S1: Study protocol

Specification for Herd immunity
April, 2019

Overview:

This study is a PhD project conducted by Ms. Hina Hakim, supervised by Dr. Holly O. Witteman, PhD, and co-supervised by Dr. Daniel Reinharz, professors, and researchers at the Faculty of Medicine at Laval University. The purpose of this study is to evaluate the effects of visualization conveying the concept of community immunity or herd immunity on risk perception (towards individual, family, community and vulnerable people in communities) (primary outcome) and on emotions, attitudes, knowledge, and behavioural intentions (secondary outcomes).

People Involved:

Hakim, H (PhD candidate)
Witteman, H (Research supervisor)
Parent, E (Designer and adviser on designing)
Aziaka, D (Developer/Programmer)
Noubi, M (Developer/Programmer)
Toudjeu, NT (Developer/Programmer)

Collaborating with:

- The members of the National Institute of Public Health of Quebec (INSPQ) whose comments we have received as we planned this project are: Eve Dubé, Marc Dionne, Rodica Gilca, Manale Ouakki and Dominique Gagnon.
- The members of Public Health Ontario whose comments we have received as we planned this project are: Beate Sander and Bryna Warshawsky.

Co-investigators:

Hina Hakim, Christine T. Chambers, Eve Dubé, S. Michelle Driedger, Teresa Gavaruzzi, Anik M. C. Giguere , Noah M. Ivers, Éric Kavanagh, Shannon MacDonald, Rita Orji, Elizabeth Parent, Jean-Sébastien Paquette, Jacynthe Roberge, Beate Sander, Aaron Scherer, Bryna Warshawsky, Kumanan Wilson, Daniel Reinharz, Holly O. Witteman

INTRODUCTION

Background and rationale

Herd immunity or community immunity exists when enough people in a given population are vaccinated. This provides an indirect protection to those who cannot get vaccinated due to a weak immune system as either they are too young, too old or immunocompromised (vulnerable people). Recent drop in vaccine uptake results in an outbreak of vaccine-preventable diseases which increases the risk of everyone to be in at high risk of being infected, especially vulnerable people. Research suggested that effective communication about herd immunity may increase vaccination intention (Betsch et al. 2017).

Visualization is a powerful communication mechanism that uses pre-attentive processing to communicate large amounts of information rapidly in understandable and compelling ways (Healey and Enns 2012). A systematic review demonstrates that there are some interventions available for conveying the concept of community immunity, and very few evaluate interventions for their effects on vaccine intentions and uptake as well as their precursors, such as knowledge, attitudes, and none on emotions (Hakim et al. 2018).

The purpose of this study is to evaluate the effects of visualization conveying the concept of community immunity on risk perception (to individual, family, community and vulnerable people in communities) (primary outcome) and on emotions, attitudes, knowledge, and behavioural intentions (secondary outcomes).

INTRODUCTION

Contexte et justification

L'immunité collective existe quand suffisamment de personnes dans une population donnée sont vaccinées. Cela fournit une protection indirecte aux personnes qui ne peuvent pas se faire vacciner en raison de la faiblesse de leur système immunitaire, du fait qu'elles soient trop jeunes, trop âgées ou immunodéprimées (personnes vulnérables). La récente baisse des taux de vaccination a entraîné une épidémie des maladies pouvant être prévenues par la vaccination, ce qui augmente le risque pour chacun d'être infecté, en particulier les personnes vulnérables. Les recherches ont suggéré qu'une communication efficace sur l'immunité collective pourrait augmenter l'intention de vaccination (Betsch et al. 2017).

La visualisation est un mécanisme de communication puissant qui utilise un traitement pré-attentif pour communiquer rapidement de grandes quantités d'informations de manière compréhensible et convaincante (Healey et Enns 2012). Une revue systématique a montré que certaines interventions permettent de véhiculer le concept d'immunité collective et que très peu évaluent leurs effets sur les intentions liées à la vaccination ainsi que sur leurs précurseurs, telles que les connaissances, les attitudes, mais aucune ne portant sur les émotions (Hakim et al 2018).

Le but de cette étude est d'évaluer les effets de la visualisation transmettant le concept d'immunité collective sur la perception du risque (pour l'individu, la famille, la communauté et les personnes vulnérables dans les communautés) (résultat principal) et sur les émotions, les attitudes, les connaissances et les intentions comportementales (résultats secondaires).

Objectives

To evaluate the effects of visualization conveying the concept of community immunity on risk perception (to individual, family, community and vulnerable people in communities) (primary outcome) and on emotions, attitudes, knowledge, and behavioural intentions (secondary outcomes). We hypothesize that the new visualization "Herdimm" (A visualization developed iteratively in our previous study) will help people understand the concept of community immunity better than the control condition or comparator:

- Control: No information about community immunity or herd immunity
- Intervention: Herdimm visualization

We will also conduct secondary analyses comparing the effects of visualization (Herdimm) versus five comparators:

- Measles: Guardian visualization (https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated)
- Measles: Visualization by theotheredmund: (https://imgur.com/gallery/8M7q8#J7LANQ4)
- Generic: Robert Koch Institut visualization (http://rocs.hu-berlin.de/D3/herd/)
- Generic: SBS News visualization: (https://www.sbs.com.au/news/two-sydney-babies-too-young-to-be-vaccinated-infected-with-measles?fbclid=IwAR1Rns_ycauiZeTeY1HC22eJe8JE6dwcQG0A9fgq305WCskIHBT_AGzExLE)
- Influenza: Public Health Agency of Canada Flu: Don't Pass It On! (https://www.canada.ca/en/public-health/services/video/the-flu-don-t-pass-it-on.html)

Objectifs

Évaluer les effets de la visualisation traduisant le concept d'immunité communautaire sur la perception du risque (pour l'individu, la famille, la communauté et les personnes vulnérables dans la communauté) (résultat principal) et sur les émotions, les attitudes, les Connaissances et les intentions comportementales (résultats secondaires).

Nous émettons l'hypothèse que la nouvelle visualisation « Herdimm » (visualisation développée de manière itérative dans notre précédente étude) aidera les gens à mieux comprendre le concept d'immunité collective que la condition de contrôle ou le comparateur :

- Contrôle : aucune éducation sur l'immunité de la communauté ou l'immunité collective
- Intervention : visualisation « Herdimm »

Nous effectuerons également des analyses secondaires comparant les effets de la visualisation (Herdimm) à ceux de six comparateurs :

- Rougeole : Visualisation de The Guardian (https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated)
- Rougeole : Visualisation par theotheredmund: (https://imgur.com/gallery/8M7q8#J7LANQ4)
- Générique : Visualisation de l'Institut Robert Koch (http://rocs.hu-berlin.de/D3/herd/)
- Générique : Visualisation de SBS (The Special Broadcasting Service) : (https://www.sbs.com.au/news/two-sydney-babies-too-young-to-be-vaccinated-infected-with-measles?fbclid=IwAR1Rns_ycauiZeTeY1HC22eJe8JE6dwcQG0A9fgq305WCskIHBT_AGzExLE)
- Influenza : Agence de la santé publique du Canada Grippe: ne le transmettez pas! (https://www.canada.ca/fr/sante-publique/services/video/la-grippe-n-en-passez-pas-les-maux.html)

METHODS: Study design, participants, interventions, and outcomes

Study/Trial Design

A factor is a variable that is controlled and varied during the course of an experiment. In the clinical trial, treatment can be a factor i.e., treatment vs. control. Factorial clinical trials are experiments that test the effect of more than one treatment using a type of design that permits an assessment of potential interactions among the treatments. The randomized controlled trial is the robust research method for determining whether a cause and effect relation exists between an intervention and an outcome. By using randomized controlled trial one can generate high-quality evidence when evaluating the effectiveness of an intervention.

