

BMJ Open Comparative effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in Kenya: a retrospective observational study

Lucas Malla,¹ Rafael Perera-Salazar,² Emily McFadden,² Mike English^{1,3}

To cite: Malla L, Perera-Salazar R, McFadden E, *et al.* Comparative effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in Kenya: a retrospective observational study. *BMJ Open* 2017;7:e019478. doi:10.1136/bmjopen-2017-019478

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-019478>)

Received 5 September 2017
Revised 20 October 2017
Accepted 23 October 2017



CrossMark

¹Nuffield Department of Medicine, University of Oxford, Oxford, UK

²Nuffield Department of Primary Health Care Sciences, University of Oxford, Oxford, UK

³Health Services Unit, Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya

Correspondence to

Lucas Malla;
lmalla@kemri-wellcome.org

ABSTRACT

Objectives Kenyan guidelines for antibiotic treatment of pneumonia recommended treatment of pneumonia characterised by indrawing with injectable penicillin alone in inpatient settings until early 2016. At this point, they were revised becoming consistent with WHO guidance after results of a Kenyan trial provided further evidence of equivalence of oral amoxicillin and injectable penicillin. This change also made possible use of oral amoxicillin for outpatient treatment in this patient group. However, given non-trivial mortality in Kenyan children with indrawing pneumonia, it remained possible they would benefit from a broader spectrum antibiotic regimen. Therefore, we compared the effectiveness of injectable penicillin monotherapy with a regimen combining penicillin with gentamicin.

Setting We used a large routine observational dataset that captures data on all admissions to 13 Kenyan county hospitals.

Participants and measures The analyses included children aged 2–59 months. Selection of study population was based on inclusion criteria typical of a prospective trial, primary analysis (experiment 1, n=4002), but we also explored more pragmatic inclusion criteria (experiment 2, n=6420) as part of a secondary analysis. To overcome the challenges associated with the non-random allocation of treatments and missing data, we used propensity score (PS) methods and multiple imputation to minimise bias. Further, we estimated mortality risk ratios using log binomial regression and conducted sensitivity analyses using an instrumental variable and PS trimming.

Results The estimated risk of dying, in experiment 1, in those receiving penicillin plus gentamicin was 1.46 (0.85 to 2.43) compared with the penicillin monotherapy group. In experiment 2, the estimated risk was 1.04 (0.76 to 1.40).

Conclusion There is no statistical difference in the treatment of indrawing pneumonia with either penicillin or penicillin plus gentamicin. By extension, it is unlikely that treatment with penicillin plus gentamicin would offer an advantage to treatment with oral amoxicillin.

Strengths and limitations of this study

- This study provides a platform to explore effectiveness of alternative treatments in routine care in a low income setting to improve health outcomes for children.
- The analysis is 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, has however been assessed through the use of alternative methods as instrumental variables.

INTRODUCTION

WHO recommendations guide treatment for millions of children with pneumonia every year across low-income and middle-income countries.¹ These guidelines are largely based on moderate certainty in evidence of effects.^{2–5} However, trials supporting recommendations for hospitalised children have included fewer participants from Africa than other settings⁶ and it is suggested that African children with pneumonia have higher mortality.⁷ Additionally, trial populations may not always include the heterogeneous populations presenting for care, many of whom at hospital level may have comorbidity.⁸ Thus, despite improving access to recommended treatments and deployment of childhood vaccines at high coverage, including those against *Haemophilus influenzae* type B and pneumococcus, clinically diagnosed pneumonia remains one of the top causes of mortality in children under 5 years of age in Kenya and other countries.⁷ According to the mortality data derived from the Global Health Observatory Data—published in the WHO website,⁹ pneumonia caused about 5.4 under five deaths per 1000

Box Clinical pneumonia classifications and treatments in use in Kenya

1. Severe pneumonia: If a child has either oxygen saturation less than 90% or central cyanosis or is grunting or unable to drink or not alert, then she/he is classified as having severe pneumonia and is put on oxygen and treated with a combination of gentamicin and penicillin. The previous WHO²³ and pre-2016 Kenyan guidelines²⁰ named this class as 'very severe pneumonia'.
2. Indrawing pneumonia: If a child has lower chest wall indrawing (but does not have any of the qualifying signs for severe pneumonia above) and is alert then she/he is classified as having indrawing pneumonia. In previous WHO²³ and pre-2016 Kenyan guidelines²⁰, this class was named as 'severe pneumonia' and treatment recommended was inpatient penicillin monotherapy. Our analyses are based on data from the period before March 2016 when inpatient penicillin monotherapy was recommended for this population. Since March 2016 in Kenya, and reflecting updated WHO guidance and results of a local trial,²⁴ it has been recommended that this group be treated in outpatient settings with oral amoxicillin as part of an expanded group of non-severe pneumonia. Note: The term indrawing pneumonia is hereafter used in this analysis to define this category of children to avoid confusion.
3. Non-severe pneumonia: If a child has none of the clinical signs in the two categories above but has cough or difficulty breathing and a respiratory rate greater than or equal to 50 breaths/min (for age between 2 and 11 months) or respiratory rate greater than or equal to 40 breaths/min (for age above 12 months) then Kenyan guidelines in the period pre-March and post-March 2016 recommend she/he is classified as having non-severe pneumonia and treated with oral amoxicillin as an outpatient.

children in 2015 (which was the highest compared with diarrhoea/dehydration and malaria which are the other top causes of under five mortality in Kenya). The comparison of mortality rates between 2000 and 2015 for pneumonia, diarrhoea/dehydration and malaria is presented in online supplementary file 1: figure A. The basic and pneumococcal vaccine coverage by 2014 for children aged 12–23 months in Kenya was at least 80%.¹⁰

In a recent change to guidance, it is now recommended that pneumonia characterised by lower chest wall indrawing be treated in outpatient settings with oral medication (box).^{11 12} Yet it remains associated with non-trivial mortality that may be higher outside trial populations.¹³ Residual mortality may be associated with causes that are not prevented by currently available conjugate vaccines and organisms, which are not susceptible to the antibiotics currently recommended. Establishing whether there are benefits of alternative treatment regimens to help reduce mortality would ideally require large, pragmatic clinical trials.^{14 15} However, these remain relatively expensive and time consuming. Observational data may support comparative effectiveness analyses of alternative treatments, may be cheaper and quicker, and may enable evaluation of interventions for which randomisation is difficult.¹⁶ We use observational data from Kenya to address an important contemporary question for the

treatment of pneumonia, a comparison of the effectiveness of gentamicin plus penicillin versus penicillin alone for the treatment of indrawing pneumonia in routine settings. The only previous clinical trial comparing these treatments was a small study of 40 patients in Malaysia.¹⁷ In so doing, we examine the potential of using data collected by providers as part of their routine practice for comparative effectiveness research in an African setting.

