treatment guidelines largely drawing on those of WHO for a number of childhood diseases including pneumonia (1, 2). These pneumonia guideline recommendations are based on synthesis of available evidence that provides moderate certainty in evidence of effects of treatments for forms of pneumonia that can result in hospitalization (1, 2). Such guidelines have been shown to be effective in reducing pneumonia related mortality and thus Kenyan clinicians are supposed to use them in routine practice to treat pneumonia (and other diseases) (3, 4). However, although the guidelines are based on the best available evidence, the evidence available from trials conducted in Africa remains limited (5). There has also been little thorough investigation of the effectiveness of treatments in non-trial populations in routine settings that may often differ from those enrolled in formal clinical trials. For example many children admitted with pneumonia may have co-morbidity that might exclude them from trials (6). These issues can prove problematic when making national guidelines where study generalisability can be contested (7).

The WHO and Kenyan pneumonia treatment guidelines are implicitly based on risk stratification of illness with children deemed at higher risk of severe illness and mortality offered broad spectrum antibiotic regimens and those at lower risk narrow spectrum antibiotics (2, 8-10). This risk stratification approach is operationalized by requiring clinicians to look for specific features in the clinical history and examination that are used to define illness severity and therefore recommended treatment (Panel 1). Previous studies conducted in Kenya have, however, indicated that clinicians do not always follow guideline recommendations in treating pneumonia (4). Variation from the guideline recommended approach can occur at the point of pneumonia severity assignment (clinicians do not follow the rules linking clinical signs and severity category) and at the point of treatment assignment

(clinicians do not follow the rules linking treatment and severity). This variability in treatment assignment provides the opportunity for comparative effectiveness evaluation if similar populations of children with pneumonia are prescribed different treatments. Clinicians may create such a situation by not following recommendations because they have inadequate

knowledge or if they believe (potentially contrary to the evidence) that certain treatments result in better health outcomes.

Panel 1: Pneumonia treatment algorithm

The pneumonia severity classification that was recommended by Kenyan guidelines up to March 2016 (9) (and previously by WHO guidelines (1)) defined the following three severity classes:

1. ***Very severe pneumonia:*** *If a child had either oxygen saturation less than 90% or central cyanosis or was grunting or unable to drink or not alert, then s/he was classified as having very severe pneumonia, put on oxygen and treated with a combination of gentamicin and penicillin.*

   {The new WHO (2) and Kenyan guidelines (9) renamed this class as "severe pneumonia" – and currently recommend treatment with a combination of ampicillin (or penicillin) with gentamicin plus oxygen}.
2. ***Severe pneumonia:*** *If a child had lower chest wall indrawing (but did not have any of qualifying signs for very severe pneumonia above) and was alert then s/he was to be classified as having severe pneumonia and be treated with benzyl penicillin only.*

   Note: *The term indrawing pneumonia is hereafter used in this protocol to define this category of children to avoid confusion.*
3. ***(Non – severe) Pneumonia:*** *If a child had none of the mentioned signs but had cough or difficulty breathing and a respiratory rate greater than or equal to 50 breaths/minute (for age between 2 and 11 months) or respiratory rate greater than or equal to 40 breaths/minute (for age above 12 months) then s/he was classified as having non severe pneumonia and treated with cotrimoxazole or amoxicillin if previously treated with cotrimoxazole.*

   {The current WHO and Kenyan guidelines collapsed severity classes 2 and 3 into one category referred to as "non –severe pneumonia". This group of patients are currently treated with oral amoxicillin – partly informed by a local trial (18)}.

