Statistical Analysis Plan: CAPS

Trial full title	An advanced cookstove intervention to prevent pneumonia in children under 5 years old in Malawi: a cluster randomised controlled trial
Short Title	Cooking and Pneumonia Study
Acronym	CAPS
SAP Version 1	30/09/2014
SAP Version 2	26/10/2014
SAP Final Version	08/01/2015
Website	www.capstudy.org
Trial Registration	http://www.controlled-trials.com/ISRCTN59448623/
Lancet Summary	http://www.thelancet.com/protocol-reviews/13PRT-4689
ISRCTN:	59448623
Trial Sponsor	Liverpool School of Tropical Medicine (LSTM)
Funders	Wellcome Trust, Medical Research Council, UKaid
Trial Statistician	Professor Brian Faragher
Trial Chief Investigators	Dr. Kevin Mortimer / Professor Stephen Gordon
Co-investigators	Professor Moffat Nyirenda / Professor Nigel Bruce / Professor John Balmes / Professor Jonathan Grigg / Dr Lesong Conteh / Dr Magi Matinga / Dr Dianne Terlouw / Dr Daniel Peter Pope
Trial co-coordinating Centre:	Tropical Clinical Trials Unit (tCTU), LSTM
SAP Author	Professor Brian Faragher/Professor Duolao Wang

Abbreviations

AE	Adverse event
CO	Carbon Monoxide
COMREC	(Malawi) College of Medicine Research Ethics Committee
CRF	Case Report Form
GACC	Global Alliance for Clean Cookstoves
ICC	Intra-cluster correlation
IMCI	Integrated Management of Childhood Illness
IRR	Incidence rate ratio
ITT	Intention-to-treat
LSTM	Liverpool School of Tropical Medicine
MAR	Missing at random
MCAR	Missing completely at random
MoH	Malawi Ministry of Health
PP	Per protocol
QECH	Queen Elizabeth Central Hospital
REC	Research Ethics Committee
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SUMS	University of California Berkley Stove Use Monitors
WHO	World Health Organization

1 Introduction

1.1 Preface

Biomass is burned by 700 million people in Africa to provide energy for cooking, heating and lighting. The smoke this generates causes considerable morbidity and mortality (4 million deaths a year worldwide). Pneumonia in children aged less than 5 years is one of the major diseases associated with biomass smoke exposure and a serious cause of avoidable mortality in less developed countries. There are now efficient biomass-burning cookstoves that substantially reduce smoke emissions and exposures. This trial will evaluate whether provision of an advanced cookstove (Philips fan assisted stove) will reduce pneumonia in young children.

1.2 General purpose of the analyses

The analyses detailed in this document will assess the efficacy and safety of advanced technology cookstoves in comparison with cooking on an open fire with biomass fuel in terms of reducing levels of biomass smoke exposure and reducing the incidence of pneumonia in children aged less than 5 years in Malawi.

2 Study Objectives and Endpoints

2.1 Study Objectives

The overall study objectives are to evaluate the following comparative research hypotheses:

1) H_O: The incidence of pneumonia in children aged less than 5 years in Malawi is the same when using an advanced cookstove relative to an open fire.

   H_A: The incidence of pneumonia in children aged less than 5 years in Malawi is different when using an advanced cookstove relative to an open fire.

2) H_O: Biomass smoke exposure in children aged less than 5 years in Malawi is the same when using an advanced cookstove relative to an open fire.

   H_A: Biomass smoke exposure in children aged less than 5 years in Malawi is different when using an advanced cookstove relative to an open fire.

3) H_O: There is no association between exposure to household air pollution (carbon monoxide) and the development of pneumonia in children aged less than 5 years in rural Malawi.

   H_O: There is an association between exposure to household air pollution (carbon monoxide) and the development of pneumonia in children aged less than 5 years in rural Malawi.

In additional, the study will include an observational survey component to estimate:

1. The prevalence and determinants of obstructive lung disease in adults in rural Malawi
2. The extent to which exposure to household air pollution explains the rate of decline in lung function in adults in rural Malawi.
3. The affordability and cost effectiveness of distributing advanced cookstoves from household, healthcare system and societal perspectives.

2.2 Endpoints

Primary outcome measure:

Incidence of pneumonia episodes in children aged less than 5 years over a 24 months study period.

Pneumonia in children aged less than 5 years will be diagnosed by physicians, medical officers or other appropriately trained staff at local healthcare facilities, blinded to intervention allocation. The WHO IMCI pneumonia assessment protocol will be used to make the diagnosis since chest X-rays are not universally available in the study areas: under this protocol, pneumonia is diagnosed by the presence of cough or difficult breathing and signs of pneumonia - fast breathing (60, 50 or 40 breaths per minute or more in those <2 months, 2-12 months and 1-5 years respectively), chest in-drawing, stridor or any general danger sign (inability to drink or breastfeed, vomiting, convulsions, being lethargic or unconscious).

Secondary efficacy outcome measures:

In children aged less than 5 years in Malawi over a 24 months study period, incidence of each of:

- *Death*

  All deaths will be recorded and an attempt made to distinguish deaths due to pneumonia from other causes; if a child dies at home and it is acceptable to do so, a verbal autopsy will be undertaken.

