Project Title: The PLEASANT Study

(Preventing and Lessening Exacerbations of Asthma in School-age children Associated with a New Term)

A cluster randomised controlled trial investigating the effect of a postal intervention in reducing unscheduled medical contacts in school age children following returning to school.

Statistical Analysis Plan

Version 1.1.0

Authored by

_____ ___/___/_____

Benjamin Sully Date

PLEASANT trial statistician

Medical Statistics Group, DTS, University of Sheffield

Approved by

_____ ___/___/_____

Neil Shephard Date

PLEASANT trial statistician

CTRU, DTS, University of Sheffield

_____ ___/___/_____

Mike Bradburn Date

PLEASANT trial Senior statistician

CTRU, DTS, University of Sheffield

_____ ___/___/_____

Professor Steven A. Julious Date

PLEASANT chief investigator

Medical Statistics Group, DTS, University of Sheffield

_____ ___/___/_____

Martyn Lewis Date

Reader in Biostatistics

PLEASANT Trial Steering Committee Statistician

Primary Care Sciences, Keele University

List of abbreviations used

AE	Adverse Event
AFT	Accelerated Failure Time
CI	Confidence Interval
CPRD	Clinical Practice Research Datalink
CTRU	Clinical Trials Research Unit
GP	General Practitioner
HTA	Health Technology Assessment
ICC	Intra-Class Correlation
ITT	Intent-To-Treat
MPR	Medicine Possessions Ratio
NHS	National Health Service
SOP	Standard Operating Procedure
TMG	Trial Management Group
TSC	Trial Steering Committee
UoS	University of Sheffield

1.1.1 Introduction, study design and key trial objectives

1.1.1.1 Study outline

The PLEASANT study is a parallel group, cluster randomised controlled trial that will compare a postal intervention to standard care in children aged 4-16 with previous diagnoses of asthma; 70 General Practices (GPs) will be randomised to each arm, and patients from these GPs will receive the appropriate intervention.

This statistical analysis plan is written in conjunction with the International Conference on Harmonisation topic E9 (Statistical principles for clinical trials), applicable standard operating procedures from the Sheffield Clinical Trials Research Unit (CTRU) and trial documents referenced in section 4.

This trial is funded by the National Health Service (NHS) Health Technology Assessment (HTA).

1.1.2 Outcome measures

1.1.2.1 Primary outcome measure

- The proportion of patients aged between 5-16 who have an unscheduled medical contact in September

-

1.1.2.2 Secondary outcome measures

- The proportion of patients who have an unscheduled medical contact in the period September – December
- The total number of medical contacts (scheduled and unscheduled) per patient in September and in the period September – December
- The time to first unscheduled medical contact in September and in the period September – December
- The proportion of patients who have a medical contact (either scheduled or unscheduled) in September and in the period September – December
- The total number of medical contacts (scheduled and unscheduled) per patient in September and in the period September – December
- The time to first medical contact in September and in the period September – December
- The proportion of patients who have an unscheduled medical contact in September and in the period September – December associated with a respiratory diagnosis
- The number of unscheduled medical contacts per patient in September and in the period September – December associated with a respiratory diagnosis
- The time to first unscheduled medical contact associated with a respiratory diagnosis in September and in the period September – December
- The number of prescriptions per patient in the month of August
- The number of prescriptions in the 12 months following the intervention
- The proportion of patients who have a scheduled medical contact (for example asthma review) in August
- The proportion of patients who have a scheduled medical contact (for example Asthma review) in the 12 months following the intervention.

All above analyses will be undertaken and reported twice: once on patients under the age of 5 and once on patients aged between 5 and 16. This is because asthma is difficult to diagnose in children below this age[1-2].

1.1.2.3 Sample size

From previous research in the CPRD practice population 30% of school age asthmatic children had at least one unscheduled medical contact within the month of September13. We postulate that the intervention may reduce the number of children who have unscheduled medical contacts from 30% to 25% (i.e. an absolute reduction of 5%). We would have an effect size of 5%. The average practice size in the CPRD is 8,294. We thus anticipate circa 100 school age asthmatic patients per practice (based on 12% of a practice being school age children and 11% of school age children having asthma). Hence, to detect a difference of 5% with 90% power and two sided significance level of 5%, with an intra-class correlation (ICC) of 0.03 to account for clustering we require 70 practices per arm. The sample size of 140 practices would equate to approximately 14,000 school age asthmatic patients.