In particular, a previous study showed that clinicians over-prescribed gentamicin, adding this to penicillin for the treatment of pneumonia characterized by lower chest wall indrawing but no other signs of severe illness instead of penicillin alone as was recommended (4)[1]. Therefore, this protocol is for a study that seeks to explore whether there is any benefit from adding gentamicin to penicillin in treating children with indrawing pneumonia. Such a benefit could accrue if bacterial causes of pneumonia that were previously (prior to introduction of new vaccines) proportionately less common (eg. *S. aureus* and gram negative bacteria) are now accounting for an increased proportion of pneumonia deaths – as in such cases, the addition of gentamicin might provide effective treatment for a broader spectrum of pathogens. Tackling this question is of importance as WHO have recently changed indrawing pneumonia treatment guidance based on trials that suggest equivalence of oral amoxicillin and injectable penicillin (12-15). New guidance recommends outpatient oral treatment for a population of children previously admitted to hospital (10). However, mortality from pneumonia has been reported to be higher in African settings (16, 17) despite the increasing use of multiple vaccines spanning: measles, pertussis, HiB and pneumococcal conjugate vaccines. It remains possible therefore that for a small number of children a broader spectrum antibiotic regimen might be of benefit. This study addresses this question that has not been the subject of prior community and pragmatic clinical trials.

Objectives

Primary

1) Experiment 1: To compare the effectiveness of injectable penicillin versus penicillin plus gentamicin (both injectable) in treatment of indrawing pneumonia; where severity level is constructed (imputed) using data recorded on each child's clinical

[1] The fact that inadequate knowledge in handling childhood pneumonia may result in inconsistent treatment allocation is supported by a survey conducted in seven developing countries showing that 56% of nurses and doctors had inadequate knowledge in managing pneumonia in children (11).

signs (hospitals use a structured record form that supports recording of signs highlighted in guidelines) such that severity classification is consistent with guideline recommendations.

Secondary

2) Experiment 2: To compare effectiveness of injectable penicillin versus penicillin plus gentamicin in treatment of indrawing pneumonia; where we use clinician assigned severity level.
3) Experiment 3: To compare effectiveness of injectable penicillin versus penicillin plus gentamicin in treatment of all cases of pneumonia admitted to hospital.

Experiment 1 will be primary as it most approximates a typical randomised trial where recruitment would be based on specified clinical signs. This scenario will provide an evaluation of alternative therapies within a guideline class (where children have very similar clinical signs) and thus is the best mimic of a prospectively designed comparative evaluation in which clinicians stick to the rules of severity classification (see (18) for an example of a RCT in Kenya that this would be similar to – where classification is based on clinical signs). Recommended treatment for this disease classification was penicillin alone, treatment with combination therapy may therefore represent over-treatment. Alternatively, the combination treatment that provides broader antimicrobial cover could provide an advantage in a small proportion of cases that would only be detected in moderately large studies – where the addition of gentamicin offers improved treatment for specific organisms not susceptible to penicillin alone.

Experiment 2 will provide a test of alternative therapies amongst those where clinicians used their own judgement (possibly including gut feeling) to classify and treat (19) and have on occasions (potentially) over-ridden or ignored the guideline recommendations. In this case

although the same label of indrawing pneumonia is given to all, the treatment selected may be an indicator of perceived severity and there may be a potential bias as a result – and the propensity score distributions (see below) may help demonstrate this and in theory may overcome this potential bias. Here if there is no clinically relevant difference between treatments within a group of patients that reflects clinicians' actual classification decisions this could reassure them that monotherapy with penicillin (or amoxicillin) would be acceptable. Lastly, experiment 3 is an extension of the logic of experiment two. To date there have been no pragmatic trials of penicillin alone compared with alternative combination therapies for all forms of inpatient pneumonia, and addressing this question may be relevant for two reasons. First, the population of children admitted with severe forms of pneumonia is now largely one that has received *H. influenzae* Type B and pneumococcal conjugate vaccines that have likely changed the aetiology of this illness. Second, if clinicians are poorly trained and unable to classify illness severity – resulting in non-adherence to guidelines - it would be useful to explore the potential impact of this across all levels of severity of pneumonia. This analysis has the largest numbers of subjects.

Methods and analysis

To answer these three questions, we will use the Kenyan Clinical Information Network (CIN) dataset that provides observational data on all admissions to 13 Kenyan County hospitals (panel 2).