METHODS

Clinical definitions of pneumonia, primary and secondary analyses

WHO and Kenyan pneumonia treatment guidelines are implicitly based on risk stratification of illness with children deemed at higher risk of mortality offered broader spectrum antibiotic regimens and those at lower risk narrower spectrum antibiotics.^{11 18–20} We present three categories of clinically diagnosed pneumonia in box. This categorisation outlines previous and recently revised WHO and Kenyan pneumonia treatment guidelines.^{11 19} What we refer to as indrawing pneumonia may be associated with low but clinically significant mortality rates.^{13 21} Prior to March 2016, recommended treatment for this group was penicillin monotherapy and our aim is to examine whether there is any advantage of broader spectrum antibiotics in this group. Since March 2016, new guidelines recommend outpatient treatment with oral amoxicillin for this group on the basis of trials suggesting equivalence of amoxicillin and penicillin. However, as indicated above, very few patients had been included in studies comparing narrow (amoxicillin or penicillin) and broader spectrum antibiotic regimens. As indicated above, beyond the confines of clinical trials among all children being treated for indrawing pneumonia, clinical outcomes (including mortality) are worse than seen in the trials⁷ and clinicians are often choosing not to use a single drug regime and are in fact often opting to use the combination of gentamicin and penicillin in the group meeting criteria for indrawing pneumonia in real life settings.²² As mortality is higher in real-life settings than in trials and as the possibility that broad spectrum antibiotics could have an advantage over monotherapy with penicillin (or amoxicillin) has not been explored in Kenya's previous trials, we feel that examining whether broad spectrum antibiotics confer an advantage is an important question.

The ability to use routine data to compare treatment effects requires that patients with similar problems receive different treatments. Previous studies conducted in Kenya and elsewhere have indicated that clinicians often do not follow guideline recommendations in treating pneumonia.²² Variation from the guideline recommended approach can occur at the point of pneumonia severity assignment (clinicians do not follow a nationally approved protocol linking clinical signs and severity category outlined in box) and at the point of treatment assignment (clinicians do not follow this protocol that links treatment and severity). This variability in adherence to protocols

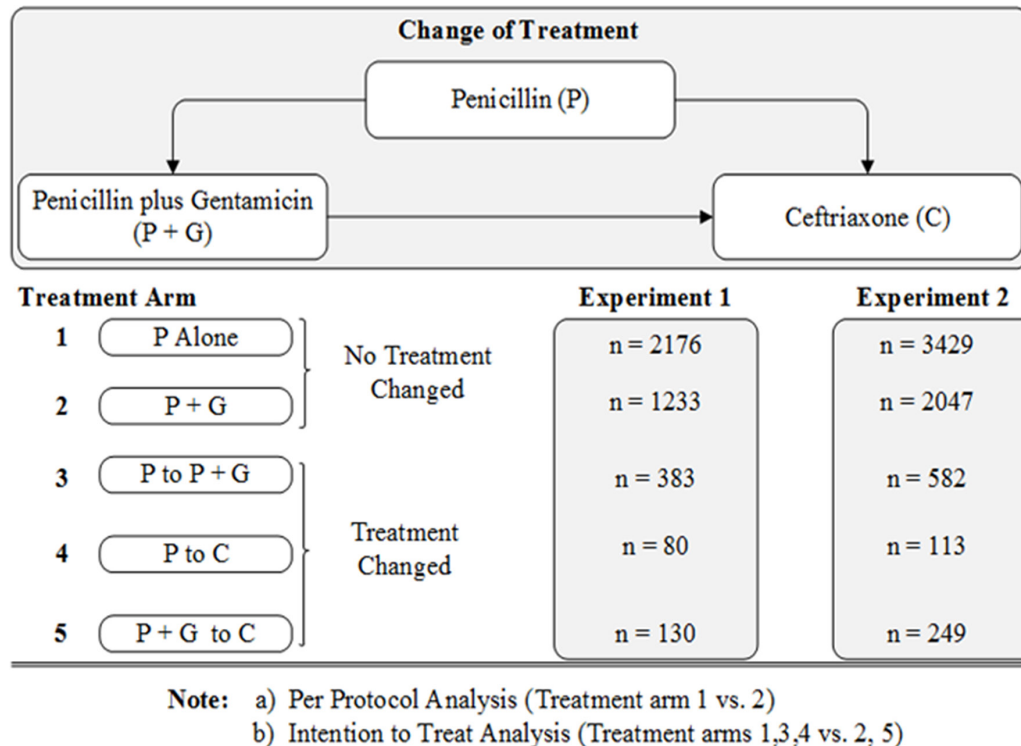


Figure 1 Summary of patients per treatment arm in Experiments 1 and 2.

provides the opportunity for comparative effectiveness evaluation. More specifically, the adherence and non-adherence to treatment protocols by clinicians allows us to classify indrawing pneumonia admissions in two ways:

1. Those with clinical signs placing them in the group of indrawing pneumonia irrespective of the category or classification assigned to the child by the clinician.
2. Those given a clinician classification of indrawing pneumonia irrespective of the actual clinical signs observed by the clinician.

Based on these two possibilities, two experiments were designed (see online supplementary file 2: analysis protocol²⁵) with specific objectives as follows¹:

1. Experiment 1: to compare effectiveness of injectable penicillin versus penicillin plus gentamicin (both injectable) in treatment of indrawing pneumonia, where the child is identified as belonging to a population of children with indrawing pneumonia on the basis of data on their recorded clinical signs. Experiment 1 population of indrawing pneumonia is therefore consistent with pre-2016 clinical guideline recommendations.
2. Experiment 2: to compare effectiveness of injectable penicillin versus penicillin plus gentamicin in a population in which we use the clinician assigned categorisation of indrawing pneumonia, which

may not be consistent with clinical guideline recommendations.