Ukoumunne et al[3] give estimates of ICCs for patients with respiratory symptoms in General Practice. Based on the work of Ukoumunne et al an ICC of 0.03 is a conservative estimate. The power of the study for ICCs of 0.01, 0.02, 0.03, 0.04 and 0.05 is respectively 99.4, 96.0, 90.0, 83.1 and 76.2%

As a further sensitivity analysis we investigated the effect of practices not sending out the letter as planned. Suppose 10 practices failed to send out the letter, these would still be included in the primary analysis under the intent to treat principle. However, the effect that could be observed would be reduced to 4.3%. Under the sample assumptions (ICC=0.03 etc) the power for the same sample size is reduced to 79.3%. This is a little under 80% but it does demonstrate that the study is reasonably robust to at least one deviation in the planned design.

1.1.2.4 Randomisation

The study is a cluster randomised trial; 70 general practices (GPs) undertaking the intervention and 70 control practices of "usual care". The randomisation will be stratified by size of GP to ensure that there is an equal sample size – in terms of number of school age asthmatic children – in each arm of the trial. Practices will be randomised to one of the two arms after they have agreed to participate. The randomisation will be carried out by the University of Sheffield (UoS) Clinical Trials Research Unit (CTRU) using a randomisation plan developed prior to the beginning of the trial.

1.1.2.5 Interim analyses and study committees

Two committees will be established to govern the conduct of the study:

1. Trial Management Group (TMG)
2. Trial Steering Committee (TSC)

All committees are governed by Sheffield CTRU standard operating procedures. The TMG consists of the Principal Investigator, co-investigators and key staff within the CTRU. The role of the TMG is to implement all parts of the trial.

The TSC consists of the Principal Investigator, key staff within the CTRU (as non-voting members), an independent chair and two independent members (including a statistician) and 2 lay members. The roles of the TSC are to provide supervision of the protocol and statistical analysis plan, provide advice on and monitor progress of the trial.

No formal interim analyses are required in the study.

1.1.3 Data sources, data evaluability and analysis populations

1.1.3.1 Data sources

The data for this study will be collected and managed by the Clinical Practice Research Datalink (CPRD), a computerised database of anonymised longitudinal medical records from primary care. The CPRD are able to capture all medical contacts along with the reason for the contact.

The PLEASANT study team at CTRU will request and collect the appropriate data from CPRD at three time points:

1. Baseline
2. 1 month post intervention
3. 12 months post intervention

The data requested from CPRD will include, for each patient:

- Age
- Gender
- Anonymised General Practice identifier
- The date of each appointment
- The type of medical contact for each appointment

- The diagnosis given for each appointment
- Any prescriptions given as a result of an appointment.

1.1.4 Data evaluability

Upon receiving the data from CPRD, CTRU will handle and prepare the data for statistical analysis. This includes forwarding data pertaining to the nature of each appointment to an adjudication panel for their review, who will in turn define appointments as being either scheduled or unscheduled. The CTRU will also merge treatment allocation data with CPRD data and calculate the number of appointments for each patient.

Detailed data management and data quality issues will be set out in a data management plan. Data will be retained in accordance with the Data Protection Act 1998 and CTRU data management Standard Operating Procedures (SOPs).

All source documents and data will be retained for a period of at least 5 years following the end of the trial.

1.1.5 Analysis populations

The analysis populations will be as follows:

Intent To Treat (ITT)	All randomised patients identified through the extraction identified by the CPRD.
Per protocol (PP)	The subset of the ITT who belong to a practice which complies to the protocol, meet the inclusion/exclusion criteria and whom the GPs did not exclude from receiving the intervention.

All analyses will be performed on both study populations.

There are three study periods to be analysed. The primary analysis will be undertaken on the primary study period; secondary analyses will use all three stages.