Panel 2: Clinical Information Network

CIN was initiated to improve data availability from secondary care in paediatrics and as a model for demonstrating the value of routine data in improving quality of care in the county (formerly district) hospitals. These hospitals typically have a single paediatrician leading services predominantly provided by junior clinical teams. Data in these hospitals are collected prospectively post discharge by trained data clerks, guided by well-defined standard operating procedures, under close supervision by the hospital medical records department and the research team. It is worth noting that the research team has no personnel checking quality of clinical process and whether clinicians correctly document what they do. However, the patient record is the formal (and legal) document describing the clinical condition and management. These documents are used for data abstraction and they include patient files with standardized Paediatric Admission Record (PAR) forms, treatment sheets, discharge summary forms, laboratory reports and clinician notes. The collected data are used to assess documentation of history, physical examination, diagnosis, laboratory investigations, treatment and discharge plans. Feedback to hospitals as part of the CIN activities has helped improve the quality of clinical data (38). The description of hospital selection and their populations of patients is detailed in Ayieko (2015) (6).

The analysis will proceed in two stages – design and outcome analysis as suggested by Rubin (2008) (20) as an objective way for analysing observational datasets.

Study Design

This will be an observational study conducting secondary analyses of data routinely collected from hospital paediatric wards in Kenya's CIN. The design process for the three experimental scenarios will be similar and broadly consists of the following steps suggested in Rubin (2008) (20):

a) Definition of inclusion and exclusion criteria.

b) Understanding the pneumonia diagnosis and treatment assignment processes. This is to help understand key and auxiliary variables required for analysis.

c) Verification of sample size if sufficient for any meaningful analyses.

d) Creation of comparable treatment arms – which will be addressed analytically aiming to overcome non – random treatment assignment and deal with missing data.

e) Outcome analysis follows after conceptualisation of design in steps a – d.

a) Inclusion and exclusion

This analysis will include all children aged 2 – 59 months and will exclude children with any co-morbidity of HIV, meningitis, tuberculosis and or acute malnutrition as there are specific antibiotic treatment rules for these children that supersede those for pneumonia. Specifically Kenyan guidelines for the inpatient treatment of pneumonia in children that are HIV infected recommend only combination therapy. Importantly therefore children with other co-morbidities such as mild anaemia, diarrhoea and malaria are not necessarily excluded from the analysis.

b) Understanding the diagnosis and treatment assignment rules for pneumonia paediatric patients

Clinicians are supposed to use guidelines widely disseminated as the 'Basic Paediatric Protocols' in Kenya (9) that are adapted from WHO guidance, based on available evidence and developed by consensus by a national guideline panel (see (21-23)). In standard practice, the process of treatment assignment happens in three steps; first, there is assessment and documentation of each clinical sign. Step two involves integration of clinical information into severity classification, and in step three severity classification is translated into a treatment assignment (see panel 1). In Kenya, as in many low and middle income countries these recommendations reflect the absence of access to further diagnostic tests. Thus pulse oximetry, blood culture or tests for inflammatory markers are not routinely available (6). As indicated above clinicians may fail to adhere to guideline recommendations by making errors

or over-riding recommendations at any of the three steps of assessment, severity classification and treatment assignment. However, based on the clinical symptoms and signs recorded it is possible to assign a severity classification (and thus expected treatment) based on the data. It is a data informed and investigator assigned classification as indrawing pneumonia that is used in the primary analysis (experiment 1).

c) Analysis Variables

Outcome variable

Mortality will be used as the outcome variable in all the three experiments.

Independent variables

These variables are grouped into key and auxiliary. Key variables are defined as those that should influence pneumonia severity classification and hence treatment based on the treatment protocol (9) (panel 1). Auxiliary variables are defined as those that might, *a priori,* be expected to influence treatment assignment based on clinical reasoning (for example they might make a clinician concerned for severe illness), although according to the formal rules (the guidelines) they are not considered reasons to alter treatment assignment. Such auxiliary variables were identified from those clinical symptoms and signs that are routinely collected within CIN. See table 1 for a summary of key and auxiliary variables that will be used in the analyses.