- *Severe pneumonia episodes*

  Severe IMCI pneumonia will be identified by the presence of any general danger sign, chest wall in-drawing or stridor in a calm child. Oxygen saturation <90% will be included as an additional and objective marker of severity.

- *All pneumonia episodes (including those not meeting IMCI criteria)*

  Additional data supporting the diagnosis of pneumonia (*e.g.* presence of pyrexia, chest X-ray findings) will be collected where this is available. The clinical information recorded in the health passports (see below) will be used to make each diagnosis of pneumonia and assess its severity; this will then be subjected to review by a fully blinded Independent Endpoint Review Committee comprising 3 paediatric specialists not otherwise involved with the trial. A further level of pneumonia diagnosis validation will be possible at the Chikhwawa field site by capitalising on the improved diagnostics (*e.g.* blood cultures) being developed by MLW at this site. Pneumonia diagnoses made within a month of each other will be counted as the same episode but otherwise as separate episodes.

- *Respiratory symptoms and burns*

  Respiratory symptoms and burns will be assessed by active surveillance in the villages every 6 months (*i.e.* at alternate field visits) using respiratory symptoms and burns questionnaires.

To improve the quality of data collection, all children recruited into the study will be issued with a new health passport if they do not currently have one or if their current passport does not have a sufficient number of blank pages. A sticker will be inserted into this passport explaining that the child is in a trial with a brief summary of the IMCI pneumonia assessment protocol and boxes to tick if the child is diagnosed with pneumonia and if so whether this was severe or not. Malaria will be tested for and treated as indicated as part of routine clinical practice and the result of this

recorded in the health passport. During or after the attendance, the trial team will be notified by text or phone call about the event by health facility-based staff or a member of the household using a phone and airtime credit provided to a nominated CAPS village representative. Deaths will be reported in the same way. Fieldworkers will review the health passports of all children in the trial at 3-monthly visits to the villages to obtain information about episodes of pneumonia and deaths not otherwise detected by this system.

In adult members of households in CAPS villages only:

- *Spirometry*

In individual households:

- *Household air pollution, personal exposure and concordance (cookstove use)*

  Household air pollution ($PM_{2.5}$, carbon monoxide) and personal exposure (randomly selected child under age of 4½: black carbon, carbon monoxide, carboxyhaemoglobin; randomly selected adult member of household: as for children plus $PM_{2.5}$, induced sputum alveolar macrophage carbon) will be measured in a random sample of up to 2000 households from each trial site. Up to 48 hours of continuous indoor air quality monitoring will be conducted every 6 months to provide a series of repeated measures from each monitoring episode. Similar methodology will be applied for the personal exposure assessments aspect except that the monitoring devices will be worn on the person and in addition, black carbon exposure, carboxyhaemoglobin and induced sputum alveolar macrophages carbon will be assessed.

  Utilisation of the advanced cookstove will be assessed using University California Berkley Stove Use Monitors (coin-sized heat detecting and recording devices that can be attached to the stoves) in 10% of trial households randomly selected from the intervention trial arm.

Fieldworkers will visit all recruited villages every 3 months for 24 months to collect primary and secondary outcome data. This will be backed up by telephone contact with a village representative every 4 weeks.

3 Study Methods

3.1 General Study Design and Plan

Study design: village-level cluster-randomised controlled intervention open-label trial with two arms of equal size.

Intervention arm: villages in which eligible households will be issued with an advanced cookstove (Philips fan assisted cookstove) with user training for cooking (replacing open fires).

Control arm: villages in which eligible households will continue using traditional cooking methods (open-fires using biomass fuels).

Blinding: because of nature of the intervention, the trial will be open-label for participants, but wherever possible outcome measures will be recorded and/or verified by assessors who are blind to the intervention allocations and the trial statistician will remain blind to the intervention allocations (see below).

Randomisation: within each district, villages that have agreed to participate will be randomly allocated to the intervention and control arms using a computer-generated randomisation schedule; randomisation will be carried out in advance of the intervention (advanced cookstoves) being delivered to the study co-ordinating centres for distribution.

Study population: within each participating village all households with children up to 4½ years old will be invited to participate; after informed consent has been given by a member of the household with authority to do so, the household will be enrolled and given a Household Trial Number and each child in the household will be given a Participant Trial Number linked to the household number.

Assessment times: all recruited households and children will have a baseline assessment at the start of the study (after randomisation but prior to distribution of intervention stoves) with follow-up assessments subsequently at 3 monthly intervals for up to 24 months.

3.2 General Study Population and Inclusion-Exclusion Criteria

General study population: the study will be conducted in a total of 150 suitable village selected across two regions of Malawi: Chikhwawa and Karonga.

Inclusion criteria: all children aged up to 4½ years old resident in each of the selected villages at the start of the study period.

Exclusion criteria: there are no specific exclusion criteria - to maximise generalisability of the findings the trial will be broadly inclusive and open to all consenting households with a child under 4½ (including households where babies are born during the trial) - children known to have HIV (around 5%) will be eligible for inclusion.

Households in villages included in CAPS with at least one adult aged 18 years or older will be eligible for inclusion in the sub-studies involving adult participants.