This study will be a multi-armed factorial randomized controlled trial. Factorial experiments allow for the estimation of main effects of multiple factors in a single experiment by combining experimental conditions (Gallis et al. 2019). Thus, such designs are useful options in health behaviour research to efficiently compare more than one intervention. In this study, our primary analysis will be no education about community immunity or herd immunity (control) versus our developed visualization (herdimm).

Each participant will be randomly allocated to focus on one of four possible vaccine-preventable diseases that is: measles, pertussis, flu (influenza), or a generic vaccine-preventable disease (hereby referred to as "generic"). All aspects of the study (visualization, questionnaire) will be focused on that one disease for that participant.

For secondary analyses, we will compare effects of disease-specific interventions (e.g., the measles version of the visualization) to the effects of disease-matched interventions (comparators) drawn from existing visualizations available online. Specifically, we will compare our measles visualization to other measles visualizations, and so on. The comparators are:

- Measles: Guardian visualization (https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated)
- Measles: Gif by theotheredmund: (https://imgur.com/gallery/8M7q8#J7LANQ4)
- Generic: Robert Koch Institut visualization (http://rocs.hu-berlin.de/D3/herd/)
- Generic: SBS News visualization: (https://www.sbs.com.au/news/two-sydney-babies-too-young-to-be-vaccinated-infected-with-measles?fbclid=IwAR1Rns_ycauiZeTeY1HC22eJe8JE6dwcQG0A9fgq305WCskIHBT_AGzExLE)
- Generic: NHS visualization (https://www.nhs.uk/Video/Pages/Vaccinationanimation2.aspx)
- Influenza: Public Health Agency of Canada Flu: Don't Pass It On! (https://www.canada.ca/en/public-health/services/video/the-flu-don-t-pass-it-on.html)

The primary and secondary outcomes of the study which are risk perception, emotion, knowledge,trust (in information provided), attitudes and a validated scale about vaccination (Betsch et al. 2018) will be assessed in an online questionnaire.

Because three online visualizations are available only in English whereas our study will be conducted in both English and French, we will use slightly different randomization patterns for English- and French-speaking participants.

MÉTHODES: Conception de l'étude, participants, interventions et résultats

Conception de l'étude

Un facteur est une variable contrôlée et variée au cours d'une expérience. Dans l'essai clinique, le traitement peut être un facteur, par exemple traitement versus contrôle. Les essais cliniques factoriels sont des expériences qui testent l'effet de plus d'un traitement en utilisant un type de conception permettant d'évaluer les interactions potentielles entre les traitements. L'essai contrôlé randomisé est la méthode de recherche robuste pour déterminer s'il existe une relation de cause à effet entre une intervention et un résultat. En utilisant des essais contrôlés randomisés, on peut générer des preuves de haute qualité lors de l'évaluation de l'efficacité d'une intervention.

Cette étude sera un essai contrôlé randomisé factoriel multi-bras. Les expériences factorielles permettent d'estimer les effets principaux de plusieurs facteurs dans une seule expérience en combinant des conditions expérimentales. Ainsi, de tels modèles sont des options utiles dans la recherche sur les comportements en matière de santé pour comparer efficacement plus d'une intervention. Dans cette étude, notre analyse principale portera sur l'absence d'éducation concernant l'immunité de la communauté ou l'immunité collective (contrôle) par rapport à notre visualisation développée (herdimm).

Chaque participant se verra attribuer au hasard une des quatre maladies pouvant être prévenues par un vaccin aux fins de l'étude : rougeole, coqueluche, grippe ou une maladie générique évitable par un vaccin (ci-après dénommée « générique »). Tous les aspects de l'étude (visualisation, questionnaire) seront centrés sur cette maladie pour ce participant.

Pour les analyses secondaires, nous comparerons les effets des interventions spécifiques à une maladie (par exemple, la version de la visualisation portant sur la rougeole) avec les effets des interventions appariées à la maladie (comparateurs) tirés des visualisations existantes disponibles en ligne. Plus précisément, nous comparerons notre visualisation de la rougeole à d'autres visualisations de la rougeole, et ainsi de suite. Les comparateurs sont :

- Rougeole : visualisation de The Guardian (https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated)
- Rougeole : Visualisation par theotheredmund: (https://imgur.com/gallery/8M7q8#J7LANQ4)
- Générique : Visualisation de l'Institut Robert Koch (http://rocs.hu-berlin.de/D3/herd/)
- Générique : Visualisation de SBS (The Special Broadcasting Service) : (https://www.sbs.com.au/news/two-sydney-babies-too-young-to-be-vaccinated-infected-with-

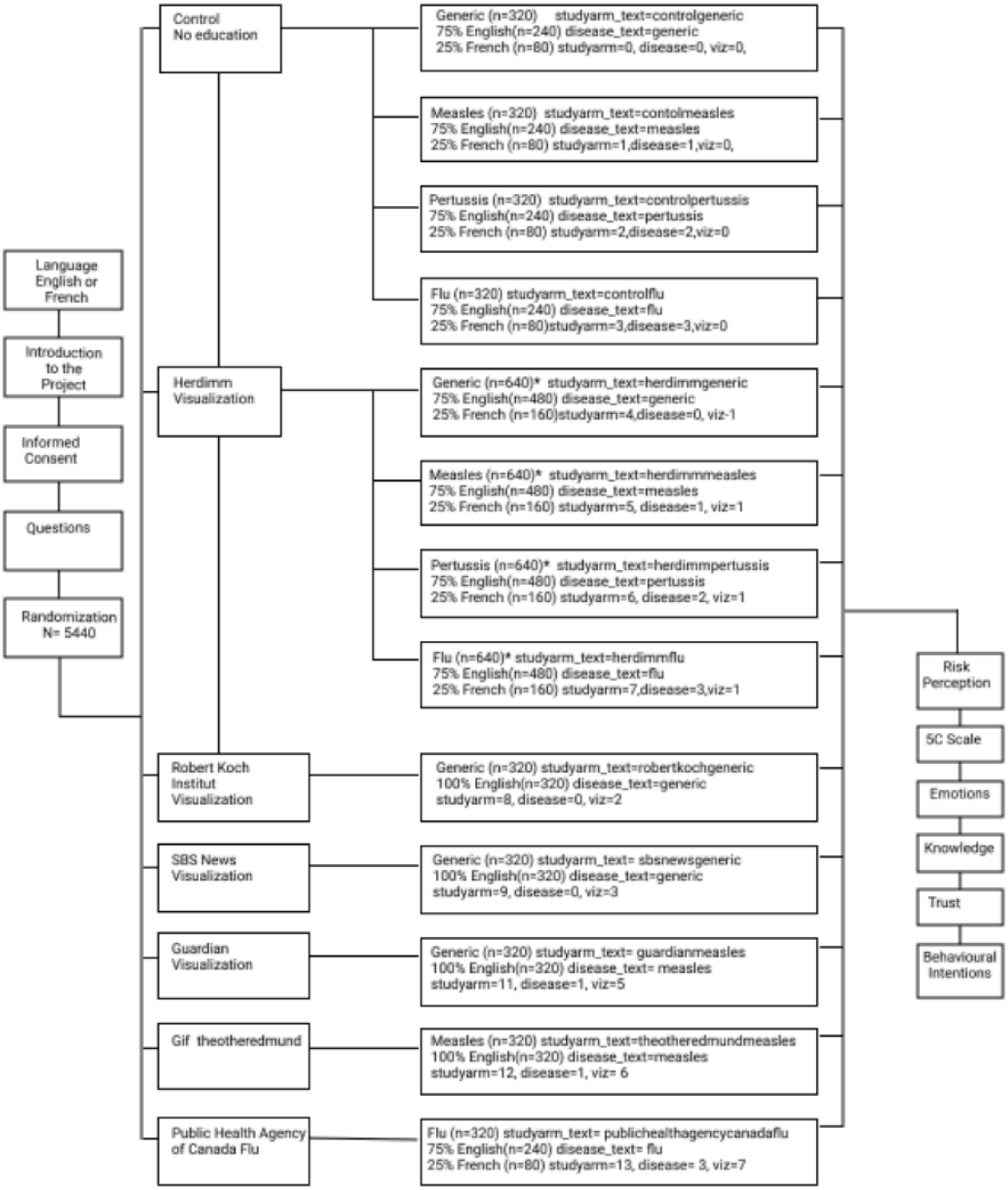
measles?fbclid=IwAR1Rns_ycauiZeTeY1HC22eJe8JE6dwcQG0A9fgq305WCskIHBT_AGzExLE)

- Générique: visualisation NHS (National Health Service) (https://www.nhs.uk/Video/Pages/Vaccinationanimation2.aspx)
- Influenza: Agence de la santé publique du Canada Grippe: ne le transmettez pas! (https://www.canada.ca/fr/sante-publique/services/video/la-grippe-n-en-passez-pas-les-maux.html)

Les résultats primaires et secondaires de l'étude (perception du risque, émotion, connaissances et une échelle validée concernant la vaccination (Betsch et al. 2018) seront évalués à l'aide d'un questionnaire en ligne.