Table 1: Summary of key and auxiliary independent variables for experiments 1, 2 and 3[2]

Experiment 1 and 2 key variables	Experiment 3 key variables	Auxiliary variables for experiments 1, 2 and 3
Age (2 – 59 months)	Age (2 – 59 months)	Gender (male/female)
Indrawing (present/absent)	Indrawing (present/absent)	Cough duration (days)
History of cough (yes/no)	History of cough (yes/no)	Crackles (present/absent)
Difficulty breathing (present/absent)	Difficulty breathing (present/absent)	Weight (Kg)
Level of consciousness – AVPU (alert/verbal response/pain response/unresponsive)	Level of consciousness – AVPU (alert/verbal response/pain response/unresponsive)	Pallor (0, +, +++)
	Oxygen ordered (yes/no)	Capillary refill (immediate, 1 – 2 secs, 3 – 6 sec, > 6 secs)
	Cyanosis (present/absent)	Fever (present/absent)
	Inability to drink/breastfeed (yes/no)	Diarrhoea (present/absent)
	Grunting (present/absent)	Convulsions (present/absent)
	Respiratory rate (breaths/min)	Vomiting (yes/no)
		Referral (yes/no)
		Length of illness (days)
		Number of fits
		Thrush (present/absent)
		Quinine/artesunate (prescribed/not prescribed)
		Weight for age z – score
		Wheeze (present/absent)
		Comorbidities (Malaria and or diarrhoea)

d) Sample size verification

Here, sample size verification uses the formula cited in (24):

$$ns = \frac{k+1}{k}\frac{\bar{p}(1-\bar{p})(Z_{\beta}+Z_{1-\alpha/2})^2}{(p_1-p_2)^2}$$, where:

ns = size of smaller group.

k = ratio of larger group to smaller group.

$p_1 - p_2$ = clinical difference in proportions of the outcome.

Z_{β} = corresponds to power of 80%

$Z_{1-\alpha/2}$ = corresponds to two-tailed significance level (1.96 for $\alpha = .05$).

$\bar{p}$ = corresponds to average of outcome proportions in two groups.

[2] Experiment 3 has more key variables than experiment 2 as it considers patient populations with "very severe, severe and non –severe pneumonia" – as classified in the previous WHO and Kenyan treatment guidelines. Therefore, in addition to variables used to classify severe pneumonia, other variables used to classify very severe and non-severe pneumonia are considered.

The value for $\bar{p}$ is estimated from studies – two of which formed evidence for earlier WHO indrawing pneumonia treatment guidelines. See table 2 that shows the number of deaths per treatment arm reported in these studies.

Table 2: Summary of some of pneumonia studies that informed previous WHO guidelines

Study	Treatment arms	Mortality	$\bar{p}$
Shann et. al(1985)	Chloramphenicol alone	48/377	0.1470
	Chloraphenicol+Penicillin	62/371	
Addo – Yobo et. al (2002)	Injectable penicillin	7/845	0.0050
	Oral amoxicillin	2/857	
Agweyu (2015)	Injectable penicillin	3/264	0.008
	Oral amoxicillin	1/263	

For assessment of sample size for indrawing pneumonia experiments, a weighted[3] $\bar{p}$ of 0.041 from these studies is used. The ratio r is varied between 1 and 3. Figure 1 was generated by fixing power and significance level at 80% and 5% respectively. Estimates of $\bar{p}(1-\bar{p})$ derived from WHO studies were substituted in the sample size formula and data simulated in order to see what detectable differences would be achieved by different sample sizes. A total sample size of about 4000 would be sufficient to detect a minimum difference of 1.5% (absolute difference e.g. a reduction of mortality from X% to X – 1.5%) in any of these experiments[4].

[Insert **figure 1**]

Statistical and outcome Analysis

Statistical analysis will proceed in the following four steps:

Step 1 – subset of patients of interest for the experiments will be obtained.

[3] Weighting was done using the total sample sizes per experiment.

[4] A sample size of at least 4000 would be required for experiment 3 as this is the minimum sample for experiments 1 and 2 which are nested in experiment 3.

- Experiment 1: First, missing clinical signs data will be multiply imputed[5] (excluding outcome data) – and then key clinical signs data used to impute (construct) a pneumonia severity level for all patients based on the algorithms in the pneumonia treatment protocol (9). Thereafter, a subset of patients with guideline-defined indrawing pneumonia (for each of the imputed datasets) will be obtained for further analyses.
- Experiment 2: A subset of indrawing pneumonia patients (with severity as indicated by the clinicians) will be obtained from the raw dataset – and clinical signs data imputed using multiple imputation (without the outcome data).
- Experiment 3: The raw dataset containing all the patients with all forms of pneumonia severity will be used –and clinical signs data imputed using multiple imputation (without the outcome data).