We defined experiment 1 as our primary analysis as we propose it would identify a population similar to that recruited to a randomised trial where the inclusion criteria would be based on specified clinical signs. Experiment 2 offers a scenario that may represent a more pragmatic study design with inclusion criteria based around a clinician-led classification.

Data source

We use data from the Kenyan Clinical Information Network (CIN) that was initiated to improve inpatient paediatric data availability from county (formerly district) hospitals. Thirteen county referral hospitals were purposively selected with direction from Ministry of Health and recruited into the CIN. These hospitals were recruited into the study at different times; four in September 2013, five in October 2013 and four in February 2014. This analysis uses data up to March 2016. On average, 25 000 paediatric admissions are captured per year. These hospitals typically have one paediatrician leading services predominantly provided by junior clinical teams. Data systems and standardised clinical forms were specifically implemented in all hospitals at the start of this work to optimise the quality of routine data. Patient data in these hospitals are collected post-discharge by trained data clerks guided by well-defined standard operating procedures, under supervision by the hospital medical records department and the research team. Clinicians admitting patients fill standardised Paediatric Admission Record forms²⁶ that have

¹All children with danger signs were excluded from experiment 1 and in general (both in experiments 1 and 2), children with the following comorbidities were excluded: HIV, meningitis, tuberculosis and/or acute severe malnutrition.

been shown to improve documentation of clinical symptoms and signs.²⁷ Together with discharge forms, treatment sheets and laboratory reports these are all part of the patient files that are the primary data source. This data collection system has been described in detail elsewhere.²⁸ Feedback to hospitals as part of the CIN activities has helped improve the quality of clinical data.²⁸ The description of hospital selection and their populations of patients is detailed elsewhere.²⁹

Statistical analysis

Defining per protocol and intention to treat populations

In typical randomised controlled trials, types of analyses to be conducted are defined beforehand—and this involves defining the type of patient populations that are included in the analyses. Intention to treat and per protocol populations derived from observational datasets have been described by Danaei *et al.*³⁰ We defined per protocol and intention to treat populations based on the dates actual treatments were recorded as prescribed for patients included in our primary and secondary analyses (experiments 1 and 2, respectively). Within each experiment, and after applying inclusion and exclusion criteria, we define the per protocol population as those whose prescription of one of the two study regimens did not change during the admission. The intention to treat population is defined by the original treatment assignment and included children in whom treatment was subsequently changed (see [figure 1](#) in the Results section).

Dealing with missing data and propensity score matching

As CIN comprises data from routine care settings, it faces challenges of non-random treatment allocation and missing data. The missing data and propensity score (PS) methods for this analysis have been detailed in the online supplementary file 2: analysis protocol linked to this work.²⁵ In brief, after exploring the patient populations, 20 datasetsⁱⁱ³¹ were derived using multiple imputation (with chained equations) for each experiment (all the variables in both the experiments had missing data less than 30%; see online (supplementary file 1: table A). Clinical signs and symptoms data considered were those recorded by clinicians before patients were admitted. The multiple imputation excluded outcome data as guidance on the use of observational datasets for comparative effectiveness analysis recommends exclusion of outcome data in the design phase.³² Following this, those with missing outcome data were excluded from the analysis (missingness in the outcome data were 0.5% and 0.8% for experiments 1 and 2). For each imputed dataset, patients in the alternative treatment groups (penicillin monotherapy vs penicillin plus gentamicin) were then matched using PS

methods to overcome non-random treatment allocation. PS define the probability of belonging to or being assigned a given treatment based on signs and symptoms.³³ PS is a distance measure³⁴ which is used as a means to overcome allocation bias as treatment outcomes in children with similar PSs can then be compared. In these analyses, we compared three approaches to reducing possible bias based on PS—optimal full matching, weighting and subclassification.^{33 34} All are aimed at creating groups of patients that are comparable in terms of the distribution of observed signs and symptoms. For each experiment, in order to select the optimum PS implementation method, absolute standardised mean differences (ASMD) were used as diagnostic checks for covariate balance and overlap^{35 36} between the alternative treatment groups. PS methods that resulted in the minimum average ASMD for the majority of the variables while retaining the largest number of patients in the analysis were considered the most appropriate.³⁴

Analytic modelling and sensitivity analyses

In sample size calculations conducted prior to the experiments (presented in greater detail elsewhere; see online supplementary file 2: analysis protocol), it was estimated that a sample size of at least 4000 would be sufficient for the planned experiments to detect a minimum difference of 1.5% in mortality between the two treatment groups. The sample size for experiment one was 4002 and experiment two 6420 (including 3312 of those that were also in experiment 1). In other words, experiment 2 largely included those in the experiment 1 population but also children not meeting eligibility criteria for experiment 1. For each of the experiments, after multiple imputation, multivariable log-binomial regression models were fitted to PS-weighted datasets and adjusting for all the variables also used in the PS models (also as a form of sensitivity analyses, treatment effects were estimated on PS-unweighted datasets). Only pooled treatment effect estimates are reported.

One possibility is that clinicians' treatment assignment is skewed such that patients who appear sicker (having a greater number of clinical signs of more severe illness) are assigned 'stronger' or broad spectrum treatment. In this situation as mentioned by Stürmer *et al.*,³⁷ specific types of treatment allocation may be more likely associated with increased mortality.³⁷ In theory, the use of PSs is supposed to account for such skewed assignment by comparing only outcomes of those with similar PSs assumed to suggest they have similar clinical profiles and thus similar risks. PS trimming attempts to tackle this problem further by excluding patients who are at the extremes of the PS distribution to create a population with clinical characteristics that are as homogeneous as possible. We use PS trimming to define a population between the 5%–95% PS percentiles in a sensitivity analysis.

In a further sensitivity analysis, we used an instrumental variable to examine the potential influence of any unmeasured variables.³⁸ An instrumental variable

ⁱⁱThe current literature³¹ recommends the use of more than five imputed datasets and, therefore, 20 should be sufficient.