Primary study period	1st September 2013 – 30th September 2013
Extended study period	1st September 2013 – 31st December 2013
Follow-up period	1st September 2013 – 31st August 2014

1.1.6 Outline of analyses

1.1.6.1 General considerations

Summaries of continuous variables will comprise the sample size used and either:

i. mean, standard deviation, minimum and maximum, or

ii. median, inter-quartile range, minimum and maximum

as appropriate for the distributional form of the data. Summaries of categorical variables will comprise the sample size used, and the number and percentage of observations in each category.

1.1.6.2 Levels of statistical significance and adjustment for multiplicity

The PLEASANT study was designed and planned using a 2-sided significance level of 5%. All analyses will be undertaken using this level of significance. As there is only one primary outcome and no interim analysis, adjustment for multiplicity is unnecessary. However adjustments will be made for the multitude of secondary outcomes. Conservative Bonferroni corrections will be made to the raw P-values and where possible k-fold cross-validation will be performed by using a leave-one-out approach.

1.1.6.3 Rules for derived variables

The number of appointments for each patient will be calculated after the panel has determined whether appointments were scheduled or unscheduled. The numbers of each will then be summed (for both the primary and extended study periods). There are instances where no medical code has been used to record the type of medical contact and instead free-form text has been entered. Such entries will always be unscheduled (because scheduled contacts are recorded so that GPs are remunerated) but it is impossible to determine the nature of the contact and therefore whether it is respiratory related or not. The number of each contact type, in terms of "relevant"/"irrelevant", "scheduled"/"unscheduled"/"unknown" and "respiratory related"/"not respiratory related"/"indeterminable" will be reported.

The proportion of patients with unscheduled medical contacts in September 2013 will be analysed using a derived variable. Any patients who have had one or more unscheduled medical contacts in this period will be coded as '1', while those who have had zero unscheduled medical contacts in this period will be coded as '0'. This binary variable will then be used as the dependent variable in the analysis. This will be done for all outcomes involving a proportion of patients.

1.1.7 Disposition

The following summary will be presented for all practices and patients:

- Centre disposition: the number and percentage of practices included in each analysis population with reason for exclusion

- Patient disposition: the number and percentage of patients included in each analysis population with reason for exclusion

The following summary will be presented for the ITT:

- Data completeness: the number of patients with complete data for key parameters by treatment group

- Data completeness by practice: the number of patients with complete data for key parameters by practice.

1.1.8 Demographics and baseline characteristics

The following summaries will be presented:

- Demographics: age; gender; practice; number of asthma admissions in September 2012, the period 1st September – 31st December 2012 and the period 1st September 2012 – 31st August 2013 (scheduled, unscheduled and both combined); time to first medical contact in September and the period 1st September – 31st December 2012 (scheduled, unscheduled and both combined).

1.1.9 Efficacy

1.1.9.1 Primary outcome

The primary outcome will be analysed by intent to treat among patients aged 5-16 as of 1st September 2012. The primary endpoint (the proportion of patients who have an unscheduled medical contact in September) will be analysed by logistic regression in which the fixed covariates will include the individual's age, gender, number of contacts the previous September, and trial arm; GP will be included as a random effect to account for the effect of clustering by practice.

The following outputs will be presented for the ITT and PP:

- The number of unscheduled medical contacts in September 2013
- The proportion of patients having unscheduled medical contacts in September 2013
- The results of the logistic regression modelling for the primary outcome, summarising the effect of all covariates fitted in the model.

1.1.10 Secondary outcomes

1.1.10.1Proportion of patients with medical contacts

For analysis of secondary outcomes involving proportions of patients in both the extended period of September-December 2013, September2-13-August 2014 and September 2014 , the same approach will be used as for the primary outcome. Similar covariates will be included in the analysis, ensuring that the baseline variable matches the outcome variable. For example, when analysing the proportion of patients who have an unscheduled medical contact in the period September – December 2013 associated with a respiratory diagnosis, the baseline covariate will be the number of contacts in the previous September – December 2012 associated with a respiratory diagnosis.

1.1.10.2Number of patients with medical contacts

For outcomes involving numbers of medical appointments or prescriptions the intervention will be analysed in an analogous approach to those involving proportions. A random effects negative binomial model will be fitted, including the same covariates as above.

1.1.10.3 Time to first medical contact

Analyses involving the time to first medical contact will all be analysed using a random effects ("shared frailty") regression model including the same covariates as described previously. Due to the expected high prevalence of ties (i.e. the same time to first contact) the Efron method for handling ties will be used.

1.1.11 Number of Prescriptions

The number of prescriptions per patient in August 2013 and in the 12 months following the intervention will also be summarised and analysed under a negative binomial random effects regression model.

1.1.11.1 Scheduled contacts

The proportion of patients who have a scheduled medical contact (e.g. asthma review) in August 2013 and in the 12 months following the intervention will be analysed using a logistic random effects regression model.

1.1.12 Testing assumptions of statistical analyses

The primary outcome will be analysed using a random effects logistic regression model. This modelling technique is very robust and makes very few assumptions. The same applies for the secondary analyses involving proportional dependent variables. The Hosmer-Lemeshow test will be used to test the goodness of fit for these models.

The secondary analyses involving number of events will be analysed using random effects negative binomial regression. Similarly to above, this method is very flexible and does not rely on assumptions.

Analyses involving time-to-event data will be analysed using random effects "shared frailty" Cox regression. The key assumption underlying this analysis method is that the hazard in one group (or one level for a continuous covariate) is a constant multiple of that in another group (level). This will be tested by fitting an interaction term between time and treatment arm: if the hazard ratio is constant, this term will be non-significant. If the hazard ratio is found to be non-constant over time the outcome will instead be analysed using Accelerated Failure Time (AFT), with goodness of fit assessed by Q-Q plots[4]. If the assumptions underlying this method are not met, residual mean survival methods will be used[5].

1.1.13 Compliance

Compliance will be based on whether or not practices comply with the intervention i.e. whether they send out the letter. To check for differences between complying and non-complying practices the demographics for each population will be. Tables displaying outcome data will also be reported split by compliance.

1.1.14 Economic analyses

Economic analyses will be included in a separate document.

1.1.15 Analysis of non-adherence

In order to identify patients who are non-adherent to regular asthma treatments the medicine possessions ratio (MPR) for each participant will be calculated as the following:

$$MPR = 100 \times \frac{\text{Number of days of medicine prescribed in last 12 months}}{365}$$

This will be calculated at baseline (the year prior to the intervention) and at follow-up (the year following the intervention). Patients with an MPR of under 80% will be classed as 'non-adherent' to medicine.

The MPR will be calculated for preventative medications only, using prescription information. The analysis will be undertaken only on patients who have a single medication which remains the same over both baseline and follow-up; patients prescribed more than one preventative medication or who switch medications between periods will be excluded from this analysis.

Informal analysis will take place to ensure that the MPRs are independent across treatment arms at baseline and also independent across time points in the control group. This will comprise histograms of the MPR and summary statistics.

The main analysis of MPR will test whether the intervention changes MPR. This will be done in two ways:

1. A test for change in proportion of patients classed as non-adherent before and after the intervention in control and intervention arms.

2. Testing the difference in change in MPR before and after the intervention between the control and intervention arms.

A separate subgroup analysis will investigate whether patients who are classed as non-adherent at baseline respond differently to the intervention to those who are classed as adherent.

1. Paired t-test, intervention group only, comparing difference between baseline and follow-up for adherent vs non-adherent.

1.1.16 References

1.1.16.1 Trial Documents

Trial Protocol (version 1.6, 25/09/2012)

1.1.16.2 Other References

1. British Thoracic Society / Scottish Intercollegiate Guideline Network. British Guideline on the Management of Asthma: A national clinical guideline. Jan 2012.

2. Bush A. Diagnosis of asthma in children under five. Primary Care Respiratory Journal. 2007;16(1):7-15

3. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC and Burney PJG. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. Health Technology Assessment 1999; Vol. 3: No. 5.

4. Bradburn MJ, Clark TG, Love SB and DG Altman DG. Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. British Journal of Cancer (2003) 89, 605–611.

5. Royston P and Parmar MKB. The use of restricted mean survival time to analyse randomized clinical trials data when the proportional hazards assumption is in doubt. Statistics in Medicine 2011; 30: 2409-2421.