Statistical Analysis Plan Contagious Misinformation Trial

The statistical analyses will be carried out after the last participant has completed the follow-up survey. Analyses will be performed by a team of researchers from Karolinska Institutet in collaboration with FOCUS1000 and COMAHS in Freetown, Sierra Leone and New York University in the USA.

The flow of the participants in the Contagious Misinformation Trial will be reported according to the Consolidation Standard of Reporting Trials (CONSORT), see figure 1.

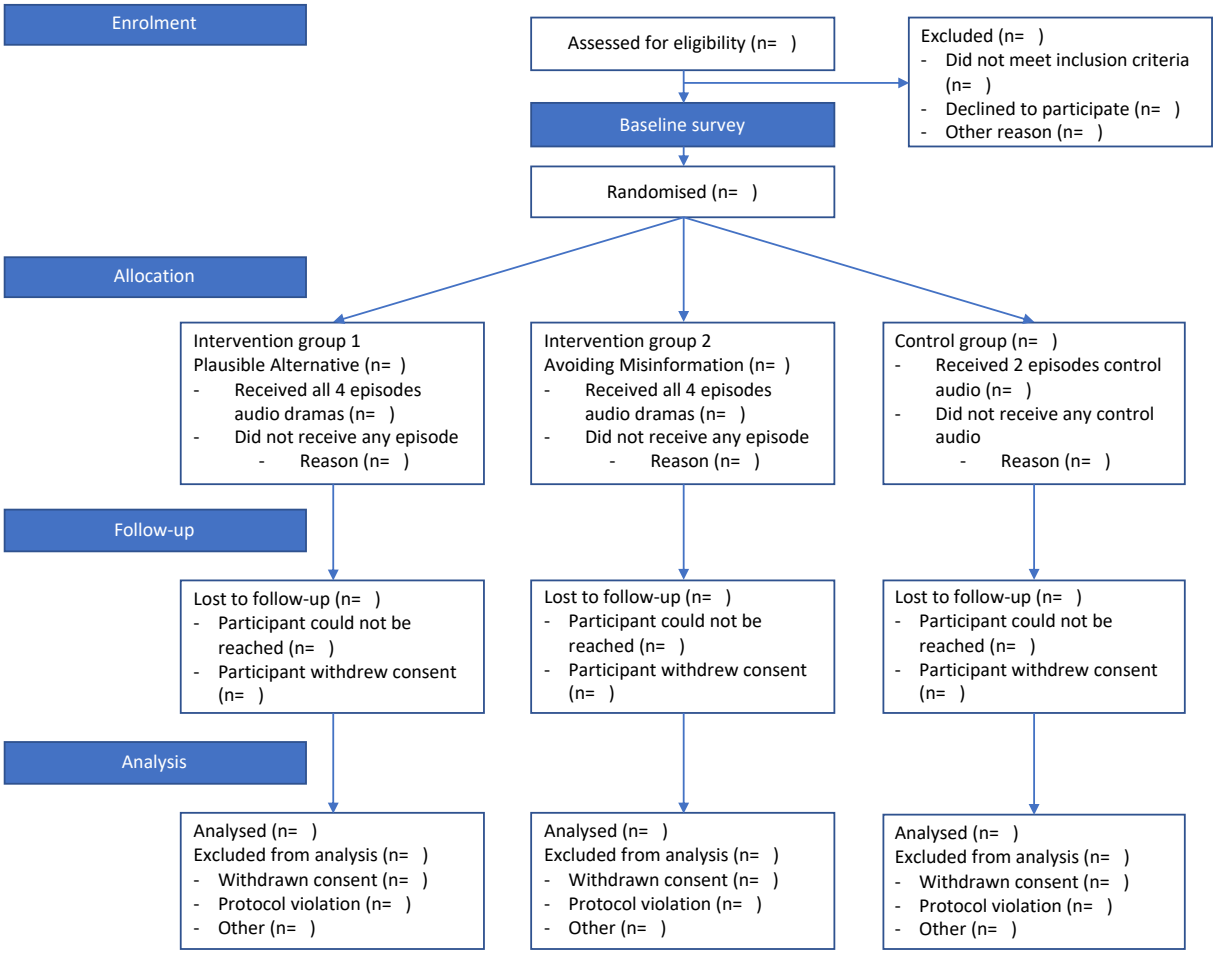

Figure 1. Participant flow diagram

Baseline characteristics

The baseline characteristics of the participants will be reported in table 1 per randomization group. In addition to demographic variables such as age, education, sex and religion, table 1 will also include the baseline prevalence of the belief that typhoid is caused by mosquitoes and the belief that typhoid and malaria come together. Differences between the study groups will be analysed using a t-test or Chi-square test (depending on the variable) and p-values will be reported in the last column.