Step 2 – patients in the alternative treatment arms will be matched using propensity score (PS) methods to overcome non – random treatment allocation. Standardised mean differences (and where necessary density plots) will be used as diagnostic checks for covariate balance and overlap (26, 27) between penicillin and penicillin plus gentamicin treatment groups. PS methods that utilise all the data (PS optimal full matching, weighting and sub classification) will be examined in experiments 1 and 2 (on each imputed dataset) and the method that results in the minimum average absolute standardised mean differences for the majority of the variables and retains the largest number of patients in the analysis will be considered appropriate (28). While only PS sub-classification will be used for experiment 3. As experiment 3 aims to investigate comparative effectiveness in all cases of pneumonia, propensity score will be used as a proxy for disease severity thus patients with lower

[5] For the three experiments, 20 datasets will be multiply imputed using chained equations (25).

propensity scores will be considered less ill while those with higher propensity scores will be considered more ill (grouped in propensity score subclasses for analysis).

Step 3: conducting outcome analysis.

For each imputed dataset (per experiment), outcome analysis will aim to investigate treatment causal effects across all the hospitals. Bayesian log binomial regression models (29) will be used to estimate overall treatment effects[6]. A hospital variable will be modelled as a fixed effect in the log binomial regression that measures treatment effects on pooled data. These models will be fitted on each imputed dataset (adjusting for other variables used in PS models) and results pooled using Rubin rules (30).

Step 4: sensitivity analysis will be conducted to investigate effects of unmeasured confounders and validity of estimates obtained through multiple imputation. Propensity score methods generate matched treated and (active) control patients whose distribution of measured covariates are as similar as possible. However, two patients with similar covariate distribution may differ in terms of unmeasured variables – and this may introduce bias in estimated treatment effects (31). On the other hand, if outcome and explanatory variables have missing data, then inclusion of outcome data in multiple imputation may contribute minor information in the substantive (outcome) model (32).

Exploring effects of unmeasured confounders

Sensitivity analysis for unmeasured confounders will involve the use of an instrumental variable (IV) (33) – weekend admission and PS trimming (34). A few IV sources in health

[6] Bayesian models will be used to overcome any bias due to sparsity of data as PS sub-classification in itself reduces the effective sample size.

studies have been described in Baiocchi (2014) (35). These include: distance to specialty, genes, insurance plan, timing of admission, calendar time and preference based IVs. Of relevance to this analysis would be timing of admission IVs. A study conducted by Berkley (2004) (36) in a Kenyan hospital demonstrated that children who were admitted during the weekend experienced higher mortality compared to those admitted during the weekdays – which is an indication of poor quality of care and treatment during the weekend. In other words, it is anticipated that children admitted during the weekdays would have better health outcomes. This, in theory, implies that the type of treatment and care received depend on the day of admission – and which later determines the type of health outcome of the patient.

Examining validity of multiple imputation

The analysis steps 1 – 3 above will exclude outcome data in the imputation model – however sensitivity analysis will include models in which the outcome variable is included in the imputation approach. This will aim to investigate if including outcome data in the imputation model has an influence[7].

Ethics and dissemination

The primary data are held by hospitals participating in the Kenyan Clinical Information Network (CIN) project with de-identifed data shared with the KEMRI-Wellcome Trust Research Programme for agreed analyses. The analyses described in this protocol are part of this larger project (CIN) which was approved by the Kenya Medical Research Institute Scientific and Ethical Review Committee (Protocol number: 2465). This committee agreed the use of de-identified patient data derived from retrospective case record review without gaining individual patient consent as is common practice in service evaluation research. The

[7] The primary interpretations will consider results of multiple imputations without outcome if results differ from those of MI with outcome – as is the standard recommendation to analysis of observational datasets in Rubin (2008) (37).

findings will be useful in understanding the external validity of current treatments – and will provide a platform on which to do more similar analyses for different (combinations of) treatments. The results of this analysis will be shared with the Kenyan Ministry of Health and will inform discussions on national pneumonia treatment guidelines to which the research team have made major prior contributions. The work will also be submitted for publication.