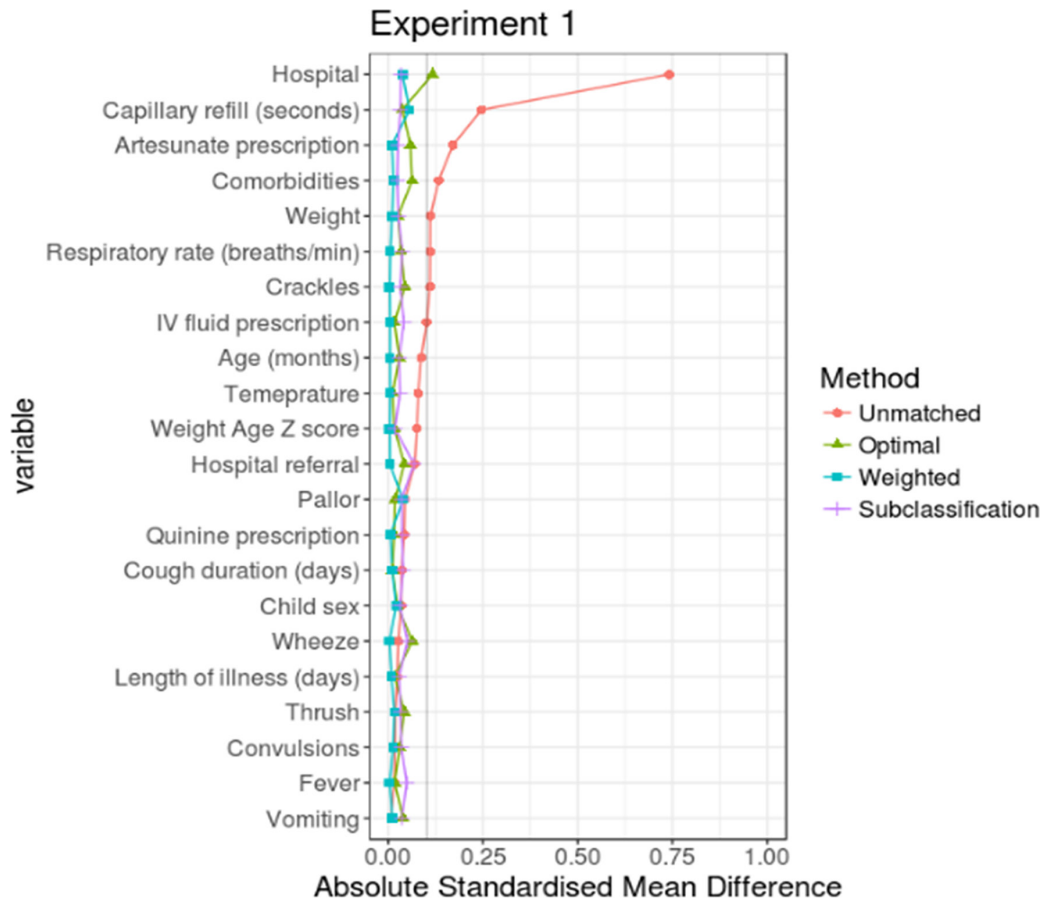


Figure 2 Comparing performance of the three propensity score (PS) implementation methods in experiment 1. The y-axis contains all the variables used in the PS models. While x-axis shows absolute standardised mean difference (ASMD) which is a measure of covariate balance between the two treatment groups. An ASMD value of $\leq 10\%$ indicates that the method has performed well in creating comparable groups.

method aims to find a proxy randomised experiment in a routine or observational dataset.³⁹ We used weekend/weekday admission as an instrumental variable as it was demonstrated in a study conducted by Berkley *et al*⁴⁰ in a Kenyan hospital that children who were admitted during the weekend experienced higher mortality compared with those admitted during the weekdays. This, in theory, implies that the type of treatment and care received depends on the day of admission—and this later determines the type of health outcome of the patient. The process of fitting the instrumental variable models has been described in online supplementary file 1. The two sensitivity approaches described above were done for both primary and secondary analyses.

RESULTS

Creating per protocol and intention to treat populations

Examining the dates treatments were given, five treatment arms (per experimental scenario) were defined—specifically those who received¹: penicillin alone without changes,² a combination of penicillin plus gentamicin without changes,³ penicillin but switched to a combination of penicillin plus gentamicin,⁴ penicillin but switched to ceftriaxone and⁵ a combination of penicillin plus

gentamicin but switched to ceftriaxone (ceftriaxone is the recommended second line treatment for severe pneumonia). Therefore, per protocol analyses would compare patients in treatment arm 1 versus 2, while intention to treat analyses would compare patients in treatment arms 1, 3 and 4 versus 2 and 5 (figure 1).

In this analysis, intention to treat populations were considered primary and are reported in experiments 1 and 2 in keeping with clinical trial reporting guidelines. These analyses include a relatively larger number of patients compared with per protocol analyses. The recommended doses of penicillin and gentamicin in these hospitals are 50 000 IU/kg and 7.5 mg/kg given four times and one time per day, respectively. Additional data suggest most clinicians prescribed these doses correctly (see online supplementary file 1: table B).