Puisque trois visualisations en ligne sont disponibles uniquement en anglais alors que notre étude sera menée en anglais et en français, nous utiliserons des schémas de randomisation légèrement différents pour les participants anglophones et francophones.

Flow chart: Overall randomization



*** Sample size doubled for an embedded study of male versus female voice narration**

Study setting

- The online randomized controlled trial will be running in Qualtrics, an online survey software.

- Eligible participants will be adult members (18 years and over) of the general public in Canada (with a minimum of 25% of the sample with French as their preferred language compared to English).

Cadre de l'étude

- L'essai contrôlé randomisé en ligne sera exécuté dans Qualtrics, un logiciel de sondage en ligne.
- Les participants admissibles seront des membres adultes (de 18 ans et plus) du grand public au Canada (au moins 25% de l'échantillon ayant le français comme langue préférée par rapport à l'anglais).

Eligibility criteria

Inclusion Criteria:

- Members of the general population
- Be at least 18 years old
- To be able to provide free and informed consent and to be able to read and understand French or English
- Do not have vision problems not corrected
- To be able to use computer

Exclusion Criteria:

- Be under 18 years old
- Not being able to provide free and informed consent
- Have difficulty understanding and answering questions in French or English
- Have uncorrected vision problems

Critères d'éligibilité

Critères d'inclusion :

- Membres de la population générale
- Avoir au moins 18 ans
- Être capable de donner son consentement libre et éclairé et savoir lire et comprendre le français ou l'anglais
- Ne pas avoir des problèmes de vision non corrigés
- Être en mesure d'utiliser un ordinateur

Critères d'exclusion :

- Avoir moins de 18 ans
- Être dans l'incapacité de donner un consentement libre et éclairé
- Avoir de la difficulté à comprendre et à répondre aux questions en français ou en anglais
- Avoir des problèmes de vision non corrigés

Interventions/ Interventions

- Explained above (in section study design)
- Expliqué ci-dessus (dans la section conception de l'étude)

Incentive / Incitatif

We will use paid panels that offer a small incentive (typically $1-1.50) to members of panels to complete surveys.

Nous aurons recours à des panels rémunérés offrant un petit incitatif (généralement de 1 à 50 USD) aux membres des panels pour leur permettre de répondre aux sondages.

Outcomes / Résultats

In this study we are evaluating the effects of visualization on risk perception (for the individual, family, community, and vulnerable people in communities) (1st outcome) and on emotions, attitudes, knowledge, and behavioural intentions (2nd outcome) with trust as a covariate.

Dans cette étude, nous évaluons les effets de la visualisation sur la perception du risque (pour l'individu, la famille, la communauté et les personnes vulnérables dans les communautés) (1er résultat) et sur les émotions, les attitudes, les Connaissances et les intentions comportementales (2^{e} résultat) en utilisant la confiance comme covariable.

Participant timeline / Calendrier

Following ethics approval from Laval University, we will collect data online from **December 2020 to February 2021**.

Suite à l'approbation éthique de l'Université Laval, nous collecterons les données en ligne du décembre 2020 au février 2021.

Sample size / Taille de l’échantillon

To determine the sample size for this between-subjects design, we assumed a small effect size (Cohen’s f=0.10) and a statistical test power of 80% and 5% type 1 error, resulting in a required N of 320 participants for each arm with 75% of English and 25% of French speaking respondents which were sufficient to run our two-way analysis of variance. We estimated the sample size by using g power (Universität Düsseldorf: gpower) for continuous outcome.

Pour déterminer la taille de l'échantillon pour ce plan inter-sujets, nous avons supposé une petite taille d'effet (f=0,10 de Cohen) et une puissance de test statistique de 80% et 5% d'erreur de type 1, ce qui a donné un N requis de 320 participants pour chaque bras avec 75% de répondants anglophones et 25% de répondants francophones qui étaient suffisants pour effectuer notre analyse de variance à deux voies. Nous avons estimé la taille de l'échantillon en utilisant la puissance (Universität Düsseldorf: gpower) pour un résultat continu.

Recruitment / Recrutement

- The online randomized controlled trial will be run in Qualtrics, an online survey software.
- Eligible participants will be adult members of the general public in Canada (with a minimum of 25% of the sample with French as their preferred language compared to English).
- Participants will be recruited using established survey panels (subcontracted in Qualtrics).
- L'essai contrôlé randomisé en ligne sera exécuté dans Qualtrics, un logiciel de sondage en ligne.
- Les participants admissibles seront des membres adultes du grand public au Canada (au moins 25% de l'échantillon ayant le français comme langue préférée par rapport à l'anglais).
- Les participants seront recrutés sur des panels d’enquêtes établis (sous-traités dans Qualtrics)

METHODS: Assignment of interventions (for controlled trials)/MÉTHODES : Attribution des interventions (pour les essais contrôlés)

Allocation/Allocation :

Sequence generation/Génération de séquences

This study will be simple randomized (computer generated randomization) to eliminate selection bias, balancing both known and unknown prognostic factors, in the assignment of interventions.

Cette étude utilisera une randomisation simple(randomisation générée par ordinateur) pour éliminer le biais de sélection, en équilibrant les facteurs pronostiques connus et inconnus, dans l'attribution des interventions.

Allocation concealment mechanism/Mécanisme de dissimulation des attributions

This will be a single-blinded study because we cannot mask participants to the fact that they have been randomized to a visualization e.g., about measles. However, participants will not necessarily know the purpose of the study arm to which they are assigned, and investigators will be blinded to study arm during data analysis (investigators are not aware of whether the next eligible participant will be receiving treatment or control intervention).

Les participants et les investigateurs seront aveugles (à double insu) quant au processus d'attribution (les investigateurs ou les participants ne savent pas si le prochain participant éligible se verra attribué un traitement ou une intervention de contrôle).

Implementation/ Implémentation

Qualtrics (computerized survey software) will use computerized randomization to automatically assign participants to interventions.

Qualtrics (logiciel d'enquête informatisé) utilisera la randomisation informatisée pour affecter automatiquement les participants aux interventions.

Blinding (masking)/ Aveuglement (masquage)

This will be a single-blinded study because we cannot mask participants to the fact that they have been randomized to a visualization e.g., about measles. However, participants will not necessarily know the purpose of the study arm to which they are assigned, and investigators will be blinded to study arm during data analysis.

Ce sera une étude à simple insu car nous ne pouvons pas masquer aux participants le fait qu'ils ont été randomisés pour obtenir une visualisation p. ex., de la rougeole. Cependant, les participants ne connaîtront pas nécessairement le but du groupe d'étude auquel ils sont affectés et les investigateurs seront aveuglés pour étudier le groupe lors de l'analyse des données.

METHODS: Data collection, management, and analysis/MÉTHODES : collecte, traitement et analyse des données

Data collection methods/Méthodes de collecte des données

After seeking eligible participants' consent, participants will be randomized into different arms of the study and the effects of visualization on risk perception (for the individual, family, community, and vulnerable people in communities) (primary outcome) and on emotions, attitudes, knowledge, and behavioural intentions (secondary outcome) will be measured by questionnaire.