Competing Interests

The authors declare they have no competing interests.

Authors Contributions

The contributions of the authors were as follows: LM did an initial draft of this manuscript with the support of RP, EM and ME. Thereafter, all authors edited subsequent versions and approved the final copy.

Funding

We are grateful for the funds from the Wellcome Trust (#097170) that support ME through a fellowship and additional funds from a Wellcome Trust core grant awarded to the KEMRI-Wellcome Trust Research Programme (#092654) that supported this work. LM is supported by a Nuffield Department of Medicine Prize DPhil Studentship and Clarendon Scholarship (Oxford University). The funders had no role in drafting or submitting this manuscript.

References

1. Hospital Care for Children [Internet].; 2005 []. Available from: http://apps.who.int/iris/bitstream/10665/43206/1/9241546700.pdf.

2. Hospital Care for Children [Internet].; 2013 []. Available from: http://apps.who.int/iris/bitstream/10665/81170/1/9789241548373_eng.pdf.

3. Sazawal S, Black R. Pneumonia Case Management Trials Group. Lancet Infect Dis. 2003;3(9):547-56.

4. Agweyu A, Kibore M, Digolo L, Kosgei C, Maina V, Mugane S, et al. Prevalence and correlates of treatment failure among Kenyan children hospitalised with severe community-acquired pneumonia: a prospective study of the clinical effectiveness of WHO pneumonia case management guidelines. Tropical Medicine & International Health. 2014.

5. Mulholland K, Carlin JB, Duke T, Weber M. Challenges of trials of antibiotics for pneumonia in low-income countries. The Lancet Respiratory Medicine. 2014;2(12):952-4.

6. Ayieko P, Ogero M, Makone B, Julius T, Mbevi G, Nyachiro W, et al. Characteristics of admissions and variations in the use of basic investigations, treatments and outcomes in Kenyan hospitals within a new Clinical Information Network. Archives of Disease in Childhood. 2015.

7. Agweyu A, Opiyo N, English M. Experience developing national evidence-based clinical guidelines for childhood pneumonia in a low-income setting - making the GRADE? BMC Pediatric. 2012.

8. Ayieko P, English M. Case Management of Childhood Pneumonia in Developing Countries. Pediatr Infect Dis J. 2007;26:432-40.

9. Basic Paediatric Protocols: November 2013 Edition [Internet].: Ministry of Health; 2013 []. Available from: http://www.idoc-africa.org/index.php/86-clinical-guide/ken-guide/134-basic-paediatric-protocols-november-2013-edition.

10. Basic paediatric protocols [Internet].: Kenyan Ministry of Health; 2016 []. Available from: https://www.tropicalmedicine.ox.ac.uk/_asset/file/basic-paediatric-protocols-2016.pdf.

11. Graham SM, English M, Hazir T, Enarson P, Duke T. Challenges to improving case management of childhood pneumonia at health facilities in resource limited settings. Bulletin of the WHO. 2015.

12. Atkinson M, et al. A multicentre randomised controlled equivalence trial comparing oral amoxicillin and intravenous benzyl penicillin for community acquired pneumonia in children PIVOT Trial. Thorax. 2007;61:1102-6.

13. Hazir T, et al. Ambulatory short-course high-dose oral amoxicillin for treatment of severe pneumonia in children: a randomised equivalency trial. Lancet. 2008;371:49-56.

14. Addo-Yobo E, Chisaka N, Hassan M, Hibberd P, Lozano JM, Jeena P, et al. Oral amoxicillin versus injectable penicillin for severe pneumonia in children aged 3 to 59 months: a randomised multicentre equivalency study. Lancet. 2004;364:1141-8.

15. Addo-Yobo E, et al. Outpatient treatment of children with severe pneumonia with oral amoxicillin in four countries: the MASS study. Tropical Medicine and International Health. 2011;16:995-1006.