3.3 Randomisation and Blinding

Within each district, villages that have agreed to participate will be randomly allocated to the intervention and control arms using Excel-generated (pseudo-)random numbers with stratification by site, distance from (or accessibility to) health centre and size of cluster.

The randomisation will be performed by the trial statistician using dummy codes “A” and “B” only to represent intervention and control groups; to ensure the statistician remains blinded, the identity (allocation) of “A” and “B” will be determined by a person independent of the study.

Randomisation will be carried out in advance of the intervention (advanced cookstoves) being delivered to the study co-ordinating centres for distribution.

3.4 Study Variables and assessment times

The following variables will be collected at the assessment times indicated in Table 1. Details of the format of each of the individual variables to be collected are shown in the CRFs (Appendix 1).

Table 1 Timing and nature of study assessments

Measure	Time (months)								
	0 (baseline)	3	6	9	12	15	18	21	24
Consent confirmation	x								
Household inclusion/exclusion criteria	x								
Child inclusion / exclusion and enrolment criteria	x	x	x	x	x	x	x	x	x
Household information	x								
Equipment (cookstove) status and use		x	x	x	x	x	x	x	x
Equipment replacement required		x	x	x	x	x	x	x	x
Fire sources		x	x	x	x	x	x	x	x
SUMS* monitor		x	x	x	x	x	x	x	x
Child symptoms (pneumonia, …)	x	x	x	x	x	x	x	x	x
Child adverse events		x	x	x	x	x	x	x	x
CO** monitoring		x	x	x	x	x	x	x	x
Household / child movement		x	x	x	x	x	x	x	x
Household / child withdrawal (inc. death of child)		x	x	x	x	x	x	x	x
Child health seeking and treatment		x	x	x	x	x	x	x	x

*: University of California Berkley Stove Use Monitors (coin-sized heat detecting and recording devices attached to stoves in 10% of trial households randomly selected from within the intervention trial arm).

**: carbon monoxide.

Assessments carried out:

- within a window of ±28 days will be categorised as "completed at scheduled time";
- outside a window of ±28 days will be categorised as "not completed at scheduled time" but will be allocated to the *intended* assessment time for statistical analysis.

Where appropriate, sensitivity-type analyses will be carried out whereby analyses will be conducted initially using only those assessments completed within the relevant ±28 days window and then repeated using all completed assessments.

3.5 Movement of study participants

A small number of households and/or children are likely to move location during the study period. These will be categorised as "study design violations".

Households or individual children who move to locations *outside* the study area (i.e. to a village that is not a study cluster) will be considered as being lost to follow-up.

Households or individual children who move to locations *inside* the study area (i.e. to a different village which *is* a study cluster) will continue to be followed-up at their new location and will remain part of the intervention group or the control group according to their original randomisation. For logistical reasons, however, all assessment times following a location change will be done at the

times relevant to the new cluster, even though these times are likely to be outside of the ±28 days window for the original cluster.

A separate and detailed guide has been developed to deal with as many possible "study design violation" variations as possible; a copy of this protocol is attached as Appendix 2. Under the "intention-to-treat" principle, the CAPS study is purely observational and naturalistic once a household has been randomised, so the general principle is to do nothing other than to continue to follow-up or monitor households / children as appropriate for the group into which they were originally randomised.

4 Sample Size

The sample size calculations for this study involved a lengthy and iterative discussion process, during which assumptions used in the calculations were amended, as described below.

Original sample size assumptions

The sample size proposed in the outline proposal to JGHT was substantially increased due to the trial design being changed to the current village-level cluster randomised structure, to reflect the effect size seen in the RESPIRE trial of a cookstove intervention on severe pneumonias and to utilise the more contemporary estimate of a health centre IMCI pneumonia diagnosis rate to the 9 per 100 child-years incidence observed in Karonga, Malawi. The full funding application considered the potential impact of the introduction of pneumococcal vaccination on baseline pneumonia rates and took into account informed estimates suggesting around a 20% reduction in pneumonia rates could be expected in the Malawi-specific context. Trials of the 9-valent pneumococcal vaccine in Soweto and The Gambia found a 17% (95% CI 4% to 28%) and 37% (95% CI 27% to 45%) vaccine efficacy on first episodes of radiologically-confirmed pneumonia in children respectively. There was higher vaccine efficacy on pneumococcus-specific disease in both trials but it was considered that the overall impact on all cause pneumonia was most relevant for the CAPS study. The 2009 Cochrane review of pneumococcal conjugate vaccines on invasive pneumococcal disease and X-ray defined pneumonia in young children pooled data from 11 publications and found the vaccine efficacy on WHO X-ray defined pneumonia was 27% (95% CI 15% to 36%) and clinical pneumonia was 6% (95% CI 2% to 9%).