Comparing performance of optimal full matching, weighting and PS subclassification in experiments 1 and 2, respectively

For each experiment, the three PS implementation methods were compared with determine the one which would result in the least ASMD for most of the variables in the analysis (even though all the three methods resulted in variables with $ASMD \leq 10\%$). For experiment 1, PS weighting performed better than PS optimal full

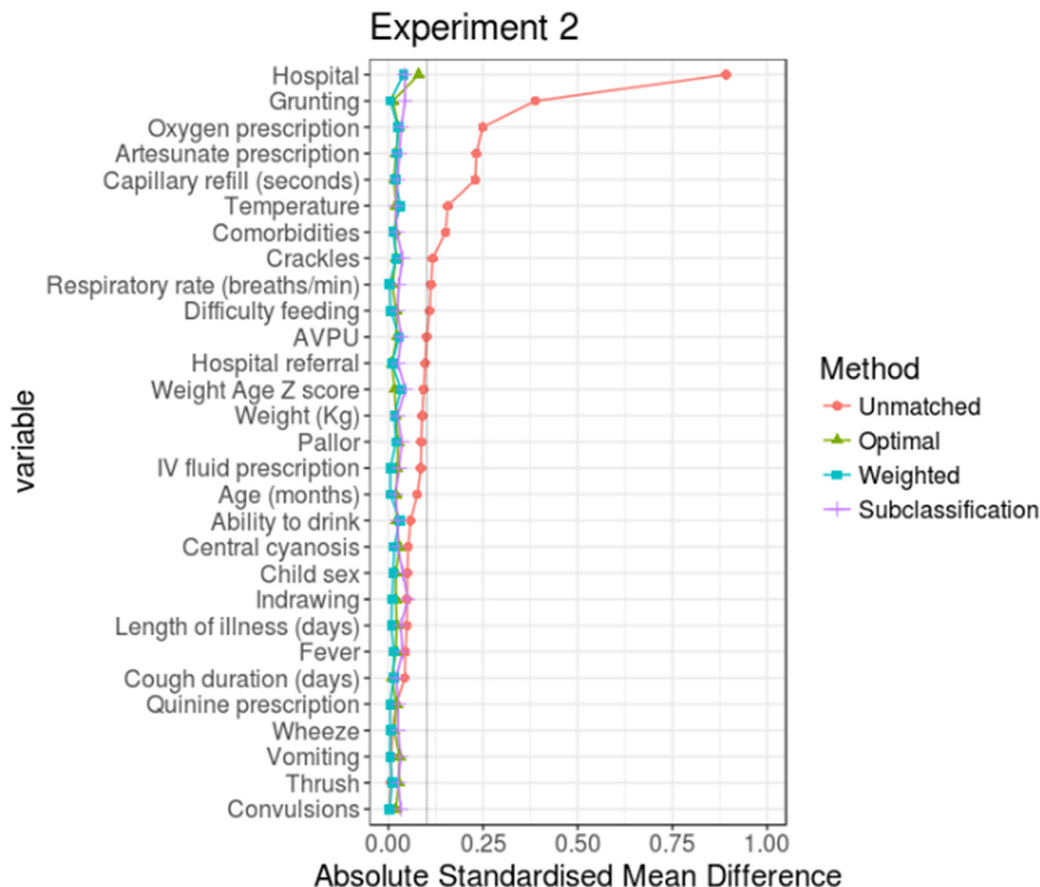


Figure 3 Comparing performance of the three propensity score implementation methods in experiment 2. AVPU, alert, verbal, pain, unresponsive.

matching and subclassification and for experiment 2, the performance of weighting was comparable to that of optimal full matching (see figures 2 and 3). In both experiments, PS subclassification reduced covariate imbalance the least. Thus, in the subsequent sections, outcome analyses are based on PS-weighted datasets for both experiments.

Outcome analysis results

Exploring mortality in raw datasets

Examining the raw datasets without PS adjustments in experiment 1, the average number of pneumonia deaths (across the 20 imputed datasets) in penicillin plus gentamicin group was 33/1363 (2.42%) and in penicillin monotherapy was 26/2639 (0.99%), and for experiment 2, the average number of deaths was 87/2296 (3.79%) and 50/4124 (1.21%) in penicillin plus gentamicin and penicillin monotherapy groups, respectively. Overall, the average number of pneumonia deaths in the penicillin plus gentamicin group was approximately two and a half to three times the number of mortality events in the penicillin monotherapy group in Experiments 1 and 2, respectively.

Modelling mortality risk ratios

The analysis considered penicillin monotherapy as the reference group and mortality as the outcome—and

therefore a risk ratio (RR) greater than one would be interpreted to favour penicillin over penicillin plus gentamicin. For both experiments, the treatment RRs estimated on the unmatched datasets were larger than the RR estimated on datasets obtained through PS weighting (see table 1 for all results). In experiment 2, the PS-unadjusted analysis showed that penicillin monotherapy was significantly more effective than penicillin plus gentamicin (1.68 (1.15 to 2.36)). However, the PS-weighted effect estimate (1.04 (0.76 to 1.40)) was much reduced and suggested that use of PS had corrected (to a degree)

Table 1 Treatment effect estimates

	Experiment 1, RR (95% CI)	Experiment 2, RR (95% CI)
Regression without PS adjustment	1.75 (0.94 to 2.77)	1.68 (1.15 to 2.36)
PS weighting	1.46 (0.85 to 2.43)	1.04 (0.76 to 1.40)
PS trimming (5%–95% restriction)	1.39 (0.74 to 2.15)	1.05 (0.74 to 1.41)
Instrumental variable	0.91 (0.41 to 2.20)	0.44 (0.34 to 1.32)

PS, propensity score; RR, risk ratio.

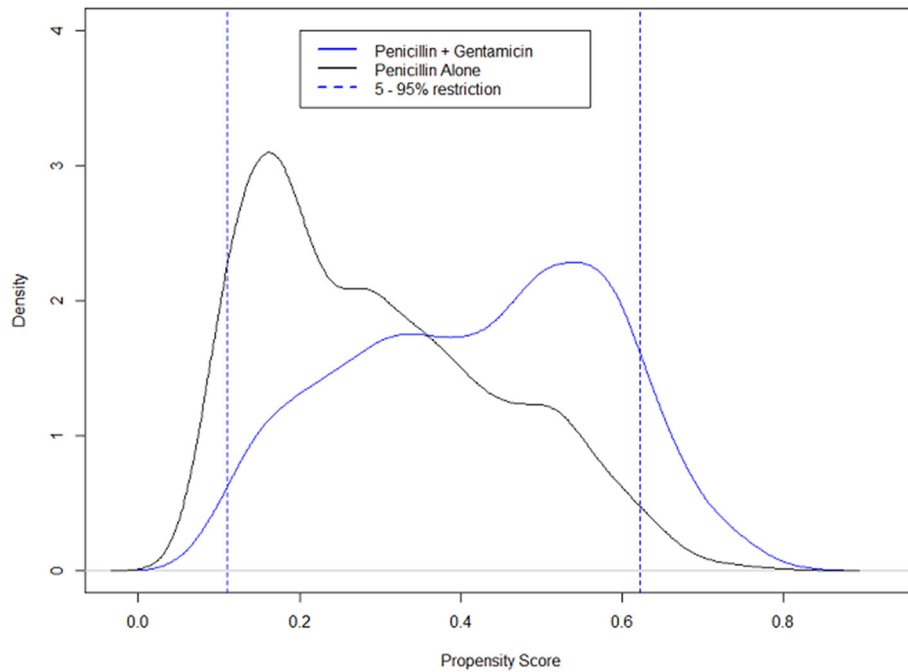


Figure 4 Experiment 1 propensity score (PS) distribution curves: The dotted lines show the distribution of PSs for patients in the 5%–95%. The continuous blue line shows the distribution of PSs for those who were given penicillin plus gentamicin. While the continuous black line shows the PS distribution for those who received penicillin alone.

for allocation bias indicating that there was no statistical difference in mortality outcomes between penicillin plus gentamicin and penicillin monotherapy treatments. We also observed that the adjusted point estimate for any effect difference in experiment 2 (1.04 (0.76 to 1.40)) was less than that in experiment 1 (1.46 (0.85 to 2.43)). This may be due to an increase in the number of covariables available for PS weighting that could be used in Experiment 2 resulting in closer matching (see online supplementary file 1: table C).

Sensitivity analysis through trimming using 5%–95% PS population restriction

After excluding 10% of the populations as a result of PS trimming in sensitivity analyses for Experiments 1 and 2, the resulting sample sizes were 3583 and 5778. The skewed assignment of children to treatment with gentamicin and penicillin is demonstrated by their higher PS scores in figure 4 for experiment 1 (and online supplementary file 1: figure B for experiment 2). As higher PS scores are associated with the presence of a greater number of clinical signs of illness, this also suggests an association between more severe illness and treatment with gentamicin and penicillin. For experiment 1, the estimated average mortality events (on PS-unadjusted datasets) were 26/1201 (2.16%) and 24/2382 (1.01%) for penicillin plus gentamicin and penicillin monotherapy groups. While the estimated events in experiment 2 were 62/2026 (3.06%) and 46/3752 (1.22%). Thus, in sensitivity analyses for both experiments, trimming excluded more mortality events in the penicillin plus gentamicin group compared with the penicillin monotherapy group. The treatment effects estimated using PS-weighted

models for the restricted populations as a result of PS trimming showed no statistical difference between the two treatments (table 1).

Sensitivity analysis through the use of weekend/weekday as an instrumental variable

In order to assess whether a timing of admission variable would form a natural and random experiment, the distributions of covariates were examined across the levels of the instrumental variable (weekend/weekday) in experiments 1 and 2. The distribution of each of the patient characteristics between weekend and weekday admissions was approximately similar (online supplementary file 1: table D) suggesting that weekend/weekday admission satisfactorily satisfies one of the criteria as a valid instrumental variable (also see online supplementary file 1 for the set of criteria for a valid IV). The weekend mortalities, in the raw datasets, seemed to be higher than weekday mortalities (see online supplementary file 1: table E).

The estimated treatment effects, both in experiments 1 and 2, suggest that there is no statistical difference in treating indrawing pneumonia with either penicillin alone or penicillin plus gentamicin. The effect estimates obtained using our IV in both experiments are less than one as compared with those obtained with PS weighting which are greater than one. Biologically, the effectiveness of gentamicin plus penicillin (when administered in correct doses) is expected to be the same or greater than that of penicillin monotherapy. Based on the magnitude and direction of effects, the use of the IV seems to demonstrate that the effects obtained through PS weighting may have had some residual bias. However, it is important to highlight that for all analyses the 95% CI obtained are

consistent with the Null Hypothesis of no different effect for the treatments.

DISCUSSION

We compared penicillin alone with penicillin plus gentamicin in treatment of indrawing pneumonia in populations with overall mortality of 1.5% and 2% in experiments 1 and 2, respectively. There were more fatal events in the penicillin plus gentamicin group than the penicillin group (approximately 2.5 times) and unadjusted analyses pointed, therefore, to a protective effect of penicillin treatment. However, adjusted analyses, both in experiments 1 and 2, that aim to account for allocation bias using PS weighting that can result from non-random treatment allocation suggest that there is no appreciable difference in outcomes between penicillin and gentamicin plus penicillin treatment of indrawing pneumonia. In addition, we conducted analyses using alternative PS methods—subclassification (results are presented in supplementary files: figures C,D and table F) and optimal full matching (results are presented in Supplementary table G) and analyses of both intention to treat and per protocol populations. All analyses showed similar findings (see online supplementary file 1: table H). We undertook two formal approaches to sensitivity analysis. First, we employed PS trimming to exclude 10% of the analysis populations in experiments 1 and 2. Effect estimates in this case are based on analyses of 90% of cases that PS suggest are best matched. Second, we used an instrumental variable. These techniques employ different approaches to account for possible confounding that might contribute to estimated treatment effects. Both these forms of analysis provided results that support the suggestion that poor outcome in this population is not associated with the antibiotic regimen received.

Our analyses were conducted using data from over 4000 children, 100 times more participants than were included in the only prior randomised controlled trial of penicillin monotherapy and penicillin plus gentamicin in treatment of pneumonia in an Asian population.¹⁷ There are continuing concerns of clinically important mortality in children with indrawing pneumonia in Africa.²¹ This has led to hesitation to adopt new WHO and Kenyan guidelines that now recommend the treatment of indrawing pneumonia as an outpatient using amoxicillin.^{11 19} Our results suggest that there are likely to be two distinct issues. First, they suggest that offering broader spectrum injectable antibiotic treatment to children with indrawing pneumonia may not improve outcomes compared with treatment with penicillin monotherapy. As other studies have suggested equivalence between oral (high dose) amoxicillin therapy and injectable penicillin therapy,^{2-5 24} it seems likely therefore that oral amoxicillin and penicillin plus gentamicin combination therapy would result in similar outcomes when used to treat indrawing pneumonia. Clinicians should therefore carefully adhere to guidelines for treatment of indrawing pneumonia and

avoid using gentamicin helping to prevent any possible toxicity.

Second, however, our results suggest that children fulfilling a definition of indrawing pneumonia based on clinical signs, and having excluded serious comorbidities, may still have an appreciable risk of mortality irrespective of their antibiotic treatment (1.5% in all children in experiment 1). When clinicians categorise children with indrawing pneumonia and imperfectly adhere to clinical sign-based guidance mortality tends to be higher (2% in all children in experiment 2). These findings point to as yet uncharacterised risk factors that could be important in determining which children need admission to hospital as current guidance indicates that all those with indrawing pneumonia can be treated as an outpatient. While offering an alternative antibiotic to amoxicillin to this group may not improve outcomes, it is possible that closer and continuing observation in hospital may help identify comorbid or alternative conditions that are contributing to this mortality and that may be treated.

The trials that informed the basis for the revised WHO guidelines²⁻⁵ showed extremely low mortality (0%–0.2%) suggesting that the populations included in such trials may not be directly representative of all those to whom guidelines are applied in routine settings. In the trial by Agweyu *et al*²⁴ conducted in Kenya (which compared penicillin versus oral amoxicillin for indrawing pneumonia), overall mortality was 0.8%.²⁴ In a parallel observational cohort providing data from the same hospitals over the same time period for children treated with penicillin alone but not included in the Kenyan, trial mortality was not significantly different but marginally higher at 1.2% (Agweyu, submitted, 2017) perhaps suggesting that even the limited exclusion criteria in this pragmatic trial might result in exclusion of some sicker children. Taken together with data from the analyses presented here it does appear that there is a need to explore whether guidelines might be modified to accommodate additional clinical risk factors for possible life-threatening illness that should prompt admission. In a population with high coverage with conjugate vaccines, this may more usefully be for more rigorous evaluation to identify alternative diagnoses or for improved supportive care than for different antibiotics.

Strengths and limitations

Conducting comparative effectiveness analyses using observational datasets can offer the advantage of larger sample sizes at lower cost than randomised controlled clinical trials. They also include patients that may not qualify for enrolment in a typical explanatory randomised controlled trial—and therefore perhaps provide more true to life estimates of treatment effects similar to those observed in highly pragmatic trials.¹⁵ However, as most observational datasets are not meant for research, they have challenges of non-random treatment allocation and missing data. We employed a rigorous ‘experimental design’ strategy as is recommended when

using observational data.³² We used PS and multiple imputation methods in an effort to minimise bias due to non-random treatment allocation and missing data and analyses suggested no appreciable difference in outcomes of indrawing pneumonia treated with penicillin alone compared with penicillin plus gentamicin. This was in contrast to unadjusted regression analyses that pointed towards better outcomes with penicillin alone suggesting the presence of allocation bias. As most observational datasets are limited to observed variables, it is important to conduct sensitivity analysis to explore if the estimated effects are potentially sensitive to unmeasured variables. We used an instrumental variable and PS trimming, both supported the idea of no appreciable difference regimens when treating indrawing pneumonia. While there are differences (in terms of magnitude) in the mortality observed in the different groups that suggest some residual bias in treatment allocation, these mortality differences are no greater than might occur by chance after PS adjustment (with the type 1 and 2 errors specified in online supplementary file 2: analysis protocol). In that sense, the PS approach may still have limitations but it does allow us to conclude no statistical difference in mortality outcomes between the two treatment arms.

The WHO recommended guidelines for treating pneumonia have considerable influence on policy and practice in low-income and middle-income countries. While the evidence base and rigour of guideline development have improved considerably, there remain few data on their effectiveness when implemented in non-trial settings. Even though well-designed, large pragmatic trials would be preferred, we demonstrate that carefully collected routine data may be useful for assessing the effectiveness of alternative treatments.¹⁵ Such analyses may become increasingly possible as electronic medical records are deployed in low-income and middle-income countries⁴¹ but it is important that such studies are carefully designed to limit as far as possible the biases that arise from non-random treatment allocation.³² Our results suggest that children with indrawing pneumonia may gain little benefit from treatment with broader spectrum antibiotic regimens. However, they also suggest that further work is needed to identify those who are at higher risk of death who might be prioritised for an inpatient diagnostic work-up and improved supportive care rather than treated as outpatients.

Acknowledgements The authors would like to thank Ambrose Agweyu for his comments that helped improve this manuscript. The authors also appreciate the valuable contribution offered by the CIN team: Lydia Thurairara & Grace Ochieng' (Kiambu County Hospital), Barnabas Kigen (Busia County Hospital), Melab Musabi & Rachel Inginia (Kitale County Hospital), Anne Kamunya & Sam Otido (Embu County Hospital), Margaret Kuria (Kisumu East County Hospital), Agnes Mithamo & Francis Kanyingi (Nyeri County Hospital), Celia Muturi, Caren Emadau & Cecilia Mutiso (Mama Lucy Kibaki County Hospital), David Kimutai & Loice Mutai (Mbagathi County Hospital), Nick Aduro (Kakamega County Hospital), Samuel Ng'arng'ar (Vihiga County Hospital), Fred Were & David Githanga (Kenya Paediatric Association) and Rachel Nyamai (Ministry of Health).

Contributors LM did an initial draft of this manuscript with the support of RP-S, EM and ME. Thereafter, all authors edited subsequent versions and approved the final copy.

Funding The authors are grateful for the funds from the Wellcome Trust (#091710) 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).

Disclaimer The funders had no role in drafting or submitting this manuscript.

Competing interests None declared.

Ethics approval This analysis is based on a larger project (CIN) which was cleared by the Kenya Medical Research Institute Ethics and Review Board (Protocol number: 2465).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The hospital-specific datasets are in custody of the hospitals participating in CIN (these datasets have been deidentified).

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Olayemi E, Asare EV, Benneh-Akwasi Kuma AA. Guidelines in lower-middle income countries. *Br J Haematol* 2017;177:846–54.
- Atkinson M, Lakhanpaul M, Smyth A, et al. Comparison of oral amoxicillin and intravenous benzyl penicillin for community acquired pneumonia in children (PIVOT trial): a multicentre pragmatic randomised controlled equivalence trial. *Thorax* 2007;62:1102–6.
- Hazir T, Fox LM, Nisar YB, 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.
- Addo-Yobo E, Chisaka N, Hassan M, 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.
- Addo-Yobo E, Anh DD, El-Sayed HF, et al. Outpatient treatment of children with severe pneumonia with oral amoxicillin in four countries: the MASS study. *Trop Med Int Health* 2011;16:995–1006.
- Mulholland K, Carlin JB, Duke T, et al. The challenges of trials of antibiotics for pneumonia in low-income countries. *Lancet Respir Med* 2014;2:952–4.
- John Hopkins Bloomberg School of Public Health. Pneumonia and Diarrhoea Progress Report 2014. 2014 <http://www.jhsph.edu/research/centers-and-institutes/ivac/resources/IVAC-2014-Pneumonia-Diarrhea-Progress-Report.pdf>
- Berkley JA, Lowe BS, Mwangi I, et al. Bacteremia among children admitted to a rural hospital in Kenya. *N Engl J Med* 2005;352:39–47.
- Global Health Observatory data repository. 2016 <http://apps.who.int/gho/data/node.main.686?lang=en>
- Kenya Demographic and Health Survey. 2014 <http://dhsprogram.com/pubs/pdf/FR308/FR308.pdf>
- Kenyan Ministry of Health. Basic paediatric protocols. 2016 https://www.tropicalmedicine.ox.ac.uk/_asset/file/basic-paediatric-protocols-2016.pdf
- Hospital Care for Children. 2013 http://apps.who.int/iris/bitstream/10665/811170/1/9789241548373_eng.pdf
- UNICEF DATA: Monitoring the situation of women and children. 2017 <https://data.unicef.org/topic/child-health/pneumonia/>
- Ye C, Beyene J, Browne G, et al. Estimating treatment effects in randomised controlled trials with non-compliance: a simulation study. *BMJ Open* 2014;4:e005362.
- Thadhani R. Formal trials versus observational studies. In: Mehta A, Beck M, Sunder-Plassmann G, eds. *Fabry Disease: Perspectives from 5 Years of FOS*. Oxford: Oxford PharmaGenesis. 14, 2006. <https://www.ncbi.nlm.nih.gov/books/NBK11597/>

16. West SG, Duan N, Pequegnat W, *et al.* Alternatives to the randomized controlled trial. *Am J Public Health* 2008;98:1359–66.
17. Hasali MA, Ibrahim MI, Sulaiman SA, *et al.* A clinical and economic study of community-acquired pneumonia between single versus combination therapy. *Pharm World Sci* 2005;27:249–53.
18. Ayieko P, English M. Case management of childhood pneumonia in developing countries. *Pediatr Infect Dis J* 2007;26:432–40.
19. Hospital Care for Children. 2013 http://apps.who.int/iris/bitstream/10665/81170/1/9789241548373_eng.pdf
20. Ministry of Health. Basic Paediatric Protocols: November 2013 Edition. 2013 <http://www.idoc-africa.org/index.php/86-clinical-guide/ken-guide/134-basic-paediatric-protocols-november-2013-edition>
21. Koh J, Wong JJ, Sultana R, *et al.* Risk factors for mortality in children with pneumonia admitted to the pediatric intensive care unit. *Pediatr Pulmonol* 2017;52:1076–84.
22. Agweyu A, Kibore M, Digolo L, *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. *Trop Med Int Health* 2014;19:1310–20.
23. Hospital Care for Children. 2005 <http://apps.who.int/iris/bitstream/10665/43206/1/9241546700.pdf>
24. Agweyu A, Gathara D, Oliwa J, *et al.* Oral amoxicillin versus benzyl penicillin for severe pneumonia among Kenyan children: a pragmatic randomized controlled noninferiority trial. *Clin Infect Dis* 2015;60:1216–24.
25. Malla L, Perera-Salazar R, McFadden E, *et al.* Comparative effectiveness of injectable penicillin versus a combination of penicillin and gentamicin in children with pneumonia characterised by indrawing in Kenya: protocol for an observational study. *BMJ Open* 2017;7:e016784.
26. Mwakuyasa S, Wamae A, Wasunna A, *et al.* Implementation of a structured paediatric admission record for district hospitals in Kenya—results of a pilot study. *BMC Int Health Hum Rights* 2006;6:9.
27. Ayieko P, Ntoburi S, Wagai J, *et al.* A multifaceted intervention to implement guidelines and improve admission paediatric care in Kenyan district hospitals: a cluster randomised trial. *PLoS Med* 2011;8:e1001018.
28. Tuti T, Bitok M, Malla L, *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 Glob Health* 2016;1:e000028.
29. Ayieko P, Ogero M, Makone B, *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. *Arch Dis Child* 2016;101:223–9.
30. Danaei G, Rodríguez LA, Cantero OF, *et al.* Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease. *Stat Methods Med Res* 2013;22:70–96.
31. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.
32. Rubin DB. For objective causal inference, design trumps analysis. *Ann Appl Stat* 2008;2:808–40.
33. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
34. Stuart EA. Matching methods for causal inference: a review and a look forward. *Statistical Science* 2010;25:1–21.
35. Austin PC. Assessing balance in measured baseline covariates when using many-to-one matching on the propensity-score. *Pharmacoepidemiol Drug Saf* 2008;17:1218–25.
36. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083–107.
37. Stürmer T, Rothman KJ, Avorn J, *et al.* Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution—a simulation study. *Am J Epidemiol* 2010;172:843–54.
38. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med* 2014;33:2297–340.
39. Baiocchi M, Cheng J, Small DS. Tutorial in Biostatistics: Instrumental Variable Methods for Causal Inference. *Statistics in Medicine* 2014;33:2297–340.
40. Berkley JA, Brent A, Mwangi I, *et al.* Mortality among Kenyan children admitted to a rural district hospital on weekends as compared with weekdays. *Pediatrics* 2004;114:1737–8.
41. English M, Irimu G, Agweyu A, *et al.* Building Learning Health Systems to Accelerate Research and Improve Outcomes of Clinical Care in Low- and Middle-Income Countries. *PLoS Med* 2016;13:e1001991.