16. IVAC. Pneumonia and Diarrhoea Progress Report 2014. John Hopkins Bloomberg School of Public Health; 2014.

17. Onyango D, Kikuvi G, Amukoye E, Omolo J. Risk factors of severe pneumonia among children aged 2-59 months in western Kenya: a case control study. The Pan African Medical Journal. 2012(13):45.

18. Agweyu A, Gathara D, Oliwa J, Muinga N, Edwards T, Allen E, et al. Oral amoxicillin versus benzyl penicillin for severe pneumonia among kenyan children: a pragmatic randomised controlled noninferiority trial. Clin Infect Dis. 2015.

19. Stopler E, Van Roeyn P, Van de Wiel M, Van Bokhoven M, Houben P, Van der Weijde T, et al. Consensus on gut feelings in general practice. BMC Family Practice. 2009.

20. Rubin DB. For Objective Causal Inference, Design Trumps Analysis. 2008;2(3):808-40.

21. Draft recommendation for management of children with severe febrile illness and impared circulation without signs of severly impaired circulation: Guideline Panel Meeting [Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idoc-africa.org/images/documents/guidelines/Panel%20recommendation%20on%20fluid%20resuscitation%20I%20final.pdf.

22. Draft recommendations for care of umbilical cord in newborns: Guideline Panel Meeting. [Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idoc-africa.org/images/documents/guidelines/Panel%20recommendations%20on%20cord%20care%20final.pdf.

23. Draft recommendations for hospital management of children (less than 5 years) with sickle cell disease: Guideline Panel Meeting 2013. [Internet].: Health Services, Implementation Research and Clinical Excellence Collaboration; 2013 []. Available from: http://www.idocafrica.org/images/documents/guidelines/Panel%20recommendations%20on%20sickle%20cell%20anaemia%20final.pdf.

24. Wittes J. Sample size calculations for randomised controlled trials. Epidemiologic Reviews. 2002;24(1).

25. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? International Journal of Methods in Psychiatric Research. 2011;20(1):40-9.

26. Austin PC. Assessing balance in measured baseline covariates when using many-to-one matching on the propensity-score. Pharmacoepidemiol Drug Saf. 2008 2008;17(12):1218-25.

27. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009 10 Nov 2009;28(25):3083-107.

28. Stuart EA. Matching methods for causal inference:A review and a look forward. Statist. Sci. 2010;25(1):1-21.

29. Fine J, Gray R. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.

30. Rubin DB. An overview of Multiple Imputation. 1988.

31. Rosenbaum PR. Design sensitivity and efficiency in observational studies. Journal of the American Statistical Association. 2010;105:692-702.

32. Roderick JAL. Regression With Missing X's: A Review. Journal of the American Statistical Association. 1992;87(420):1227-37.

33. Chiba Y. Bias analysis of the instrumental variable estimator as an estimator of the average causal effect. Contemporary Clinical Trials. 2010 January 2010;31(1):12-7.