Analysis of the primary outcomes

The primary analysis is an intention-to-treat analysis (ITT), meaning that participants will be analysed according to their assigned group at baseline (i.e. intervention group 1, intervention group 2 or control group). There are two primary outcomes in this trial, as assessed at follow-up:

1. The belief that typhoid is caused by mosquitoes
2. The belief that typhoid and malaria come together

Both outcomes are assessed through yes/no questions in the follow-up survey. In this ITT analysis, we will analyse the data through logistic regression models. Crude models will be fitted, followed by models adjusted for various covariates. In all models, the effect of the intervention on the outcome at follow-up is adjusted for the status of the outcome at baseline. Using post-hoc tests, we will compare intervention group 1 with intervention group 2 for the two primary outcomes, based on the adjusted models. Odds Ratios and the 95% Confidence Intervals will be reported in tables; table shells for the proposed analyses are provided in tables 2 in the appendix. Odds Ratios of the covariates will be reported in separate tables in a supplementary file. Participants who were lost to follow-up will be included in this ITT analysis, using the multiple imputations method for their missing data. All other analyses in this plan will use complete cases only.

Furthermore, the primary outcomes will be analyses in a per-protocol analysis, where we use complete cases only and exclude participants who were lost to follow-up, had protocol violations or withdrew their consent. The models will be the same as for the ITT analysis.

Model	**Predictor**	**Outcomes**
Model 1-2	Baseline outcome Randomization Control group Intervention group 1 Intervention group 2	Belief that typhoid is caused by mosquitoes (yes/no) Belief that typhoid and malaria come together (yes/no)
Model 3-4	Baseline outcome Randomization Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Belief that typhoid is caused by mosquitoes (yes/no) Belief that typhoid and malaria come together (yes/no)

Analysis of secondary outcomes

<u>Secondary outcome 1: Backfire effect</u>

The first secondary outcome will evaluate whether there was a backfire effect from the interventions. In order to analyse this, we will create a variable that compares the baseline belief in the two primary outcomes with the follow-up belief in the two primary outcomes. This means we will have 3 possible values in this variable:

1. No change in belief in misinformation between baseline and follow-up

2. Decrease in belief in misinformation between baseline and follow-up
3. Increase in belief in misinformation between baseline and follow-up

We will fit ordinal logistic regression to analyse the potential backfire effect. Since our outcome variable is already adjusted for the baseline score, we will not adjust the models for the baseline score. Similar to the analysis of the primary outcome, a post-hoc test will be carried out to compare intervention group 1 and intervention group 2. Odds Ratios and 95% Confidence Intervals will be reported in a table, see table shell 3 in the Appendix.

Model	**Predictor**	**Outcome**
Model 1-2	Randomization Control group Intervention group 1 Intervention group 2	Backfire effect: Belief that typhoid is caused by mosquitoes (yes/no) Backfire effect: Belief that typhoid and malaria come together (yes/no)
Model 3-4	Randomization Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Backfire effect: Belief that typhoid is caused by mosquitoes (yes/no) Backfire effect: Belief that typhoid and malaria come together (yes/no)

<u>Secondary outcome 2: As-treated-analysis</u>

The second secondary outcome will follow the same analysis plan as the primary outcome and the first secondary outcome, with the difference that only those participants that have listened to the audio dramas will be included in these analyses. This will be tested in the follow-up survey with three questions that ask about characters and main events of the audio dramas. Only participants that answer these questions correctly will be included in the as-treated-analysis. In case participants of the control group have answered these questions correctly, they will also be included in the appropriate intervention group. Logistic regression models will be fitted for the primary outcomes, and ordinal logistic regression will be fitted for the backfire effect analysis. Odds Ratios with 95% Confidence Intervals will be reported in tables, see table shells 4 and 5 in the Appendix.

<u>Secondary outcome 3: Knowledge about preventive methods for typhoid</u>

The question 'Can you name up to three ways how you can prevent yourself from getting typhoid?' in the baseline and the follow-up survey will be used for this analysis. This is an open question; the data collector ticks the appropriate boxes as the participant answers the question. Based on the answers, we create a score, whereby one point will be assigned to every correct answer, and one point will be subtracted for every wrong answer. The score can therefore range between -3 (when respondents give 3 incorrect examples of preventive methods) to +3 (when respondents give 3 correct examples of preventive methods). Ordinal logistic regression models will be fitted, first a model with only the randomization group as a predictor, after which a multivariate model will be fitted that includes baseline score on the preventive methods question and demographics. Post-hoc test will be carried out to compare

intervention group 1 and intervention group 2. Results will be reported as demonstrated in table shell 6 in the appendix.

Model	**Predictor**	**Outcome**
Model 1	Baseline outcome Randomization Control group Intervention group 1 Intervention group 2	Preventive methods score: -3 to +3
Model 2	Baseline outcome Randomization Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Preventive methods score: -3 to +3

Secondary outcome 4: Health-related discussions

The question 'did you discuss the audio fragments with friends and/or family?' in the follow-up survey is used for the analysis of this secondary outcome. Similar to the previously described analyses, logistic regression models will be fitted. In the multivariate model, the question 'how often do you discuss health issues with family or friends?' from the baseline survey will be adjusted for, as well as demographics. Post-hoc tests will be applied to compare intervention group 1 with intervention group 2. Results will be reported in table 7 in the appendix.

Model	**Predictor**	**Outcome**
Model 1	Randomization Control group Intervention group 1 Intervention group 2	Discussed audio with family/friends Yes No
Model 2	Randomization Control group Intervention group 1 Intervention group 2 Baseline score health-related discussions with family/friends Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Discussed audio with family/friends Yes No

Secondary outcome 5: Non-WhatsApp group

The last secondary outcome will follow a same strategy as the primary outcomes, but will instead focus on the 60 people in the trial that did not have WhatsApp. This group has similarly been randomised intervention group 1 (n=30) and intervention group 2 (n=30). The two primary outcomes will be compared with the control group of the larger trial (n=200) as well as to each other. Furthermore, the same analyses will be carried out comparing Non-WhatsApp intervention group 1 with WhatsApp intervention group 1, as well as Non-WhatsApp intervention group 2 with WhatsApp intervention group 2. Characteristics of the sample will be reported. Results will be summarized descriptively, and a Chi-square analysis will be carried out for the two primary outcomes and reported like table 8A and 8B in the appendix. Depending on statistical power, we will also carry out logistic regression.

Secondary outcome 6: Self-efficacy

There will be three questions about self-efficacy in the follow-up survey, relating to the three main preventive methods of typhoid: cooking food properly, drinking only treated water and washing hands after going to the toilet. The scale will be converted to reflect 'no self-efficacy', 'some self-efficacy' and 'high self-efficacy'. Through ordinal logistic regression, we will compare the average self-efficacy between the intervention groups and the control group.

Model	**Predictor**	**Outcome**
Model 1-3	Randomization Control group Intervention group 1 Intervention group 2	Self-efficacy: - Cooking food - Drinking treated water - Washing hands
Model 4-6	Randomization Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Self-efficacy: - Cooking food - Drinking treated water - Washing hands

Secondary outcome 7: Risk perception and preventive practices

We will analyse (likely in a separate paper) the associations between risk perception and typhoid-preventive practices and the influence of the intervention on those associations. Will test the three hypotheses put forward by Brewer et al (2014):

1. Behaviour Motivation Hypothesis
2. Risk Reappraisal Hypothesis
3. Accuracy Hypothesis

Behaviour Motivation Hypothesis

This hypothesis argues that risk perception can influence someone's likelihood to take action. We will analyse this through testing the associations between risk perception at baseline and preventive behaviour at follow-up. Risk perception is a

Model	**Predictor**	**Outcome**

1	Risk Perception at baseline (likely, unlikely) Randomisation Control group Intervention group 1 Intervention group 2	Behaviour at follow-up
2	Risk Perception at baseline (likely, unlikely) Randomization Control group Intervention group 1 Intervention group 2 Behaviour at baseline Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Behaviour at follow-up

Risk Reappraisal Hypothesis

Model	**Predictor**	**Outcome**
1	Behaviour at follow-up Randomisation Control group Intervention group 1 Intervention group 2	Change in risk perception between baseline and follow-up
2	Behaviour at follow-up Randomization Control group Intervention group 1 Intervention group 2 Behaviour at baseline Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Change in risk perception between baseline and follow-up

Accuracy Hypothesis

Model	**Predictor**	**Outcome**
1	Risk perception at baseline Randomisation Control group Intervention group 1 Intervention group 2	Behaviour at baseline
2	Risk perception at baseline Randomization	Behaviour at baseline

	Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	
3	Risk perception at follow-up Randomisation Control group Intervention group 1 Intervention group 2	Behaviour at follow-up
2	Risk perception at follow-up Randomization Control group Intervention group 1 Intervention group 2 Age (18-30, 31-49, >50) Sex (M/F) Education (no/prim/second +) Religion (Islam, Christianity, other) Income (4 levels)	Behaviour at follow-up

Secondary outcome 7: Subjective vs objective learning

The last secondary outcome uses the answers to the follow-up question 'do you feel you learned something from the audio drama?' and compare that the actual learning (i.e. through the two primary outcomes). A Chi-square test will be done, across all groups, and per intervention group.

Post hoc analysis

If the analyses yield significant results, post hoc analyses will be carried out to determine if there are statistically significant differences between the two different intervention groups.

Appendix – Table Shells

Table 1. Baseline characteristics by randomization group

	Intervention group 1	Intervention group 2	Control group	p-value
Age (years) 18-30 31-49 >50				
Sex Male Female				
Education No education Primary Secondary +				
Religion Islam Christianity Other				
Income (Le) 30.000-300.000 300.000-1.000.000 1.000.000-5.000.000 >5.000.000				
Baseline Prevalence: Belief in mosquitoes causing typhoid				
Baseline Prevalence: Belief that typhoid and malaria come together				
Total				

Table 2. Intention-to-treat analysis of the primary outcomes by intervention group

	Primary outcome 1, intervention group 1				
	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1				
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 1 vs intervention group 2				
Randomization Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1 vs intervention group 2				
Randomization					

Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Adjusted for: baseline score on outcome, age, sex, education, religion, income					

Table 3. Backfire effect

	Primary outcome 1, intervention group 1				
	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1				
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 1 vs intervention group 2				
Randomization Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1 vs intervention group 2				
Randomization Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Adjusted for: age, sex, education, religion, income					

Table 4. As-treated-analysis of the primary outcomes by intervention group

	Primary outcome 1, intervention group 1				
	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Follow up Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 2				
Follow up Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1				
Follow up Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 2				
Follow up Control Group		1.0 (Reference)		1.0 (Reference)	

Intervention Group 2					
	Primary outcome 1, intervention group 1 vs intervention group 2				
Follow up Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1 vs intervention group 2				
Follow up Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Adjusted for: baseline score on outcome, age, sex, education, religion, income					

Table 5. As-treated-analysis of backfire effect

	Primary outcome 1, intervention group 1				
	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1				
Randomization Control Group Intervention Group 1		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 2				
Randomization Control Group Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 1, intervention group 1 vs intervention group 2				
Randomization Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
	Primary outcome 2, intervention group 1 vs intervention group 2				
Randomization Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Adjusted for: age, sex, education, religion, income					

Table 6. The influence of the intervention on knowledge of preventive methods for typhoid

	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Randomization Control Group Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Baseline score -3 -2 -1 0 +1 +2 +3		-	-	1.0 (Reference)	

Age (years) 18-30 31-49 >50		-	-	1.0 (Reference)	
Sex Male Female		-	-	1.0 (Reference)	
Education No education Primary Secondary +		-	-	1.0 (Reference)	
Religion Islam Christianity Other		-	-	1.0 (Reference)	
Income (Le) 30.000-300.000 300.000-1.000.000 1.000.000-5.000.000 >5.000.000		-	-	1.0 (Reference)	

Table 7. The influence of the intervention on health-related discussion with family and/or friends

	No (%) Respondents	Crude OR (95% CI)	P value	Adjusted OR (95% CI)	P value
Randomization Control Group Intervention Group 1 Intervention Group 2		1.0 (Reference)		1.0 (Reference)	
Baseline score: discuss health with fam/friends No Yes		-	-	1.0 (Reference)	
Age (years) 18-30 31-49 >50		-	-	1.0 (Reference)	
Sex Male Female		-	-	1.0 (Reference)	
Education No education Primary Secondary +		-	-	1.0 (Reference)	
Religion Islam Christianity Other		-	-	1.0 (Reference)	
Income (Le) 30.000-300.000 300.000-1.000.000 1.000.000-5.000.000 >5.000.000		-	-	1.0 (Reference)	

Table 8A. Chi-square analysis of primary analysis of primary outcome 1 by intervention groups among the Non-WhatsApp participants

	Intervention group 1	Control group
Baseline prevalence		
Follow-up prevalence		

	Intervention group 2	Control group
Baseline prevalence		
Follow-up prevalence		
	Intervention group 1	WhatsApp Intervention group 1
Baseline prevalence		
Follow-up prevalence		
	Intervention group 2	WhatsApp Intervention group 2
Baseline prevalence		
Follow-up prevalence		

Table 8B. Chi-square analysis of primary analysis of primary outcome 2 by intervention groups among the Non-WhatsApp participants

	Intervention group 1	Control group
Baseline prevalence		
Follow-up prevalence		
	Intervention group 2	Control group
Baseline prevalence		
Follow-up prevalence		
	Intervention group 1	WhatsApp Intervention group 1
Baseline prevalence		
Follow-up prevalence		
	Intervention group 2	WhatsApp Intervention group 2
Baseline prevalence		
Follow-up prevalence		