The sample size calculation in this full proposal was based on a 45% reduction in pneumonias, a figure greater than the expected impact of pneumococcal vaccination, to allow for unpredictable factors such as herd protection effects from nationwide vaccination. The conservative assumption was made that only 5% of control group children would develop pneumonia of sufficient severity to require treatment at a health centre every year, and the effect size considered to be of minimum clinical importance was a 20% reduction in pneumonia risk (approximating to the reduction in physician-diagnosed pneumonia seen in RESPIRE (RR 0.84 or 0.78 after multiple imputation) and also smaller than seen in RESPIRE on severe pneumonias (RR 0.67)). A conservative between-cluster coefficient of variation of 0.1 was also adopted. As the intervention in this study was to be an advanced cookstove that reduces smoke emissions and exposures by 80 to 90% while the *plancha* stove used in RESPIRE just vented emissions to the outdoor environment, there was actually a much greater potential for impact with the advanced cookstove than was seen with the *plancha*.

Assuming in addition that villages in the study area contained on average of 85 children, the required sample size was estimated to be 59 villages (clusters) per group each with an average of 77 children (allowing for 10% loss to follow up) followed for an average of 1.7 years (affected by age of child at enrolment). This sample size would have provided 80% power to detect a 20% reduction in the risk of pneumonia in the intervention group from 5% to 4% per annum and 90% power to detect a reduction to 3.8% (using a conventional α=0.05), and would have provided a *potential* total of between 154,462 and 17,051 child years of follow up.

Final sample size assumptions

The final sample size was re-considered in the light of improved data that became available from the planned study sites in Chikhwawa and Karonga and a re-consideration of some of the assumptions included in the original calculations.

The total number of children under the age of 5 years in Chikhwawa was estimated as being 5,027 in 2008 and it was considered reasonable to assume that this number had since increased to approximately 5,600. As there were 50 villages (clusters) in this district, this gave an average number of children per cluster of 112 for Chikhwawa.

The total number of children under the age of 5 years in Karonga was estimated as being 4,750 in 2008 and it was considered reasonable to assume that this number has since increased to approximately 5,000. As there were 278 villages (clusters) of 20-30 households in this district, this gave giving an average number of children per cluster of 18 for Karonga.

This disparity in the average cluster size between the two study districts had implications for the power of the study. As the variation in cluster size increases, the statistical power of a cluster randomised trial reduces.

As the clusters in Karonga are relatively small, one possible option was to combine villages to form larger clusters. This would have been potentially counter-intuitive as cluster randomised trials work best with a large number of small clusters, but this needed to weighed against the negative impact of having a large range of different cluster sizes. To identify the "best" compromise between these two conflicting influences on the statistical power of this study, the sample size calculations were re-worked extensively, with "best" in this context being defined as having sufficient clusters to provide both a feasible design structure (manageable number of clusters to be randomised) and an acceptable level of statistical power.

The compromise recommended was to collapse the existing 278 villages in Karonga to just 100 clusters which, when combined with Chikhwawa, provided a total of 150 clusters (75 clusters per group), with a total number of children across all 150 clusters of approximately 10,600 (average 70.7 children per cluster). Assuming that actual cluster sizes would range between 50 and 150 (a conservative estimate), the coefficient of variation in cluster size was estimated to be in the region of 30 - 35%, requiring the intra-cluster correlation (ICC) value assumed for the sample size calculations be increased by 20%.

The proposed outcome measure for the CAPS study is the incidence of pneumonia cases in children aged less than 5 years recorded in each cluster over the two years of the study period. This measure requires no adjustment for loss to follow-up as the eligible number of children in an individual cluster can reasonably assumed to be constant. For each child who reaches their 5th birthday and hence becomes ineligible for the study, they will be replaced by (at least) one new born child. For the same

reason, the number of child-years of follow-up in each cluster will be the number of children in the cluster at the start of the study period multiplied by two.

A conservative value of 0.1 was again assumed for the ICC (intra-cluster correlation), which was increased by 20% to 0.12 to allow for between-cluster size variation.

It was estimated that:

a total of 150 clusters containing a total of 10,600 eligible children randomised in equal numbers to the two intervention groups would provide, over the whole study period, 21,200 years of follow-up and 90.3% power to detect a reduction in the (annual) risk of pneumonia from 5% in the control group to 4% in the intervention group (proportionally, a 20% reduction in risk), assuming an effective ICC value of 0.10 and a coefficient of variation in cluster size of 30-35%. The same sample size would provide 80.4% power to detect a reduction in the (annual) risk of pneumonia from 5% in the control group to 4.125% in the intervention group (proportionally, a 17.5% reduction in risk), under the same assumptions.

Additional justification for sample size for air pollution exposure and lung function sub-studies

Incidence-exposure analyses (children)

2000 children will be included in the incidence-exposure study and followed up until the end of the CAPS study period, giving an expected total of 2000*2.0= 4000 years of follow-up ("exposure"). It will be assumed that the mean levels of CO in children will be 16.31 (22.77) ppm. The anticipated annual pneumonia incidence rate averaged across the trial arms is 4.5%, which corresponds to an expected incidence rate of 7.53% per child. Assuming a Poisson model, the expected total number of pneumonia episodes is 150 – but as 5 children are predicted by this model to have more than one episode, the expected number of children who will experience at least one pneumonia episode will be 145. On a simple comparison of the 145 children who will experience pneumonia against the 1855 who will not, this study will have 90% power to detect a mean difference of 6.53 (40%) ppm or greater in mean CO levels between these two groups. If it is necessary to match each child who experiences pneumonia with just one child who does not on one or more confounding factors, the minimum detectable difference between the two groups will be 8.92 (55%) ppm.