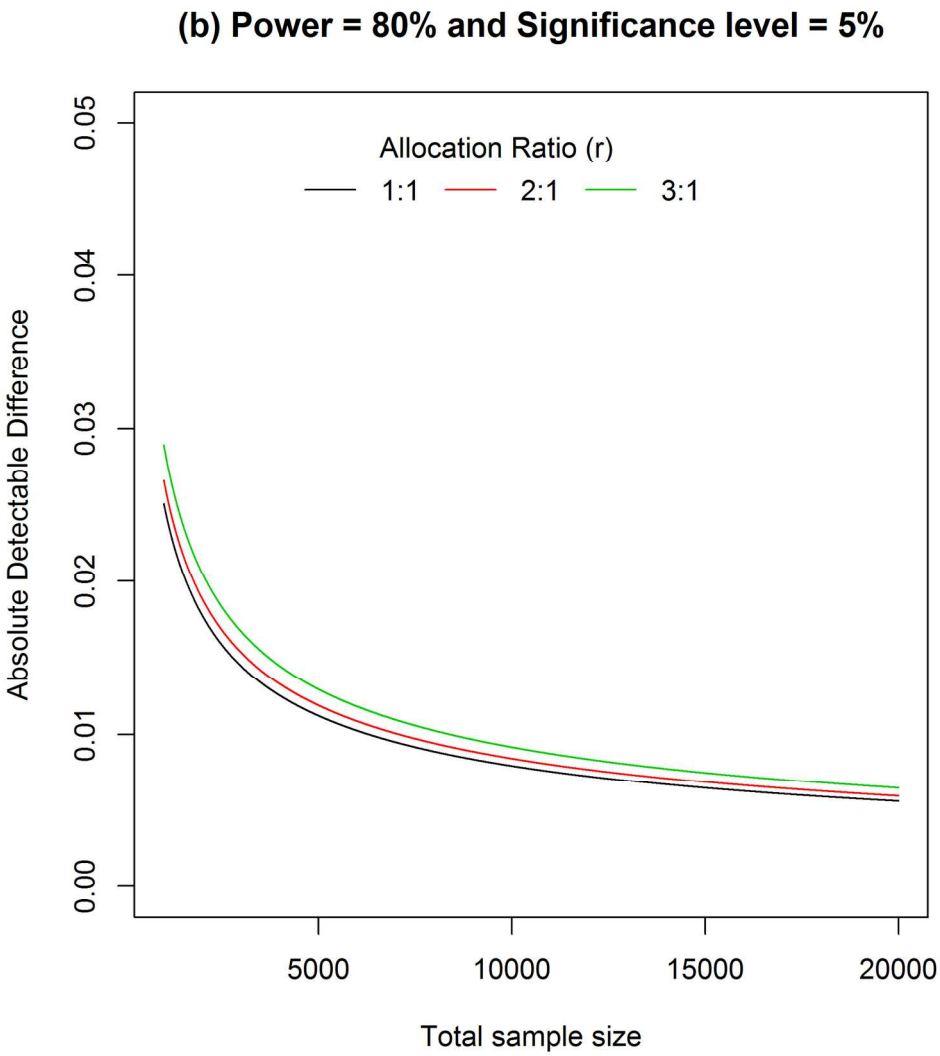
34. Lee BK, Lessler J, Stuart EA. Weight Trimming and Propensity Score Weighting. PLoS ONE [Electronic Resource]. 2011.

35. Baiocchi M, Cheng J, Small DS. Tutorial in Biostatistics: Instrumental Variable Methods for Causal Inference. Statistics in Medicine. 2014;33(13):2297-340.

36. Berkley JA, Brent A, Mwangi I, English M, Maitland K, Marsh K, et al. Mortality among Kenyan children admitted to a rural district hospital on weekends as compared with weekdays . Pediatrics. 2004;114:1737-8.

37. Rubin D.B. For Objective Causal Inference, Design Trumps Analysis. The Annals of Applied Statistics. 2008;2(3):808-40.

38. Tuti T, Bitok M, Malla L, Paton C, Muinga N, Gathara D, et al. Improving documentation of clinical care within a clinical information network: an essential initial step in efforts to understand and improve care in Kenyan hospitals. BMJ Global Health. 2016.



169x192mm (300 x 300 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/ item	ItemNo	Description	Line Number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1 – 3
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	Not Applicable
	2b	All items from the World Health Organization Trial Registration Data Set	
Protocol version	3	Date and version identifier	Not Applicable
Funding	4	Sources and types of financial, material, and other support	304 – 308
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	300 – 303
	5b	Name and contact information for the trial sponsor	Not Applicable
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	308
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	Not Applicable
Introduction			

Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	57 – 103
	6b	Explanation for choice of comparators	57 – 103
Objectives	7	Specific objectives or hypotheses	104 – 152
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	158 – 170
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	155
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	172 – 177
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	56 – 103
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	Not applicable
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	Not applicable
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	Not applicable

Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	197
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Not applicable
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	214 – 230
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	Not applicable
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	Not applicable
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	Not applicable

Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	Not applicable
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	Not applicable
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	Not applicable
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	155
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	Not applicable
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	155
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	231 – 291

	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	231 – 291
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	Not applicable
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	Not applicable
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	Not applicable
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	Not applicable
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	292 – 298
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	Not applicable

Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	Not applicable
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	Not applicable
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	Not applicable
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	298
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	Not applicable
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	Not applicable
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	295 – 296
	31b	Authorship eligibility guidelines and any intended use of professional writers	
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	

Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated.