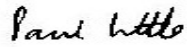
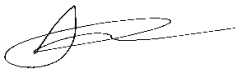
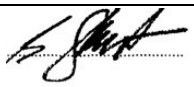
Insert trial image logo here.

Immune Defence Trial

Statistical Analysis Plan

Trial name:	Immune Defence
Trial registration number:	ISRCTN17936080
Protocol title and version number:	Immune Defence protocol v8.1 15-11-22
SAP version number:	Version 2
SAP date:	19/10/2023

To be approved and reviewed by:

	Name	**Signature**	**Date**
Chief Investigator(s)	Paul Little		24/10/23
	Adam Geraghty		25/10/23
Trial Statistician (author)	Taeko Becque		19/10/2023
Senior Clinical Trial Statistician	Beth Stuart		25/10/23

FUNDED BY
NIHR | National Institute for Health and Care Research

Table of Contents

List of Abbreviations

Abbreviation		Abbreviation	

1 Introduction

1.1 Purpose of Statistical Analysis Plan (SAP)

This statistical analysis plan (SAP) describes in detail the methods that will be used to analyse the data collected as part of the Immune Defence trial. This will form the basis of the final trial publication. The final analysis will follow the SAP to ensure that the analyses are conducted in a scientifically valid manner and to avoid post hoc decisions which may affect the interpretation of the statistical analysis. Any deviations from the SAP will be detailed in the final report.

1.2 Trial background and rationale (short synopsis)

A range of viruses circulate each winter and cause respiratory infections (RTIs) (the viruses that cause colds, sore throats, sinus, chest or ear infections, flu). These can lead to people being off work, seeking help from the NHS, and being admitted to hospital in the winter months. The combined effect of both the normal winter viruses (and also the COVID-19 virus in the current pandemic) are likely to cause a major problem for the NHS not only during the coming 2020-21 winter season but in subsequent years. There is promising evidence that using nasal sprays, or alternatively reducing stress and increasing exercise, could help people's immune defences, reduce the number of people getting infections, and reduce how severe illnesses are and how long they last. The NIHR has funded the RECUR Programme to develop and test interventions to find out if they reduce the incidence of infections. The researchers have developed a website called Immune Defence which will help us to see if using nasal sprays or getting more physically active and reducing stress can help people get fewer and less severe infections.

1.3 Objectives

This study will estimate the effectiveness and cost-effectiveness of commonly available nasal sprays and a brief physical activity and stress management intervention in preventing and reducing the incidence, severity and duration of RTIs among patient at risk of serious infection in the COVID pandemic.

1.3.1 Primary objective

To assess whether three trial interventions (1) a microgel nasal spray 2) a nasal saline spray, or 3) support for physical activity and stress management reduce the duration of illness days due to respiratory tract infections (RTIs) among at-risk individuals when compared to usual care

1.3.2 Secondary objectives

To assess whether three trial interventions (1) a microgel nasal spray, 2) saline nasal spray, or 3) support for physical activity and stress management reduce:

i. the incidence of all respiratory tract infections
ii. health service contacts
iii. hospital admissions
iv. health service resource use (and to estimate cost-effectiveness of each intervention)

v. antibiotic use
vi. the incidence of COVID-like infections (during winters when COVID is circulating)

1.4 Definition of endpoints

A complete list of measures is provided in table 1 at the end of the document.

1.4.1 Definition of primary endpoint

The primary outcome will be the number of days of illness due to respiratory tract infections (RTIs) in total over 6 months. This will be based on the 6-month questionnaire, which asks the total number of days of illness since randomisation.

1.4.2 Definition of secondary endpoints

The following measures are at 6 and 12 months unless otherwise stated:

- number of infections with RTIs
- number of days with symptoms of RTIs rated moderately bad or worse
- number of days where work/normal activities were impaired
- number of consultations with the health service for RTIs
- number of courses of antibiotics taken for RTIs
- occurrence of Covid-19 infection

Beliefs about antibiotics and intention to consult are measured using a 6-point Likert scale

- patient beliefs in the efficacy of antibiotics
- intention to consult in the future

Physical activity measures:

- Short form International Physical Activity questionnaire[5]
- Sitting behaviours questionnaire[6]

Mental health measures:

- Perceived Stress Scale[7]
- PHQ-8[8] as a measure of current depression
- GAD-7[9] (Generalised Anxiety Disorder scale)

Side effects: Side effects at 3, 6 and 12 months (e.g., nasal irritation, stinging, nose bleeds, headache or sinus pain, and trips/falls while exercising).

1.5 Analysis principles

All analyses will be reported according to CONSORT 2010 on planning, implementing and reporting statistical analyses.

2 Design considerations

2.1 Description of trial design

An open, randomised, 4-arm trial evaluating:

i) usual care plus brief advice
ii) a microgel nasal spray
iii) saline nasal spray
iv) lifestyle intervention

We wish to provide estimates overall, and also for three strata of patients:

a) recurrent infections, no risk factors (stratum 1)
b) risk factors, no recurrent infections (stratum 2)
c) risk factors plus recurrent infections (stratum 3)

2.2 Trial power and sample size

2.2.1 Original sample size calculation

Primary outcome: number of days of illness in total due to RTIs. In the target population we assumed infection rates are not likely to be lower than 15%-20% over a 6-month winter/spring season. The primary comparisons are the independent comparisons between each intervention group and control (which only require an alpha of 0.05). However we wish to have the most power for this outcome since we wish to compare, not only each group with control, but each intervention group with each other, and therefore for this outcome we will allow for multiple testing and an alpha of 0.01.

Using the data from PRIMIT [1] , to detect a 1 day difference among individuals having an infection (hazard ratio 1.2) for alpha of 0.01 and 90% power requires 147 individuals per group, and allowing for at least 15% of individuals to contract an infection during a 6 month winter/spring period 980 individuals per group are needed and assuming 4 groups and 80% follow-up (which we achieved using similar methods to the PRIMIT trial), then 4900 individuals are needed. A 1 day difference is smaller than the difference found for both saline and Vicks First Defence in the previous trial[2], and the minimum required in discussion with our PPI collaborators.

Thus 5000 is the minimum sample required, but as in the original application preferably we wish to estimate these outcomes in each stratum (i.e., 14,700 patients), we aimed for 15,000 participants in total, which is feasible since the method of recruitment is as with PRIMIT by mailed invitation. During the first season, which was during the second winter of the COVID pandemic, we anticipated we could possibly recruit 5000 patients, enough to provide useful information about the role of these interventions during a future pandemic.

Occurrence of infections. We will have less power for this outcome, but if our primary aim, as with the primary outcome, is to compare each intervention with control, using the arguments of Cook and Farewell, since each analysis of intervention versus control is independent, this should not require a conservative Bonferroni correction[3], and we can use an alpha of 0.05. Using these assumptions, the above sample size will provide more than 80% power to estimate a 25% reduction in the incidence of infections from 20% to 15% in each

stratum (smaller than the reduction found in the initial studies using Vicks), and more than 90% power if the incidence of infections is 15% using all strata combined.

2.2.2 Updated sample size calculations

Revised estimates for the primary outcome based on data of the incidence of infections from the first two seasons (2020/21; 2021/22):

1. Stratum 1 (recurrence, no risk factors): Currently, in this group about 71%, compared to our original assumption of 15%, had an infection. Basing our recalculation on the more conservative lower limit of the 95% confidence interval of this estimate, we might assume that at least 65% of this group get an infection. Based on the original number per group of 147 (i.e., changing no other assumptions), we'd need 226 per group, and for 4 groups and with 80% follow up 1130. For 80% power, based on the original assumptions, we would require 111 per group. If 65% get an infection, then we would need 111/0.65=171 per group. For 4 groups with 80% follow up we would require 855 participants.
2. Stratum 2 (risk factors, no recurrence): In this group, the observed infection rate has been around 40% compared to our assumption of 15%. On this basis, based on the original sample size calculation of 147 per group, we would require 147/0.40=368 per group. With 4 groups and 80% follow up this gives a total of 1472 for 90% power. For 80% power, we would require 1388 participants.
3. Stratum 3 (risk factors plus recurrence): The infection rate is currently around 62% and follow up is just under 80%. On the basis of the original calculation, assuming at least 60% having an infection, we would need 245 per group, time 4 groups with 80% follow up gives 1225.

Revised estimates for the number of infections (secondary outcome):

1. In stratum 1 (recurrence, no risk factors): Assuming an infection rate of 65% and for 15% absolute reduction to 50% we require 227 complete cases in each group for 90% power (1135 with four groups and 80% follow up) and 170 for 80% power (850 in total)
2. In stratum 2 (risk factors, no recurrence): To detect a change from 40% to 30% we estimate we need 476 compete cases in each group for 90% power (2380 with four group and 80% follow-up) and 356 for 80% power (1780 in total).
3. In stratum 3 (risk factors plus recurrence): To detect a change from 60% to 45% for 90% power we need 231 complete cases in each group or 1155 for four groups and 80% follow-up.

2.3 Randomisation details

Following completion of the baseline measures, participants were randomised by the Immune Defence website to one of the 4 study arms in the ratio 1:1:1:1. The intervention and data collection software generates a randomisation sequence and a computer algorithm block randomised participants to the trial groups. As the randomisation is automated, the randomisation sequence was concealed from the trial team.

Patients were stratified on the basis of whether they are in a higher risk group (over 65 and/or having comorbid condition) and whether or not they have recurrent RTIs (≥3 in the last year) to three strata: stratum 1 (recurrence, no risk factors); stratum 2 (risk factors, no recurrence); stratum 3 (risk factors plus recurrence).

2.4 Inclusion criteria

1) Patients aged ≥18 years with a risk factor:
   a) Known weakened immune system due to a serious illness or medication (e.g. chemotherapy);
   b) Known heart disease;
   c) Known asthma or lung disease;
   d) Known diabetes ;
   e) Known mild hepatic impairment;
   f) Known stroke or neurological problem;
   g) Obesity (BMI≥30)
   h) Patients with >=3 episodes of an RTI in the last year

   AND

   Normally experience one or more RTIs per year

2) Patients aged ≥65 (*criterion removed for season 3*)

   AND

   Normally experience one or more RTIs per year

2.5 Timing of planned analyses

An early analysis of the primary and secondary outcomes at the 6-month timepoint was agreed with the TSC and the funder in view of severe winter pressures in the NHS, and will be carried out at 6 months after the last recruited participant. The last recruited participant will reach 6 months on 22/09/2023. Participants have 28 days to complete the outcome measures (online) and a further 2 weeks for paper/telephone follow up of non-responders. The final 6-month dataset will be complete on 03/11/2023. At this point, the 6m dataset will be locked for all participants and this dataset will be used for all 6 month outcomes. Additionally, all participants recruited during seasons 1 and 2 will have reached the 12m timepoint (end of the study) and this dataset will also be locked and used for all 12 month outcomes in participants recruited during season 1 and 2. The remaining analyses will be carried out at 12 months when all participants have completed the study. The 12 month dataset for participants recruited in season 3 will be this final locked dataset. The six month analysis will not include cost effectiveness analysis which will apply only to the full completed data.

2.5.1 Interim analyses and early stopping

No interim analyses are planned.

2.5.2 Stopping rules

No stopping rules are planned.

3 Statistical considerations

3.1 Definition of analysis populations

The primary analysis population will include all patients who were randomised regardless of their previous occurrence of infection.

3.1.1 Intention-to-treat analysis population

This population includes all patients that were randomised regardless of intervention adherence. All summaries and analysis will be on the ITT population unless otherwise specified.

3.1.2 Per-protocol analysis population

The per-protocol population includes all ITT patients who have adhered to their interventions defined as follows:

1) Nasal sprays – used nasal spray at first sign of infection and used it every time or most times in this situation. Or, they did not use their spray because they did not have the first signs of infection.
2) Exercise and stress – Completed core digital intervention content and answered yes to 6 month adherence question for physical activity and/or answered yes to 6 month adherence question for stress management.

3.2 Analysis software

Analysis will be carried out using Stata version 17 or higher, except for the imputation of the primary outcome which will be carried out using R version 4.3.1.

3.3 Methods for handling data

3.3.1 Withdrawal from trial

All data up until the point of patient withdrawal from the trial will be used in analyses unless the patient withdrew consent and does not wish for the data already collected prior to withdrawal to be used for the trial.

3.3.2 Missing data

Multiple imputation with chained equations will be used for the primary outcome. Imputation models will be run separately within each randomised group. For the 6 month outcome, the imputation model will include all variables in the analysis model (baseline days of illness and stratification variables), days of illness from monthly questionnaires in months 1 to 6, and variables which predict missingness: age, sex, IMD decile, baseline belief in antibiotics and baseline intention to consult. Baseline days of illness due to RTIs will be truncated at the 95th percentile. The mice package for zero inflated Poisson or negative

binomial outcomes in R uses either a Bayesian regression or bootstrap approach to impute missing values.[10]

Sensitivity analyses to the imputation will include: (1) assuming missing days of illness are on average 1 day better or worse than the observed days of illness; (2) complete cases only. Intercurrent events, e.g., death, will be treated as missing data. Secondary outcomes will be analysed using complete cases analysis.

3.3.3 Outliers

If outliers are found, defined by being above the upper quartile+1.5*IQR or below the lower quartile-1.5*IQR, then firstly the source data will be checked. If the source data shows that the data is correct, then the outliers may be excluded from the analysis to explore any differences in results.

3.3.4 Assumption checking and alternative methods

For count outcomes, assumptions for Poisson or negative binomial models will be checked, specifically the goodness-of-fit ratio of the Pearson χ^2 statistic to its corresponding degrees of freedom[4]. Plots of the standardized deviance residuals to the predicted counts will also be used to visually assess model fit. Where high numbers of zeros are anticipated due to the nature of the outcome measure, zero-inflated models will be fitted (see details below).

For continuous outcomes, the assumptions for linear regressions will be checked using QQ-plots for normality and residual vs fitted value plots for linearity and homoscedasticity. If linear modelling assumptions are not met, data transformations such as Box-Cox will be applied and back transformed at specific values. If no suitable parametric distribution fits the data, a non-parametric approach will be used.

3.3.5 Data transformations

No further data transformations will be used.

3.4 Definition of key derived variables

3.4.1 Primary outcome

The total number of days of illness due to RTIs at 6 months will be defined as 0 for those who did not experience infection in the first 6 months and use the 6-month question for those who did experience at least one infection, as this has the best completion. Missing outcomes will be imputed as specified in section 3.3.2.

3.4.2 Key secondary outcome

Incidence of infection with RTIs

Number of infections with RTIs at 6 months will be defined as 0 for those who did not experience an infection in the first 6 months and use the 6-month question for those who did experience one or more infections. This is similarly defined at 12 months.

3.4.3 Secondary outcomes

The numbers of days of bad symptoms, number of days of work/activities lost, consultations and antibiotics will be defined as 0 for those who did not experience an infection in the first 6 months and use the 6-month question for those who did experience at least one infection. These will be similarly defined at 12 months. The 6- and 12- month questions will be used rather than the monthly questionnaires as these have better completion than the monthly questionnaires.

Number of days with symptoms of RTIs rated moderately bad or worse

The total number of days of moderately bad or worse symptoms due to RTIs.

Number of days where work/activities were impaired

The total number of days off work/other activities due to RTIs.

Number of consultations

Number of times consulting a healthcare practitioner about an RTI.

Number of courses of antibiotics

Number of courses of antibiotics taken for RTIs.

Occurrence of Covid infection

Occurrence of Covid infection.

Belief in antibiotics

Belief in the effectiveness of antibiotics for RTIs measured on a scale of 1 (not at all effective) to 6 (extremely effective)

Intention to consult

Likelihood of seeing a doctor with similar infections in the future, measured on scale 1 (not at all likely) to 6 (extremely likely)

International Physical Activity Questionnaire (IPAQ)

International Physical Activity Questionnaire (IPAQ) plus strength and balance items

MET minutes achieved in each category (walking, moderate activity and vigorous activity) and total MET minutes of physical activity a week are calculated as follows:
Walking MET-minutes/week = 3.3 * walking minutes * walking days
Moderate MET-minutes/week = 4.0 * moderate-intensity activity minutes * moderate days

Vigorous MET-minutes/week = 8.0 * vigorous-intensity activity minutes * vigorous-intensity days
Total physical activity MET-minutes/week = sum of Walking + Moderate + Vigorous MET-minutes/week scores.

It is recommended that bouts of activity lasting less than 10 minutes duration are not counted. It is also recommended that activity bouts of greater than 3 hours are truncated, i.e., in each category a maximum of 21 hours of activity are permitted a week (3 hours x 7 days).

Sitting behaviours questionnaire (domain specific sitting)

Total time spent sitting per day on weekdays and weekends summed over five domains: Travel, Work, TV, Computer, Leisure.

Perceived Stress Scale

The 14 items are scored from 0 to 4. Questions 4, 5, 6, 7, 9, 10, 13 are reverse scored and summed to produce total scores ranging from 0 to 56. For interpretation, score categories are: Low Stress (scores 0 - 18); Moderate Stress (scores 19 - 37); High Stress (scores 38 - 56)

PHQ-8

The total score is the sum of the 8 items and ranges from 0 to 24. If more than 1 item is missing, then the value of the scale will be set to missing. If only one individual item score is missing, the missing item score will be imputed by the rounded (to 0 dp to match other score items) mean of all other available item scores for that participant to calculate the PHQ-8 total score. A score of 10 or greater is considered major depression, 20 or more is severe major depression.

GAD-7

The GAD-7 is scored on a scale of 0 to 21 with higher values indicating worse anxiety. A cut-off of 10 will be used to indicate generalised anxiety disorder. If only one individual item score is missing, the missing item score will be imputed by the rounded (to 0 dp to math other score items) mean of all other available item scores for that participant to calculate the GAD-7 total score. If more than one individual item is missing, the GAD-7 total score will not be calculated and will be left as missing.

3.5 General principles for reporting and analysis

Analyses will in general be reported using a significance level of 5%, corresponding to 95% confidence intervals unless otherwise stated. Each active intervention arm will be compared to usual care and to each other. For the primary analysis, comparisons of interventions to usual care will be reported using a significance level of 5%, corresponding to 95% confidence intervals.

For the comparisons of interventions to each other, analyses will be reported using a significance level of 1%, corresponding to 99% confidence intervals. For regression analyses, adjusted results will be reported. Descriptive statistics will be reported to 1 decimal place as number and percentage for categorical variables, and mean and standard deviation for continuous variables, or median and interquartile range for variables with a skewed distribution. The groups will be labelled Usual Care, First Defence, Nasal Saline, Lifestyle.

4 Planned analyses and reporting

4.1 Disposition of the study population

CONSORT flow diagram (following CONSORT guidelines) which should include:

- Screening data – total number screened, reasons for not entering trial)
- Summary of eligibility data to be presented - total number assessed for eligibility, breakdown of patients screened against the eligibility criteria
- Recruitment information - number consented, recruited/randomised, receiving allowed treatment, withdrawing/lost to follow up at each time point
- Analysis population - number included, reasons excluded

4.2 Protocol deviations

A protocol deviation is an unanticipated or unintentional divergence or departure from the expected conduct of a study inconsistent with the protocol, consent documents or other study procedures. Of particular importance are major deviations (violations) which may expose participants to increased risk; compromise the integrity of the entire study or affect participant eligibility. Major protocol deviations will be listed with information on treatment group and the type of deviation.

4.3 Baseline and demographic characteristics

Patient socio-demographic measures include: gender, age; marital status; years of education; ethnicity; BMI; smoking status; index of multiple deprivation; number of children under 16 years in the household; number of household members.

Clinical/behavioural measures collected from patient report: number of RTIs in last year; total days with symptoms, days with moderately bad or worse RTI symptoms, days lost to work/other activities; visits to the doctor for RTIs in the past year; number of antibiotic prescriptions for RTIs in the past year; COVID symptoms in the last 12 months; results of COVID testing (PCR or antibody) in the last 12 months; other health problems; influenza vaccine in current season; COVID vaccine in the last 12 months.

Baseline outcome measures as listed in Table 1, including Beliefs in antibiotics, International Physical Activity Questionnaire, Domain-Specific Sitting, Perceived Stress Scale, PHQ-8, and GAD-7.

4.4 Primary analysis

The primary analysis will be a zero-inflated Poisson regression, or zero-inflated negative binomial regression if over-dispersed, of number of days of illness due to RTIs at 6 months, adjusting for baseline days of illness due to RTIs and stratification variables (comorbidity, recurrent infection and their product). Baseline days of illness due to RTIs will be truncated at the 95^{th} percentile of non-zero values for all primary and secondary analyses.

The three intervention groups will be compared with usual care and clinical effectiveness will be concluded if the point estimate (IRR) favours the 'intervention' and the 95% confidence interval excludes 1. The interventions will also be compared with each other, clinical effectiveness will be concluded if the 99% confidence interval excludes 1.

The primary analysis will be based on all patients randomised. A sensitivity analysis will include only the patients who have at least one infection in a normal year, as defined by the baseline variable. The primary analysis will use the 6-month question, as this is the primary outcome and has the best completion. As a sensitivity analysis, we will sum the monthly number of days of illness due to RTIs over 6 months.

The primary analysis will use multiple imputation with chained equations. The imputation model will be implemented as specified in section 3.3.2. Sensitivity analyses to the imputation will include: (1) assuming missing days of illness are on average 1 day better or worse than the observed days of illness, and also (2) a complete cases analysis.

The primary analysis model will be fitted overall adjusting for strata, and separately within each stratum.

4.5 Secondary analyses

The following outcome measures will be analysed at 6 and 12 months based on complete cases:

4.5.1 Incidence of infection with RTIs

The number of infections will be analysed using zero-inflated Poisson/negative binomial regression, depending on model fit, adjusting for baseline number of infections and the same covariates as in the primary analysis.

For the incidence of infection data, the primary comparison will be between each intervention group and usual care, and if any of these demonstrate an intervention is effective at the 5% level then we will conduct exploratory comparisons between groups.

4.5.2 Number of days with symptoms of RTIs rated moderately bad or worse

Number of days will be analysed using zero-inflated Poisson/negative binomial regression, adjusting for the baseline number of days with bad/worse symptoms and the same covariates as in the primary analysis.

4.5.3 Number of days where work/activities were impaired

Number of days will be analysed using zero-inflated Poisson/negative binomial regression, adjusting for baseline number of days of work lost and the same covariates as in the primary analysis.

4.5.4 Contact with health service

Number of contacts with the health service will be analysed using zero-inflated Poisson/negative binomial regression, adjusting for baseline number of contacts with the health service and the same baseline variables as in the primary analysis.

4.5.5 Courses of antibiotics

The number of courses of antibiotics taken will be analysed using zero-inflated Poisson/negative binomial regression, adjusting for baseline number of courses of antibiotics and the same baseline variables as in the primary analysis.

4.5.6 Occurrence of Covid-19

Occurrence of Covid-19 will be analysed using logistic regression, adjusting for baseline covid vaccination status (yes/no) and the same covariates as in the primary analysis.

4.5.7 Belief in antibiotics

Belief in antibiotics (on a scale 0 to 6) will be analysed using ordinal logistic regression, adjusting for baseline belief in antibiotics and the same variables as in the primary analysis.

4.5.8 Intention to consult

Intention to consult (on a scale 0 to 6) will be analysed using linear regression, or ordinal logistic regression if normality assumptions are not satisfied, adjusting for baseline intention to consult and the same baseline variables as in the primary analysis.

4.5.9 International Physical Activity Questionnaire (IPAQ)

Total MET minutes will be analysed using linear regression, adjusting for baseline MET minutes and the same baseline covariates as in the primary analysis.

4.5.10 Sitting behaviours questionnaire (domain specific sitting)

Total minutes sitting per day on weekdays and weekends will be analysed using linear regression, adjusting for baseline minutes sitting and the same baseline covariates as in the primary analysis.

4.5.11 Perceived Stress Scale

Perceived Stress Scale will be analysed using linear regression, adjusting for baseline PSS score and the same baseline covariates as in the primary analysis.

4.5.12 PHQ-8

Depression (defined as a score >=10 on PHQ-8) will be analysed using logistic regression, adjusting for baseline PHQ-8 score and the same covariates as in the primary analysis.

4.5.13 GAD-7

Anxiety (defined as a score >=10 on GAD-7) will be analysed using logistic regression adjusting for baseline GAD-7 score and the same covariates as in the primary analysis.

4.5.14 Side effects

Side effects (nosebleeds, nasal irritation, headache, trips/falls) will be analysed using logistic regression.

4.5.15 Cold avoidance behaviours

Frequency of using cold avoidance behaviours, such as washing hands, wearing a face covering, social distancing, will be summarised descriptively by randomised arm.

4.5.16 Use of nasal sprays

Use of nasal sprays will be summarised descriptively in the First defence and Saline nasal spray arms. How the nasal sprays were used (preventatively, at the start of the illness or both) will also be summarised descriptively.

4.5.17 Adverse events

Adverse events will be summarised descriptively by randomised arm. This will include reporting of facial pain and trigeminal neuralgia.

4.6 Additional analyses

4.6.1 Subgroup analyses

Subgroup analyses will be performed by repeating the primary analysis and including a treatment by covariate interaction term for the following subgroups:

- recurrent infections (>3/year vs <=3 per year)
- age: <=65 vs >65
- presences of comorbidities – presence/absence of serious comorbidities
- number of serious comorbidities – 2 or more vs 1 or fewer
- use of nasal spray – preventively/onset of illness/both
- recruitment season - 1, 2 or 3
- Flu vaccination within last 12 months – yes/no
- Covid vaccination within last 12 months – yes/no
- For the Lifestyle vs usual care comparison: low physical activity (IPAQ = inactive or minimal) and psychological distress (PHQ-8>=5, GAD-7>=5) vs all others

4.6.2 Per-protocol analysis

A per-protocol analysis comparing those who adhere to the interventions, as defined in section 3.1.2, compared to usual care will be carried out.

4.6.3 Sensitivity analyses

The primary analysis will be based on all patients randomised, which includes those who were included as 'at-risk' due to COVID in season one as well as those who report at least one infection in a normal year. We will also conduct a sensitivity analysis to include just those who report at least one infection in a normal year.

The primary analysis will use the 6-month question, as this is the primary outcome and has the best completion. The question asks for all days of illness due to RTIs since randomisation. As a sensitivity analysis, we will sum the monthly number of days of illness due to RTIs reported at each month over 6 months, accepting this may risk double-counting some days of illness.

The primary analysis will use multiple imputation with chained equations. The imputation model will include all variables in the analysis model and any baseline covariates associated with the missingness of the number of days of illness. Sensitivity analyses to the imputation will include: (1) assuming missing days of illness are on average 1 day better or worse than the observed days of illness, and also (2) a complete cases analysis.

If the frequency of one or more baseline cold avoidance behaviours is observed to be very different between arms, we will undertake a sensitivity analysis adjusting for the frequency of these behaviours.

4.6.4 Further analyses

The health economic analysis plan and the process evaluation plan will be outlined in separate documents.

4.7 Safety reporting

This study shall adhere to the EU Medical Device Regulation (MDR) Article 2 (58): , where a "Serious Adverse Event" (SAE) is defined as any adverse event that led to:

a) death,

b) serious deterioration in the health of the subject, that resulted in any of the following:

i. life-threatening illness or injury,

ii. permanent impairment of a body structure or a body function,

iii. hospitalisation or prolongation of patient hospitalisation,

iv. medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function,

v. chronic disease,

c) foetal distress, foetal death or a congenital physical or mental impairment or birth defect.

SAEs will be listed and summarised descriptively by severity, relatedness and expectedness.

5 Tables, listings and figures templates

Table 1: Outcome and process measures

Measures	Screening	Baseline	Illness diaries	4 weeks	8 weeks	12 weeks	16 weeks	20 weeks	6 months	28 weeks	32 weeks	36 weeks	40 weeks	44 weeks	48 weeks	12 months
Screening	x															
Respiratory illnesses:																
Number and type of infection		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Checklist of symptoms if infection occurred			x	x	x	x	x	x	x	x	x	x	x	x	x	X
Days with infection		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Days with mod /severe infection		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Symptom severity			x													
Days off work /normal activities		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Appointments with HCP		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
If tested for COVID and the result		x		x	x	x	x	x	x	x	x	x	x	x	x	x
COVID vaccine		x				x			x							x
Vitamin D									x							x
Courses of antibiotics		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Severity/other health problems		x														
Psychological Process measures		x		x					x							x
Adherence to treatments/intervention			x						x							x
Beliefs in antibiotics		x							x							x
Intention to consult		x							x							x
Demographics		X														
IPAQ		x							x							x
Domain specific sitting		x							x							x
PHQ-8		x							x							X

GAD-7		x							x							X
EQ5D		X	x			x			x							X
Out of pocket expenses						x			x							x
Cold avoidance behaviours		x				x			x							x
Use of nasal sprays		x				x			x							x
Adverse events						x			x							x

6 References

1. Little P, Stuart B, Hobbs FDRea. An internet-delivered handwashing intervention to modify influenza-like illness and respiratory infection transmission (PRIMIT): a primary care randomised trial. *Lancet* 2015;DOI: http://dx.doi.org/10.1016/S0140-6736(15)60127-1

2. Hull D, Rennie P, Noronha A, et al. Effects of creating a non-specific, virus-hostile environment in the nasopharynx on symptoms and duration of common cold. *Acta otorhinolaryngologica Italica* 2007;27:73-77.

3. Cook R, Farewell V. Multiplicity Considerations in the Design and Analysis of Clinical Trials. *J Royal Stats Soc* 1996;Series A 159:93-110.

4. Payne EH, Gebregziabher M, Hardin JW, Ramakrishnan V, Egede LE. An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. Commun Stat Simul Comput. 2018 Jul 5;47(6):1722-1738. doi: 10.1080/03610918.2017.1323223. PMID: 30555205; PMCID: PMC6290908.

5. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35(8):1381-95. doi: 10.1249/01.MSS.0000078924.61453.FB [published Online First: 2003/08/06]

6. Marshall AL, Miller YD, Burton NW, et al. Measuring total and domain-specific sitting: a study of reliability and validity. *Med Sci Sports Exerc* 2010;42(6):1094-102. doi: 10.1249/MSS.0b013e3181c5ec18 [published Online First: 2009/12/10]

7. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24(4):385-96. [published Online First: 1983/12/01]

8. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606-13. [published Online First: 2001/09/15]

9. Ruiz MA, Zamorano E, Garcia-Campayo J, et al. Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. *Journal of affective disorders* 2011;128(3):277-86. doi: 10.1016/j.jad.2010.07.010 [published Online First: 2010/08/10]

10. Kleinke, K. and Reinecke, J. (2013), Multiple imputation of incomplete zero-inflated count data. Statistica Neerlandica, 67: 311-336. https://doi.org/10.1111/stan.12009

7 SAP revision history

Version number	Revision history	Author	Date