Respiratory symptoms and lung function analyses (adults)

Baseline spirometry measurements will be recorded for the 2000 adults aged 18 and above recruited into this sub study (replicating the sample size taken for the BHS and BOLD study currently ongoing in the urban setting of Chilomoni ward) along with relevant demographic/clinical characteristics considered to potentially influence the development of chronic respiratory disease. Participants will be stratified into two age groups: 18-39 years and 40 years or above. If 500 males and 500 females (total 1000 individuals) fall into each age group, an estimate of obstructive lung disease prevalence in each gender / age stratum will be obtained with a precision (95% CI) of ±2.6% to ±3.8% (assuming a prevalence of 10% to 25%). Allowing for unequal age and gender distributions, refusals and inability to provide spirometry measurements of acceptable quality, a sample of just 300 participants in any one gender / age stratum will provide an estimate of obstructive lung disease prevalence in this stratum with a precision (95% CI) of ±3.3% to ±5.0% (again assuming a prevalence of 10 to 25%) [this minimal sample size is informed by the BOLD protocol].

The same 2000 adults will be followed with repeated spirometry measurements for two years (the full duration of inclusion in CAPS). Assuming an ICC of up to 0.25 for possible clustering effects within villages, this study will have 90% power to detect a correlation between CO/particulate matter exposure and change in FEV1 level of 0.102 (or greater) in both age groups combined and 0.144 (or greater) in each age group separately.

5 General Considerations

5.1 Timing of Analyses

A single definitive statistical analysis of the primary and secondary outcome measures of efficacy will be performed by the trial statistician when all of the following have been achieved:

- all children recruited into the study have been followed-up for 24 months, or have reached their 5th birthday, or have been deemed to be lost to follow-up;
- all CRFs have been entered onto the computer database at the relevant study co-ordinating centre;
- all data on the two computer databases have been checked for completeness, and the accuracy of all data entries have been verified;
- the databases from the two study centres have been appropriately concatenated;
- the concatenated database has been verified and signed off by the trial statistician;
- the concatenated database has been locked.

There are *no* planned interim analyses for efficacy as such. However, the DSMB has the authority to request the independent medical statistician to perform an unplanned analysis of the overall pneumonia rate should they have concerns about this at any time during the study; as pneumonia episodes are both primary outcome events and SAEs routinely reported to DSMB, any evaluation of pneumonia on safety grounds will also constitute an interim analysis of efficacy.

Provision is made (section 5.5 below) to re-evaluate the sample size calculation assumptions should this be considered appropriate, which may also require the DSMB independent medical statistician to calculate an interim blinded estimate of effect size.

A final definitive analysis of all safety measures will also be performed by the trial statistician when the above criteria are met. In addition, a single planned interim for safety will be performed by the independent DSMB statistician at the half-way point in the study (defined as being when 50% of recruited children have completed one-year of follow-up).

5.2 Analysis Populations

5.2.1 Intention-to-treat (ITT) Population

The primary statistical analysis for efficacy will be performed using the *intention-to-treat* principle. The population for this analysis will be all children for whom a confirmed and validated consent to participate in the study was obtained and who underwent at least one post-baseline follow-up assessment. Study group membership for this analysis will be determined solely by the randomised allocation of the cluster (village) where the child was resident at the time of recruitment/consent.

5.2.2 Per Protocol (PP) Population

The secondary statistical analysis will be performed using the *per protocol* principle. The population for this analysis will be all children for whom a confirmed and validated consent to participate in the study was obtained and who underwent at least one post-baseline follow-up assessment covering a period during which they continuously exposed to the intervention (advanced cookstove or traditional cooking methods) randomly allocated to the household in which the child was resident at the time of recruitment/consent. Children who move to a house where either this intervention is no longer being used or where neither intervention is being used will be included in the per protocol analysis only up to the point in the follow-up period for which the child was exposed to their randomised intervention.

5.2.3 Safety Population

Statistical analyses for safety will be performed using the population of all children for whom a confirmed and validated consent to participate in the study was obtained; children included in this population may not have undergone at least one formal post-baseline follow-up assessment if a safety-related event occurs prior to this being scheduled.

5.2.4 Allocation to populations

The inclusion / exclusion status for each participating child will be determined for all three of the above analysis populations by the trial statistician while blind to intervention group allocations. The exact process used to assign each population status will be documented for each child prior to the commencement of any analyses and any reasons for eliminating any children from a particular population will be fully stated.

5.3 Covariates and Subgroups

The study is not (statistically) powered for any formal sub-group analyses involving the primary and secondary outcome measures.

Important differences may be present between the two study sites, which differ considerably geographically. Adjustment for site differences will be made in the statistical analyses by including a dummy covariate for site (1=Chikhwawa, 2=Karonga) where appropriate. Sub-group analyses will be carried out for each study site separately, but as this cannot be done sensibly by incorporating appropriate interaction terms in the statistical models, all sub-group analyses will be considered to be exploratory.

Effect size estimates for the primary and secondary outcome measures will be routinely adjusted for clustering effects, by using a dummy covariate containing unique codes for each cluster included in the analyses as a clustering variable.

Effect size estimates will be estimated in two ways, in the following order (see section 7 below):

- unadjusted for any other factors (with the single exception of clustering effects)
- adjusted for the pre-specified covariates (irrespective of statistical significance)
-

Covariates that are expected to possibly influence outcome include:

- region (Chikhwawa / Karonga)
- age of child at recruitment (age<2 years and age>= 2 years)
- sex of child
- distance to nearest health centre (<median, ≥median)
- number of children aged less than 5 years living in household
- number of people in household who smoke regularly (Yes/No)
    - other sources of fire or smoke (other than cooking) to which child exposed on daily (or almost daily) basis (Yes/No to Burning rubbish)
- socioeconomic status of household (tertile)
- number of previous pneumonia episodes experienced by child (including times of occurrence) (Once or more)
- status of childhood vaccinations. (Yes/No)

Exploratory analyses of any observed reductions in household air pollution, personal exposure and stove will be performed using subgroups for which there is a sufficient sample sizes for a meaningful analysis.

5.4 Missing Data

Pneumonia episodes (those meeting IMCI criteria; severe; those not meeting IMCI criteria)

These outcome measures will be analysed using the number of known episodes recorded for each participating child and the time period for which follow-up information is available for that child. In essence, therefore, these measures will be analysed as "episodes per unit time" (for more detail, see section 7 below). Thus, no child will have missing data for these outcome measures irrespective of whether they complete the full 24 month follow-up period; children who do not have even a single post-baseline follow-up assessment will be excluded from all analyses.

Death

This outcome measure will be analysed as "time to death (measured from date of start of intervention)". Time will be recorded either as actual time to when death occurred or time for which child known to be alive. Again, no child will have missing data for these outcome measures irrespective of whether they complete the full 24 month follow-up period.

All other measures

All other measures will be reported using appropriate summary statistics (with 95% confidence intervals) using just those individuals for whom values were recorded; numbers of missing observations will be reported alongside these statistics.

Analyses of trend over time (i.e. of changes between the scheduled assessment times) will include time of each assessment (measured from the date of the start of the intervention) as a continuous variable with appropriate partitioning of the within and between-subject variances (i.e. using panel data regression methods). No attempt will be made to either insert missing observation estimates or to perform multiple imputation methods if any assessments were not completed for any of the study participants. However, as missing assessments are more likely to occur later in the study period than at the beginning (e.g. due to withdrawals and/or losses to follow-up), "last value carried forward" analyses may be performed as a form of sensitivity analysis to provide an informal estimate of the possible biases in the main analyses.

For measures used to adjust effect size estimates, missing observations will be handled using multiple imputation methods, providing it is considered valid to assume that these observations are either missing at random (MAR) or missing completely at random (MCAR). The variables included in the imputation process will be confined to just those that will be included as covariates in the relevant analyses.

5.5 Interim Analyses and Data Monitoring

5.5.1 Purpose of Interim Analyses

A single blinded interim analysis will be carried out to determine whether there are grounds to stop the trial for safety and/or to establish whether the incidence of pneumonia episodes being observed is compatible with the assumptions made in the sample size calculations (section 4 above).

There are *no* planned interim analyses for efficacy as such. However, the DSMB has the authority to request the independent medical statistician to perform an unplanned analysis of the overall pneumonia rate should they have concerns about this at any time during the study; as pneumonia episodes are both primary outcome events and SAEs routinely reported to DSMB, any evaluation of pneumonia on safety grounds will also constitute an interim analysis of efficacy.

Provision is made (section 5.5 below) to re-evaluate the sample size calculation assumptions should this be considered appropriate, which may also require the DSMB independent medical statistician to calculate an interim blinded estimate of effect size.

5.5.2 Planned Schedule of Interim Analyses

The single planned interim analysis will be carried out by the independent statistician on the DMC at the half-way point in the study (defined as being when 50% of recruited children have completed one-year of follow-up).

5.5.3 Stopping Rules

The Peto-Haybittle rule will be used to inform a discussion about whether there are grounds for terminating the study prematurely for safety reasons, with statistical significance set at the ($p<0.001$) level in both instances.

If there is concern that the observed overall event rate is very different from that anticipated, a blinded revision of the total sample size estimate will be done using the methods advocated by Gould; this simple formula is based on the predicted and current (*i.e.* at time of the interim analysis) proportion of participants overall who have experienced a pneumonia event, and has a considerable advantage in that it preserves the power of the study and does not affect the type I error rate materially. If subsequently there is concern that the observed overall event rate is much higher than anticipated and hence there might be concerns over safety, a blinded comparison of the event rates in the two study groups will be carried out using the methods advocated by Wassmer et al and by Posch and Bauer.

5.5.4 Practical Measures to Minimise Bias

All interim analyses will be carried out by the independent statistician on the DSMB. The results of these analyses will be reported only to the Chair of the DSMB, who will determine which other members of the DSMB should also receive a copy of the analyses. The Chair of the DSMB will determine which members of the DSMB require to see the safety interim analysis and/or the re-evaluation of the sample size calculations.

No copies of the interim analyses will be made available to individuals involved with the running of the trial until either the study is completed or a decision has been taken by the DSMB to either terminate the trial or to alter the total study follow-up period / sample size requirement.

Where possible and/or appropriate, interim analysis results will be provided to the DSMB aggregated across both study groups. If the DSMB requests interim analysis summary statistics for each of the two study groups separately, these will be provided initially in a blinded fashion (i.e. with the two groups labelled merely as “A” and “B”). If the DSMB then requests that these summaries to be unblinded, the identities of “A” and “B” will be provided to the Chairman by the independent trial statistician.

The trial statistician will provide the independent statistician with a blinded copy of the trial database for the interim analysis but will not otherwise be involved in this analysis or any discussions about the implications of the findings of the interim analysis. The trial statistician will carry out the full analysis of the completed/terminated trial database, but will remain blinded until all analyses of outcome have been completed. Blinded analysis findings will be reported to the DSMB and to the PI; the Chair of the DSMB will determine the point at which it will be appropriate for the PI to become unblinded to the study findings.

5.5.5 Documentation of Interim Analyses

A copy of the database used to carry out the single interim analysis will be locked and stored securely on the central file-server at LSTM.

6 Summary of Study Data

All continuous variables will be summarised using the following descriptive statistics: (non-missing) sample size (n), mean or median, standard deviation or range (maximum and minimum).

All categorical variables will be summarised using the frequency and percentage (based on the non-missing sample size) for each observed category.

Where relevant, individual participant data will be listed sorted by study group (intervention / control), cluster (village) and, if appropriate, by assessment time.

All summary tables will be structured with an initial column showing the appropriate summary statistics for all participants combined, followed by separate columns for the control (conventional cooking methods) and intervention (advanced cookstove) groups separately (in that order); sample sizes and/or numbers of missing observations will also be reported. A final column will show, where appropriate, estimates of effect size with their 95% confidence intervals.

Separate Tables will be provided for the ITT and PP analyses.

6.1 Subject Disposition

A CONSORT diagram will be provided indicating, by study group, the total number of clusters (villages), the total number of children in these clusters, the number for whom consent was given to participate in the study, the number who were lost to follow-up (with reasons: consent withdrawn, moved out of study area, died), the number who reached their 5th birthday during the study period, and finally the number who completed the full 24 months of the follow-up period.

6.2 Protocol Deviations

See section 3.5 and Appendix 2.

6.3 Demographic and Baseline Variables

The demographic and baseline variables are fully documented in Appendix 1.

Summary statistics for these variables will be produced as described above in section 6; no effect size or group difference statistics will be produced for these variables.

6.4 Treatment Compliance

The transition from cooking over an open fire to using an advanced cookstove represents a large change in an activity that usually takes up a considerable part of the day, can be part of the social fabric of the village and is sometimes associated with particular beliefs and superstitions. However, high levels of cookstove adoption were found in an exploratory RCT in Ntcheu and in an acceptability study in Lesotho. With careful community engagement, support from community leaders and training in the use of the stove, initial innovation adoption is therefore likely to be successful. It is expected that the advantages of the advanced cookstove in terms of reduced time needed for cooking, fuel consumption, smoke emissions and improved safety will help maintain high levels of use. The RESPIRE study helped to sustain high levels of compliance by providing a maintenance and repair service for the Plancha stove, a practice that will be adopted in this study. We will also assess compliance with the intervention through self-reporting and by Stove Use Monitors in 10% of intervention households. Compliance with the protocol and SOPs by field staff will be maximised through training events to include GCP training, and periodic quality control audits. The availability and condition of each cookstove distributed will be assessed at each scheduled assessment visit.

Efforts will be made to minimise loss to follow-up through active community engagement, responding promptly to trial-related difficulties, repairing and replacing stoves as needed and providing additional benefits to the participating villages (e.g. mobile phone access).

Each household will be allowed one new cookstove during the study period should the cookstove allocated at the start of the intervention period become irreparably damaged or lost.

The numbers of lost and irreparably damaged cookstoves will be summarised, broken down by assessment month.

7 Efficacy Analyses

7.1 Primary Efficacy Analysis

The primary measure of efficacy (section 2.2) is the *incidence of pneumonia episodes in children aged less than 5 years over a 24 months study period*. Stated alternatively, this is the number of episodes of pneumonia experienced by each recruited child during the period of time for which they were followed; this period will be 24 months for most children, but will be less for those who reach their 5^{th} birthday during follow-up or who are lost to follow-up, so incidence for each child will be adjusted appropriately for actual length of follow-up.

Generalised estimating equation (GEE) model will be used to evaluate the primary response variable. Firstly, the number of pneumonia events experienced by each child will be analysed assuming a Poisson distribution, with length of follow-up as the exposure variable and with adjustment for clustering effects between households within the same village. This model will generate a crude incidence rate ratio (IRR) with its 95%CI. Secondly, the GEE model with treatment and pre-specified covariates as in Section 5.3 will be used to derived the adjusted IRR with its 95%CI. In the GEE model analyses, Exchangeable covariance structure will be used. The main conclusion regarding the primary endpoint will be drawn from the unadjusted analysis based on the ITT population.

In addition, two additional analyses will be considered:

- time from randomisation to first pneumonia episode post-intervention will be analysed using Cox regression model with frailty for the cluster effect;
- time from a prior episode to each pneumonia episode post-intervention (multiple episodes of pneumonia) will be analysed using GEE model with negative binomial distribution.

Both unadjusted and adjusted treatment effects will be derived from the above analysed. The adjusted analyses will be based on the pre-specified covariates detailed in Section 5.3.

7.2 Secondary and Exploratory Efficacy Analyses

Mortality

If sufficient deaths occur during the study, Cox regression model with frailty for the cluster effect will be used to compare the times to death between the two study groups using Cox regression methods. Unadjusted and adjusted hazard rate ratios with their 95% confidence intervals will be derived from the Cox model. As children will not all be followed-up for the same length of time (or even for a uniform minimum time), an analysis of actual death rates will not be possible.

Incidence-exposure analyses (children)

Initially, the mean CO levels of those children who did and who did not experience any pneumonia episodes will be compared using linear regression models. The association between personal exposure to CO and actual number of pneumonia episodes will then be assessed using Poisson

regression analyses with length of time of follow-up as an exposure variable; exposure response curves will be constructed for personal exposure to CO against pneumonia episodes. Finally, if sufficient pneumonia episodes are observed, GEE model with negative binomial distribution will be used to evaluate time between episodes. All analyses will include appropriate adjustments for clustering within villages; terms will be included in each model for treatment arm and important confounders and covariates considered *a priori* to strongly influence outcome (section 5.3)

Respiratory symptoms and lung function analyses (adults)

(a) The BOLD data will be analysed in accordance with the BOLD protocol (www.boldstudy.org). Response rates, the characteristics of the study participants, and COPD prevalence estimates will be reported with 95% CIs. Poisson binomial regression models will be used to explore factors associated with COPD prevalence, with appropriate adjustment for clustering within villages.

(b) Longitudinal data (such as spirometry) will be analysed using the using the GEE models with treatment, visit, interaction between as fixed effect, with and without adjustment for potential confounders including age, gender, location, socioeconomic status, as covariates, and village as cluster effect. The treatment difference at different visits together with their 95%Cis will derived from the GEE models. Exchangeable covariance structure will be used in the GEE model analyses.

All statistical analyses will be performed using the SAS 9.3 statistical software packages.

8 Safety Analyses

The advanced cookstove intervention being using in this trial is a non-medical intervention and is not known to increase the risk of any adverse event. It is a particularly low-risk intervention that offers potential safety benefits (e.g. reduced risk of burns and fires). Nevertheless, data will be collected about adverse events.

For this study:

- an adverse event (AE) is defined as being any unfavorable and unintended sign, symptom, syndrome or illness that develops of worsens during the period of observation in the study
- a serious adverse event (SAE) is defined as an adverse event that results in a) death, b) a life-threatening adverse event, c) hospitalization or prolongation of an existing hospitalization, d) disability or incapacity, e) congenital anomaly in the offspring of a participant.

Data about AEs that are not serious will be collected at the routine three monthly field visits. Study participants will be asked to report SAEs immediately to the trial coordinating centre using a pro-forma in accordance with a specific SOP. This information will then be passed immediately to Dr. Kevin Mortimer and Professor Stephen Gordon who will conduct a causality assessment (not related/improbable, possible, probable, definite), assess seriousness and expectedness, take any appropriate medical action and inform COMREC and the LSTM REC of any events deemed related to the trial intervention within 7 days of knowledge of the event. All other SAEs will be reported as part of an annual report to COMREC and LSTM REC. All SAEs will be followed to resolution.

Frequencies of SAEs and AEs will be tabulated both overall and for the two study groups separately. Where appropriate, frequencies will be compared between the groups using Fisher exact tests, with the exception of death which will be analysed as a secondary efficacy variable (section 7.2).

9 Reporting Conventions

Statistical significance will be set at the conventional 5% level ($\alpha = 0.05$) for all analyses.

Effect sizes will be presented as mean group differences, group incidence rate ratios or group incidence rate differences, group hazard ratios as appropriate, with their 95% confidence intervals, but p-values will also be reported for completeness.

All p-values will be reported to 4 decimal places; p-values less than 0.0001 will be reported as "<0.0001".

The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 2 decimal places.

Appendix 1: CRF

Appendix 2: Protocol

Appendix 3: The list of tables and listings to be produced

Section 1	**Baseline information**
Table 1.1	Household information at baseline
Table 1.2	Children's characteristics at baseline
Section 2	**Household Follow Up**
Table 2.1	Consent
Table 2.2	Equipment Status
Table 2.3	Fire Sources
Table 2.4	Information copied from ANC card
Section 3	**Child Follow Up**
Table 3.1	Pneumonia
Table 3.2	Follow-up on child symptoms and adverse events
Table 3.3	Child burn
Section 4	**Adverse events**
Table 4.1	Summary of adverse events
Table 4.2	Raw listings of adverse events
Section 5	**Efficacy analysis**
Table 5.1	Primary outcome
Table 5.2	Secondary outcomes
Table 5.3	Covariate adjusted analysis
Table 5.4	Subgroup analysis
Section 6	**Household and child withdrawal**
Table 6.1	Household withdrawal
Table 6.2	Child withdrawal