Après avoir recherché le consentement des participants éligibles, les participants seront randomisés dans différentes parties de l'étude et sur les effets de la visualisation sur la perception du risque (pour l'individu, la famille, la communauté et les personnes vulnérables dans les communautés) (résultat principal) et sur les émotions, les attitudes, les Connaissances et les intentions comportementales (résultat secondaire) seront mesurées par questionnaire.

Data management/Traitement des données

All the information we collect will be confidential and used for research purposes only. In other words:

- Participant credentials will not be associated with the results of the study and no individual information will be presented in reports, publications or presentations.
- All data will be presented in aggregate form without individual identifiers.
- Data will be stored on Qualtrics servers located in Canada. When we work with data, the only people who will have access to our Qualtrics account will be the investigators and our team members who have complete and relevant ethics training. When the data is stored on our computers, each of our computers will be protected by a password. The data will be stored on a secure server, to which access is maintained and reserved for the members of the team (members of the team affiliated with Université Laval).

After the end of the study:

- Anonymized data (answers to questions in the survey including socio-demographic information) will be deposited in a public repository (Dataverse de l'Université Laval (Laval University)) which will allow data sharing with the scientific community. No information that would allow anyone to identify a person will be deposited in this public repository.
- Any other electronic data will be destroyed in June 2027.

Toutes les informations que nous recueillons seront confidentielles et utilisées uniquement à des fins de recherche. En d'autres termes:

- Les informations d'identification des participants ne seront pas associées aux résultats de l'étude et aucune information individuelle ne sera présentée dans des rapports, des publications ou des présentations.
- Toutes les données seront présentées sous forme agrégée, sans identificateurs individuels.
- Les données seront enregistrées sur les serveurs de Qualtrics situés au Canada. Lorsque nous travaillons avec des données, les seules personnes qui auront accès à notre compte Qualtrics seront les enquêteurs et les membres de notre équipe qui ont toute la formation éthique complète et pertinente. Lorsque les données seront stockées sur nos ordinateurs, chacun de nos ordinateurs sera protégé par un mot de passe. Les données seront stockées sur un serveur sécurisé, auquel l'accès est maintenu et réservé aux membres de l'équipe (membres de l'équipe affiliés à l'Université Laval).

Après la fin de l'étude :

- Les données anonymisées (réponses aux questions de l'enquête y inclus les informations sociodémographiques) seront déposées dans un dépôt public (Dataverse de l'Université Laval) qui permettra le partage des données avec la communauté scientifique. Aucune information qui permettrait d'identifier une personne ne sera déposée dans ce dépôt public.
- Toute autre donnée électronique sera détruite en juin 2027.

Statistical methods/Méthodes statistiques

The type of comparison between intervention and control groups will be analyzed and reported as a superiority trial. The comparison will be done by using a linear regression for statistical analysis of continuous data.

Le type de comparaison entre le groupe d'intervention et le groupe de contrôle sera analysé et présenté comme un essai de supériorité. La comparaison sera effectuée en utilisant une régression linéaire pour l'analyse statistique des données continues.

Variables/ outcomes	Visualization (Herdimm)	Control 1(no visualization/ no education about community immunity)
Risk perception		
Emotions		
Attitudes		
Knowledge		
Behavioural intentions		
Trust		

Variables/ outcomes	Visualization (Herdimm)	Measles: Guardian visualization	Measles: theotheredmund visualization	Influenza: Public Health Agency of Canada Flu: Don’t Pass It On!
Risk perception				
Emotions				
Attitudes				
knowledge				
Behavioural intentions				
Trust				

Variables/ outcomes	Visualization (Herdimm)	Generic: Robert Koch Institut visualization	Generic: SBS News visualization	Generic: NHS visualization
Risk perception				
Emotions				
Attitudes				
Knowledge				
Behavioural intentions				
Trust				

ETHICS AND DISSEMINATION/ÉTHIQUE ET DIFFUSION

Research ethics approval/Approbation de l'éthique de la recherche

This study is approved from the Laval University ethics committee (Approval No. 2017-137 Phase II A-6/ 10-09-2020).

Approbation de l'éthique de la recherche Cette étude est approuvée par le comité d'éthique de l'Université Laval (Approbation n° 2017-137 Phase II A-6/ 10-09-2020).

Consent/Consentement

The free and informed consent will be seeked before the start of the study. Participants will be able to withdraw from the study at anytime if they don't want to take part in the study. Participants who withdrew from the study in the middle and has not completed the study their data won't be analysed.

Le consentement libre et éclairé sera recherché avant le début de l'étude. Les participants pourront se retirer de l'étude à tout moment s'ils ne veulent pas participer à l'étude. Les participants qui se sont retirés de l'étude au milieu et qui n'ont pas terminé l'étude ne seront pas analysés.

Confidentiality/Confidentialité

- Explained in the **data management** section above.
- Expliqué dans la section sur le **traitement des données** ci-dessus.

Declaration of interests/Déclaration d'intérêts

- No competing interests.
- Pas de conflits d'intérêts.

Access to data/Accès aux données

The data will be kept in the principal investigator's laboratory located on the campus of Laval University and only members of the research team will have access to data.

The data will be stored on Qualtrics Canadian servers, located in Canada and subject to Canadian data protection laws. When we work with data the only people who will have access to our Qualtrics account will be investigators and our team members who have all complete and relevant ethics training. When the data will be held on our computers each of our computers will be protected by the password. The data will be stored on a secure server, to which access is maintained and restricted to team members (team members affiliated at Laval university).

Les données seront conservées dans le laboratoire du chercheur principal situé sur le campus de l'Université Laval et seuls les membres de l'équipe de recherche auront accès aux données.

Les données seront stockées sur des serveurs canadiens de Qualtrics, situés au Canada. Lorsque nous travaillons avec des données, les seules personnes qui auront accès à notre compte Qualtrics seront les enquêteurs et les membres de notre équipe qui ont toute la formation éthique complète et pertinente. Lorsque les données seront stockées sur nos ordinateurs, chacun de nos ordinateurs sera protégé par un mot de passe. Les données seront stockées sur un serveur sécurisé, auquel l'accès est maintenu et réservé aux membres de l'équipe (membres de l'équipe affiliés à l'université Laval).

Dissemination policy/Politique de diffusion

We intend to publish our results in a peer reviewed journal for healthcare professionals, the public, and other relevant groups to access.

Nous avons l'intention de publier nos résultats dans une revue évaluée par des pairs, à laquelle les professionnels de la santé, le public et d'autres groupes concernés pourront accéder.

Factors:

Factors:	**Facteurs :**
Visualization (Herdimm)	Visualization (Herdimm)
Control 1(no visualization/ no education about community immunity)	Contrôle 1 (pas de visualisation / pas d'éducation sur l'immunité de la communauté)
Control 2 (Guardian visualization)	Contrôle 2 (visualization Guardian)
Control 3(Robert Koch Institut visualization)	Contrôle 3 (visualization Robert Koch Institut)
Control 4(theotheredmund visualization)	Contrôle 4 (visualization theotheredmund)
Control 5 (SBS news visualization)	Contrôle 5 (SBS news visualization)
~~Control 6 (NHS visualization)~~	~~Contrôle 6 (NHS visualization)~~
Control 7 (Flu: Public Health Agency of Canada Flu: Don't Pass It On!)	Contrôle 7 (Flu: Public Health Agency of Canada Flu: Don't Pass It On!)

Measures

Measures:	**Mesures :**
Risk perception (primary outcome)	Perception du risque (résultat principal)
Emotions (secondary outcome)	Émotions (résultat secondaire)
Attitudes (secondary outcome)	Attitudes (résultat secondaire)
Knowledge (secondary outcome)	Connaissances (résultat secondaire)
Behavioural intentions (secondary outcome)	Intentions comportementales (résultat secondaire)
Trust (Covariates)	Confiance (covariables)

Randomized variables:

Randomized variables:		**Valeurs randomisées :**
visualization	0= No education 1 = Herdimm 2 = Guardian visualization 3 = Robert Koch Institut visualization 4= theotheredmund visualization 5=SBS news visualization ~~6= NHS visualization~~ 7= Flu: Public Health Agency of Canada	0= No education 1 = Herdimm 2 = visualization Guardian 3 = visualization Robert Koch Institut 4= visualization theotheredmund 5= SBS news visualization ~~6= NHS visualization~~ 7= Flu: Public Health Agency of Canada

References:

Betsch, Cornelia, Robert Böhm, Lars Korn, and Cindy Holtmann. 2017. “On the Benefits of Explaining Herd Immunity in Vaccine Advocacy.” *Nature Human Behaviour* 1 (3): s41562–017 – 0056.

Betsch, Cornelia, Philipp Schmid, Dorothee Heinemeier, Lars Korn, Cindy Holtmann, and Robert Böhm. 2018. “Beyond Confidence: Development of a Measure Assessing the 5C Psychological Antecedents of Vaccination.” *PloS One* 13 (12): e0208601.

Gallis, John A., Gary G. Bennett, Dori M. Steinberg, Sandy Askew, and Elizabeth L. Turner. 2019. “Randomization Procedures for Multicomponent Behavioral Intervention Factorial

Trials in the Multiphase Optimization Strategy Framework: Challenges and Recommendations." *Translational Behavioral Medicine*. https://doi.org/10.1093/tbm/iby131.

Hakim, Hina, Thierry Provencher, Christine T. Chambers, S. Michelle Driedger, Eve Dube, Teresa Gavaruzzi, Anik M. C. Giguere, et al. 2018. "Interventions to Help People Understand Community Immunity: A Systematic Review." *Vaccine*, December. https://doi.org/10.1016/j.vaccine.2018.11.016.

Healey, Christopher G., and James T. Enns. 2012. "Attention and Visual Memory in Visualization and Computer Graphics." *IEEE Transactions on Visualization and Computer Graphics* 18 (7): 1170–88.

"Universität Düsseldorf: Gpower." n.d. Accessed July 28, 2020. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html.

S2: Study questionnaire

contag_phase3b

Start of Block: Introduction

Intro **Project Description**

Project Description

Introduction
Thank you for participating in this study. This project is a collaboration with University Laval, Public Health Ontario (PHO) and The National Institute of Public Health of Quebec (INSPQ).

In this study, we will first give you some information to do with contagious diseases and then ask you to answer some questions.

The survey will take about 15 to 20 minutes. Your name will not be recorded anywhere in this survey. Researchers will not be able to associate your name with your answers. If you choose to leave the study, you can stop at any time. You may choose not to answer any questions you don't want to answer. If, for some reason, you cannot complete the survey, you may **restart** the survey at a later time by clicking on the survey link in your email invitation. You will have to **restart** the survey from the beginning.

If you become worried about your health while taking this survey, please talk to a healthcare provider.

About the researchers
This study is a PhD project conducted by Ms. Hina Hakim, PhD student, supervised by Holly O. Witteman, PhD, professor and researcher in the Faculty of Medicine at Laval University. Ms. Hakim is co-supervised by Dr. Daniel Reinharz, professor and researcher in the Faculty of Medicine at Laval University.

Purpose of the study
The purpose of this study is to evaluate the effects of information to do with contagious diseases on risk perception, knowledge, emotions, attitudes, beliefs, and behavioural intentions.

What your participation entails
Participating in this study will involve following:

- Answering survey questions.
- You may be asked to view or listen to information, which will require either the ability to read or hear.

- You may be asked to briefly visit another website in the middle the survey.

Risks and benefits

The risks associated with participating in the study, specifically:

- Information and questions, we ask may make you feel uncomfortable or anxious. We suggest you to consult your healthcare professional, as the case may be;
- You may feel some discomfort having to focus on a computer screen.
- You may feel some fatigue from concentrating throughout the survey.

There is no personal benefit to participate in this study, however, the information we gather will help us to evaluate the effects of the information, which may help in improving the way health information is shared, and population health.

Your help means a lot to us. We thank you again for taking the time to complete this survey.

Confidentiality and data protection

All information we collect will be confidential and used only for research purposes.
In other words:

- Participant credentials will not be associated with the results of the study and no individual information will be presented in reports, publications or presentations.
- All data will be presented in aggregate form without individual identifiers.
- Data will be stored on Qualtrics servers located in Canada. When we work with data, the only people who will have access to our Qualtrics account will be the investigators and our team members who have complete and relevant ethics training. When the data is stored on our computers, each of our computers will be protected by a password. The data will be stored on a secure server, to which access is maintained and reserved for the members of the team (members of the team affiliated with Université Laval).

After the end of the study:

- Anonymized data (answers to questions in the survey including socio-demographic information) will be deposited in a public repository (Dataverse de l'Université Laval (Laval University)) which will allow data sharing with the scientific community. No information that would allow anyone to identify a person will be deposited in this public repository.
- Any other electronic data will be destroyed in June 2027.

Contact people

If you have any questions about the research study, or if you experience a problem as a result of participating in the study, please contact:

Holly Witteman
Laval University
Tel: (418) 656-2131 ext: 403981
Email: holly.witteman@fmed.ulaval.ca

If you have concerns about your rights as a participant in this study or any complaints, you can communicate with the University Laval Ombudsman at:

Pavillon Alphonse-Desjardins
2325, rue de l'Université, Local 3320
Québec, QC G1V 0A6
Tel: 418-656-3081
Toll free: 1-866-323-2271
Email: info@ombudsman.ulaval.ca

This project has been approved by Laval University Research Ethics Board: Approval No. 2017-137 Phase II A-6/ 10-09-2020

Description du projet

Introduction

Merci de participer à cette étude. Ce projet est réalisé en collaboration avec l'Université Laval, Santé publique Ontario (SPO) et l'Institut national de la santé publique du Québec (INSPQ).

Dans cette étude, nous vous donnerons d'abord quelques informations sur les maladies contagieuses et nous vous demanderons ensuite de répondre à quelques questions.

L'enquête durera environ 15 à 20 minutes. Votre nom ne sera enregistré nulle part dans cette enquête. Les chercheurs et les chercheuses ne pourront pas associer votre nom à vos réponses. Si vous choisissez de quitter l'étude, vous pouvez y mettre fin à tout moment. Vous pouvez choisir d'ignorer les questions auxquelles vous ne voulez pas répondre. Si, pour une raison quelconque, vous ne pouvez pas compléter le sondage, vous pourrez le recommencer ultérieurement en cliquant sur le lien du sondage dans votre invitation par courrier électronique. Vous devrez alors **recommencer** le sondage depuis le début.

Si vous ressentez de l'inquiétude pour votre santé en répondant à cette enquête, parlez-en à un.e professionnel.le de la santé.

À propos des chercheurs
Cette étude est un projet de doctorat conduit par Mme Hina Hakim, candidate au doctorat en santé communautaire, supervisée par Mme Holly Witteman, Ph.D., professeure et chercheuse à la Faculté de médecine de l'Université Laval. Mme Hakim est co-dirigée par le Dr Daniel Reinharz, professeur et chercheur à la Faculté de médecine de l'Université Laval.

Aperçu de l'étude
Le but de cette étude est d'évaluer l'effet de certaines informations relatives aux maladies contagieuses sur la perception du risque, la connaissance, les émotions, les attitudes, les croyances et les intentions comportementales.

Ce que votre participation implique
Participer à cette étude impliquera de :

- Répondre aux questions du sondage
- Il se peut qu'on vous demande de lire ou d'écouter des informations, ce qui nécessiterait la capacité de voir ou d'entendre
- Il se peut que l'on vous demande de visiter brièvement un autre site web au milieu de l'enquête.

Risques et avantages
Il existe certains risques associés à la participation à l'étude, notamment :

- Les informations et questions posées pourraient vous rendre mal à l'aise ou anxieux. Nous vous suggérons de consulter votre professionnel.le de la santé;
- Vous pouvez ressentir un certain inconfort à fixer un écran d'ordinateur.
- Vous pourriez ressentir de la fatigue due à la concentration demandée tout au long du sondage.

Il n'y a aucun avantage personnel à participer à cette étude. Toutefois, les données recueillies nous aideront à évaluer les effets des informations transmises, ce qui pourrait aider à améliorer le partage des informations sur la santé ainsi que la santé de la population.

Votre aide compte beaucoup pour nous. Nous vous remercions encore de prendre le temps de répondre à ce sondage.

Confidentialité et protection des données
Toutes les informations que nous recueillons seront confidentielles et utilisées uniquement à des fins de recherche.

En d'autres termes:

- Les informations d'identification des participant.e.s ne seront pas associées aux résultats de l'étude et aucune information individuelle ne sera présentée dans des rapports, des publications ou des présentations.
- Toutes les données seront présentées sous forme agrégée, sans identificateurs individuels.
- Les données seront enregistrées sur les serveurs de Qualtrics situés au Canada. Lorsque nous travaillons avec des données, les seules personnes qui auront accès à notre compte Qualtrics seront les enquêteurs et les membres de notre équipe qui ont toute la formation éthique complète et pertinente. Lorsque les données seront stockées sur nos ordinateurs, chacun de nos ordinateurs sera protégé par un mot de passe. Les données seront stockées sur un serveur sécurisé, auquel l'accès est maintenu et réservé aux membres de l'équipe (membres de l'équipe affiliés à l'Université Laval).

Après la fin de l'étude :

- Les données anonymisées (réponses aux questions de l'enquête y inclus les informations sociodémographiques) seront déposées dans un dépôt public (Dataverse de l'Université Laval) qui permettra le partage des données avec la communauté scientifique. Aucune information qui permettrait d'identifier une personne ne sera déposée dans ce dépôt public.
- Toute autre donnée électronique sera détruite en juin 2027.

Contact
Si vous avez des questions sur le projet de recherche, ou si vous rencontrez un problème en participant à l'étude, veuillez contacter:

Prof. Holly Witteman
Université Laval
Téléphone: (418) 656-2131 ext: 403981
Courriel: holly.witteman@fmed.ulaval.ca

Si vous avez des plaintes ou des préoccupations au sujet de vos droits en tant que participant à cette étude, vous pouvez communiquer avec l'Ombudsman de l'Université Laval à:

Pavillon Alphonse-Desjardins
2325, rue de l'Université, Local 3320
Québec, QC G1V 0A6
Téléphone: 418-656-3081
Sans frais: 1-866-323-2271
Courriel: info@ombudsman.ulaval.ca

« Ce projet a été approuvé par le Comité d'éthique de la recherche de l'Université Laval : No d'approbation 2017-137 Phase II A-6/ 10-09-2020 »

---

t_intro Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_intro Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Introduction

Start of Block: Consent for participation:

X→

Consent

1. I have been informed about the nature and purpose of this study and the research procedures.
2. I was informed in English which is a language I understand fluently.
3. I understand that my participation in this study is voluntary and that I can withdraw at any time without penalty or consequences of any kind.
4. I understand that the information collected is confidential and will only be used for research purposes.
5. I have read the project description and informed consent form and I voluntarily agree to participate in this study.

○ Yes I Consent (1)

○ No I do not Consent (0)

Consent

1. J'ai été informé.e de la nature et du but de cette étude et des procédures de recherche.
2. J'ai été informé.e en français, une langue que je comprends et parle couramment.
3. Je comprends que ma participation à cette étude est volontaire et que je peux me retirer à tout moment sans pénalité ni conséquence d'aucune sorte.

4. Je comprends que les informations recueillies sont confidentielles et ne seront utilisées qu'à des fins de recherche.
5. J'ai lu la description du projet et le formulaire de consentement éclairé et j'accepte volontairement de participer à cette étude.

○ Oui, je consens (1)

○ Non, je n'y consens pas (0)

Skip To: End of Survey If I have been informed about the nature and purpose of this study and the research procedures. I wa... = No I do not Consent

t_consent Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_consent Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

End of Block: Consent for participation:

Start of Block: Standard Sociodemo Canada

yearofbirth In what year were you born?

▼ Prefer not to answer (888) ... 2010 (2010)

yearofbirth Quelle est votre année de naissance?

▼ Je préfère ne pas répondre (888) ... 2010 (2010)

---

bornincanada Were you born in Canada? (Check one only.)

○ Prefer not to answer (888)

○ Yes (1)

○ No - What year did you arrive in Canada? (0) ____

bornincanada Êtes-vous né(e) au Canada? (Cochez une seule réponse)

○ Je préfère ne pas répondre (888)

○ Oui (1)

○ Non - en quelle année êtes-vous arrivé(e) au Canada? (0) ____

Display This Question:

If If False

hp Were you born outside North America? (Check one only.)

○ Prefer not to answer (888)

○ Yes (1)

○ No - What year did you arrive in North America? (0) ____

hp Êtes-vous né(e) hors de l'Amérique du Nord? (Cochez une seule réponse)

○ Je préfère ne pas répondre (888)

○ Oui (1)

○ Non - en quelle année êtes-vous arrivé(e) au Canada? (0) ____

language What language would you feel most comfortable speaking in with your healthcare provider? (doctor, nurse, dentist, pharmacist)? (Check all that apply.)

☐ Prefer not to answer (888)

☐ □ English (1)

☐ □ French (2)

☐ □ Other, please specify: (3) ____

language Dans quelle(s) langue(s) êtes-vous à l’aise pour parler avec un.e professionnel.le de la santé (médecin, infirmière, dentiste, pharmacien)? (Cochez tout ce qui s’applique)

☐ Je préfère ne pas répondre (888)

☐ Anglais (1)

☐ Français (2)

☐ Autre, précisez : (3) ____

ethnicity Which of the following best describes your racial or ethnic group(s)? (Check all that apply.)

☐ Prefer not to answer (888)

☐ Asian - East (e.g., Chinese, Japanese, Korean) (1)

☐ Asian - Central (e.g., Kazakhstani, Uzbekistani) (2)

☐ Asian - South (e.g., Indian, Pakistani, Sri Lankan) (3)

☐ Asian - South-East (e.g., Malaysian, Filipino, Vietnamese) (4)

☐ Black - African (e.g., Ghanaian, Kenyan, Somali) (13)

☐ Black - Caribbean (e.g., Barbadian, Jamaican) (14)

☐ Black - North American (e.g., Canadian, American) (15)

☐ First Nations (16)

☐ Indigenous or aboriginal person from outside of North America (e.g., Maori, Quechua) (5)

☐ Inuit (8)

☐ Latin American (e.g., Chilean, Mexican, Salvadorian) (9)

☐ Metis (Metis Nation in Canada) (10)

☐ Middle Eastern (e.g., Egyptian, Iranian, Lebanese) (12)

☐ North African (e.g., Moroccan, Tunisian) (11)

☐ White - European (e.g., English, Italian, Portuguese, Russian) (6)

☐ White - North American (e.g., Canadian, American) (7)

☐ Other, please specify: (17) ____

ethnicity Laquelle ou lesquelles des catégories suivantes représente(nt) vos groupes raciaux ou ethniques? (Cochez tout ce qui s'applique.)

☐ Je préfère ne pas répondre (888)

☐ Asiatique de l'est (p.ex. : Chinois, Japonais, Coréen) (1)

☐ Asiatique du centre (p.ex. : Kazakh, Ouzbek) (2)

☐ Asiatique du sud (p.ex.: Indien, Pakistanais, Sri Lankais) (3)

☐ Asiatique du sud-est (p.ex.: Malaisien, Philippin, Vietnamien) (4)

☐ Noir Africain (p.ex. : Ghanéen, Kenyan, Somalien) (13)

☐ Noir Caribéen (p.ex. : Barbadien, Jamaïcain) (14)

☐ Noir Nord-américain (p.ex.: Canadien, États-Unien) (15)

☐ Premières Nations (16)

☐ Autochtone ou aborigène de hors de l'Amérique du Nord (p.ex.: Maori, Quechua) (5)

☐ Inuit (8)

☐ Latino-américain (p.ex. : Chilien, Mexicain, Salvadorien) (9)

☐ Métis (Nation métisse du Canada) (10)

☐ Moyen-oriental (p.ex. : Égyptien, Iranien, Libanais) (12)

☐ Maghrébin (p.ex. : Marocain, Tunisien) (11)

☐ Blanc Européen (p.ex.: Français, Anglais, Italien, Portugais, Russe) (6)

☐ Blanc Nord-Américain (p.ex.: Canadien, États-Unien) (7)

☐ Autre, précisez : (17) ____

disability Do you consider yourself to have a disability (e.g, physical disability, sensory disability (e.g., hearing or vision loss), chronic illness, learning disability, mental illness, developmental disability or other)? (Check one only.)

○ □ Prefer not to answer (888)

○ □ No disability (0)

○ □ Yes, at least one disability (1)

disability Considérez-vous avoir une limitation ou un handicap (p.ex. : limitation ou déficience physique, limitation ou déficience sensorielle (p.ex. perte de vision ou d'audition), maladie chronique, trouble d'apprentissage, trouble de santé mentale, trouble du développement ou autre)? (Cochez une seule réponse.)

○ Je préfère ne pas répondre (888)

○ Non, aucune limitation ou handicap (0)

○ Oui, au moins une limitation ou handicap (1)

techdisability Do you have a disability that limits your access to technology and computer based tools (p.ex. : computer, tablet, interactive terminal)? (Check one only.)

○ □ Prefer not to answer (888)

○ No (0)

○ Yes (1)

techdisability Selon vous, avez-vous une limitation ou un handicap qui vous limite l’accès à certains outils technologiques ou informatiques (p.ex. : ordinateur, tablette, borne interactive)? (Cochez une seule réponse.)

○ Je préfère ne pas répondre (888)

○ Non (0)

○ Oui (1)

---

sexatbirth What sex were you assigned at birth, meaning on your original birth certificate? (Check one only.)

○ Prefer not to answer (888)

○ Female (1)

○ Male (2)

sexatbirth Quel sexe vous a-t-on assigné à la naissance, sur votre certificat de naissance original? (Cochez une seule réponse.)

○ Je préfère ne pas répondre (888)

○ Féminin (1)

○ Masculin (2)

---

genderidentity Do you consider yourself: (Check one only.)

○ Female (1)

○ Male (2)

○ Indigenous or other cultural gender minority identity (e.g., two-spirit) (3)

○ Something else (e.g. non-binary, gender fluid) (4)

○ Prefer not to answer (888)

genderidentity Vous considérez-vous comme : (Cochez une seule réponse.)

○ Une femme (1)

○ Un homme (2)

○ Un(e) minorité de genre culturelle, autochtone ou autre (p.ex. : bispiritualité, « *two-spirit* ») (3)

○ Autre (p.ex. : non-binaire, « *gender fluid* ») (4)

○ Je préfère ne pas répondre (888)

---

income What was your total family income before taxes last year? (Check one only.)

○ □ Prefer not to answer (888)

○ □ 24 999 or less (1)

○ □ 25 000 to 49 999 (2)

○ □ 50 000 to 99 999 (3)

○ □ 100 000 or more (4)

○ □ Do not know (5)

income Quel était le revenu total (avant impôts) de votre ménage l’an dernier? (Cochez une seule réponse.)

○ Je préfère ne pas répondre (888)

○ 24 999 et moins (1)

○ 25 000 à 49 999 (2)

○ 50 000 à 99 999 (3)

○ 100 000 ou plus (4)

○ Je ne sais pas (5)

X→

supportedbyincome How many people does this income support?

○ Number of people: (1) ____

○ Prefer not to answer (888)

supportedbyincome Combien de personnes ce revenu supporte-t-il?

○ Nombre de personnes : (1) ____

○ Je préfère ne pas répondre (888)

X→

educationlevel What is the highest level of education you have completed? (Check one only.)

- ○ ☐ Prefer not to answer (888)
- ○ ☐ Elementary School (completed or not) (1)
- ○ ☐ High School Diploma (2)
- ○ ☐ Apprenticeship or trade certificate or diploma (3)
- ○ ☐ College or polytechnical school certificate or diploma (4)
- ○ ☐ University degree, bachelor level or below (5)
- ○ ☐ University graduate degree (Master’s level) (6)
- ○ ☐ University graduate degree (Doctorate level) (7)
- ○ ☐ Do not know (8)

educationlevel Quel est votre plus haut niveau de scolarité complété? (Cochez une seule réponse.)

- ○ Je préfère ne pas répondre (888)
- ○ Études primaires (complétées ou non) (1)
- ○ Diplôme d’études secondaires (2)
- ○ Certificat ou diplôme d'apprentissage ou de métier (3)
- ○ Attestation ou diplôme d’études professionnelles (DES, ASP, DEP) (4)
- ○ Études universitaires de premier cycle (Certificat, Baccalauréat) ou moins (5)
- ○ Études universitaires de deuxième cycle (DESS, Maîtrise) (6)
- ○ Études universitaires de troisième cycle (Doctorat) (7)
- ○ Je ne sais pas (8)

t_sociodemographic Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_sociodemographic Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Standard Sociodemo Canada

Start of Block: Individualism and Collectivism Scale - Triandis1998 - EN/FR

ih1 **Evaluate how well each of the following statements applies to you:**

I'd rather depend on myself rather than others.

- ○ Never or definitely not =1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ih1

Évaluez à quel point chacun des énoncés suivants s'applique à vous :

Je préfère dépendre de moi-même que des autres

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

ih2 I rely on myself most of the time, I rarely rely on others.

○ Never or definitely not = 1 (1)

○ 2 (2)

○ 3 (3)

○ 4 (4)

○ 5 (5)

○ 6 (6)

○ 7 (7)

○ 8 (8)

○ Always or definitely yes = 9 (9)

ih2 Je me fie à moi-même la plupart du temps, rarement aux autres.

○ Jamais ou définitivement non = 1 (1)

○ 2 (2)

○ 3 (3)

○ 4 (4)

○ 5 (5)

○ 6 (6)

○ 7 (7)

○ 8 (8)

○ Toujours ou définitivement oui = 9 (9) ih3 I often do “my own thing.”

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ih3 Je fais souvent « ma propre affaire. »

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

ih4 My personal identity, independent of others, is very important to me.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ih4 Mon identité personnelle, indépendante des autres, est très importante pour moi.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

iv1 It is important that I do my job better than others.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

iv1 Il est important pour moi de faire mon travail mieux que les autres.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

iv2 Winning is everything.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

iv2 Gagner, c’est tout ce qui compte.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

iv3 Competition is the law of nature.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

iv3 La compétition est la loi de la nature.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

iv4 When another person does better than I do, I get tense and aroused.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

iv4 Quand une autre personne fait mieux que moi, je deviens tendu et agité.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

ch1 If a coworker gets a prize, I would feel proud.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ch1 Si un collègue remportait un prix, je me sentirais fier.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

ch2 The well-being of my coworkers is important to me.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ch2 Le bien-être de mes collègues est important pour moi.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

ch3 To me, pleasure is spending time with others.

○ Never or definitely not = 1 (1)

○ 2 (2)

○ 3 (3)

○ 4 (4)

○ 5 (5)

○ 6 (6)

○ 7 (7)

○ 8 (8)

○ Always or definitely yes = 9 (9)

ch3 Pour moi, le plaisir c’est de passer du temps avec les autres.

○ Jamais ou définitivement non = 1 (1)

○ 2 (2)

○ 3 (3)

○ 4 (4)

○ 5 (5)

○ 6 (6)

○ 7 (7)

○ 8 (8)

○ Toujours ou définitivement oui = 9 (9) ch4 I feel good when I cooperate with others.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

ch4 Je me sens bien lorsque je coopère avec les autres.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

cv1 Parents and children must stay together as much as possible.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

cv1 Les parents et les enfants doivent rester ensemble autant que possible.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

cv2 It is my duty to take care of my family, even when I have to sacrifice what I want.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

cv2 C’est mon devoir de prendre soin de ma famille, même si je dois sacrifier mes propres désirs.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

cv3 Family members should stick together, no matter what sacrifices are required.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

cv3 Les membres d’une même famille devraient être solidaires peu importe les sacrifices requis.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

cv4 It is important to me that I respect the decisions made by my groups.

- ○ Never or definitely not = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Always or definitely yes = 9 (9)

cv4 Il est important pour moi de respecter les décisions prises par mes groupes.

- ○ Jamais ou définitivement non = 1 (1)
- ○ 2 (2)
- ○ 3 (3)
- ○ 4 (4)
- ○ 5 (5)
- ○ 6 (6)
- ○ 7 (7)
- ○ 8 (8)
- ○ Toujours ou définitivement oui = 9 (9)

t_ic Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_ic Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Individualism and Collectivism Scale - Triandis1998 - EN/FR

Start of Block: media

Display This Question:
If viz = 0

control Now we need your answers to some questions about contagious diseases.

control Nous avons maintenant également besoin de votre avis sur certaines questions concernant les maladies contagieuses.

Display This Question:
If viz = 1
Or viz = 2

robertkochgeneric Now you will visit another website with some information about contagious diseases.

When you are done, **please come back here to finish the survey.**

robertkochgeneric Vous allez maintenant visiter un autre site web contenant des informations sur les maladies contagieuses.

Lorsque vous aurez terminé, **revenez ici pour terminer ce sondage, s’il vous plaît.**

Display This Question:

If viz = 1

And disease = 0

And voice = 1

herdimmgenericmale Please click here to go to the other website

herdimmgenericmale Veuillez cliquer ici pour accéder à l'autre site web

Display This Question:

If viz = 1

And disease = 0

And voice = 2

herdimmgenericfemale Please click here to go to the other website

herdimmgenericfemale Veuillez cliquer ici pour accéder à l'autre site web

Display This Question:

If viz = 1

And disease = 1

And voice = 1

herdimmmeaslesmale Please click here to go to the other website

herdimmmeaslesmale Veuillez cliquer ici pour accéder à l'autre site web

Display This Question:

If viz = 1

And disease = 1

And voice = 2

herdimmmeaslesfemale Please click here to go to the other website

herdimmmeaslesfemale Veuillez cliquer ici pour accéder à l'autre site web

Display This Question:

If viz = 1

And disease = 2

And voice = 1

herdimmpertussismale Please click here to go to the other website

herdimmpertussismale Veuillez cliquer ici pour accéder à l'autre site web

---

Display This Question:

If viz = 1

And disease = 2

And voice = 2

herdimmpertussisfema Please click here to go to the other website

herdimmpertussisfema Veuillez cliquer ici pour accéder à l'autre site web

---

Display This Question:

If viz = 1

And disease = 3

And voice = 1

herdimmflumale Please click here to go to the other website

herdimmflumale Veuillez cliquer ici pour accéder à l'autre site web

---

Display This Question:

If viz = 1

And disease = 3

And voice = 2

herdimmflufemale Please click here to go to the other website

herdimmflufemale Veuillez cliquer ici pour accéder à l'autre site web

---

Display This Question:

If Q_Language = EN

And viz = 2

robertkoch [Please click here to go to the other website]

robertkoch NA

Display This Question:

If Q_Language = EN

And viz = 3

sbsnewsgeneric Now we ask you to please watch the video below with some information to do with contagious diseases and then continue the survey.

sbsnewsgeneric NA

Display This Question:

If Q_Language = EN

And viz = 5

guardianmeasles
Now we ask you to please watch the video below with some information to do with contagious diseases and then continue the survey.

guardianmeasles NA

Display This Question:

If Q_Language = EN

And viz = 6

theotheredmundmeasle Now we ask you to please watch the video below with some information to do with contagious diseases and then continue the survey.

theotheredmundmeasle NA

Display This Question:

If viz = 7

publichealthagencyca Now we ask you to please watch the video below with some information to do with contagious diseases and then continue the survey.

publichealthagencyca Nous vous demandons maintenant de visionner la vidéo ci-dessous, qui contient des informations sur les maladies contagieuses, puis de poursuivre le sondage.

t_media Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_media Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: media

Start of Block: Risk Perception

riskpercep1 Imagine there are two groups of 100 people each. In one group, there is 1 UNvaccinated person surrounded by 99 vaccinated people. In the other group, there is 1 vaccinated person surrounded by 99 UNvaccinated people. Who is at higher risk of getting infected with ${e://Field/diseasedescEN}?

	The 1 UNvaccinated person surrounded by 99 vaccinated people is at higher risk.	They are both equally at risk.	The 1 vaccinated person surrounded by a group of 99 UNvaccinated people is at higher risk.
1 ()			

riskpercep1 Imaginez deux groupes composés de 100 personnes dans chacun d’eux. Dans un des groupes, il y a 99 individus vaccinés et une personne NON vaccinée. Dans l’autre groupe, il y a 1 personne vaccinée et 99 NON vaccinées. Qui serait le plus à risque de contracter ${e://Field/diseasedescFR-CA}?

	La personne NON vaccinée dans un groupe de 99 personnes vaccinées est plus à risque.	Elles ont le même niveau de risque.	La personne vaccinée dans un groupe de 99 personnes NON vaccinées est plus à risque.
1 ()			

X→

riskpercep2 A person’s decision to be vaccinated or not against ${e://Field/diseasedescEN} affects only them, individually.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

riskpercep2 La décision d'une personne de se faire vacciner ou non contre ${e://Field/diseasedescFR-CA} ne touche que cette personne individuellement.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

riskpercep3 People who are vaccinated against ${e://Field/diseasedescEN} are less likely to pass it on to others.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

riskpercep3 Les personnes qui sont vaccinées contre ${e://Field/diseasedescFR-CA} sont moins susceptibles de la transmettre.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

riskpercep4 My decision to be vaccinated against ${e://Field/diseasedescEN} has **no impact** on anyone else’s chances of catching ${e://Field/diseasedescEN}.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

riskpercep4 Ma décision de me faire vacciner ou non contre ${e://Field/diseasedescFR-CA} **n'a aucune incidence** sur les chances des autres de contracter ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

riskpercep5 My decision NOT to be vaccinated against ${e://Field/diseasedescEN} has **no impact** on anyone else’s chances of catching ${e://Field/diseasedescEN}.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

riskpercep5 Ma décision de ne PAS me faire vacciner contre ${e://Field/diseasedescFR-CA} **n'a aucune incidence** sur les chances des autres de contracter ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

riskpercep6 If I get vaccinated against ${e://Field/diseasedescEN}, it lowers the risk of vulnerable people in my community (babies, young children, older people, cancer patients) getting sick.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

riskpercep6 Si je me fais vacciner contre ${e://Field/diseasedescFR-CA}, cela réduit le risque que des personnes vulnérables de ma communauté (bébés, jeunes enfants, personnes âgées, patients atteints de cancer) tombent malades.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

t-riskpercp Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t-riskpercp Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Risk Perception

Start of Block: 5CScale

X→

confidence1 I am completely confident that vaccines are safe.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

confidence1 Je suis complètement confiant que les vaccins sont sans danger.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

---

confidence2 Vaccinations are effective.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

confidence2 Les vaccins sont efficaces.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

X→

confidence3 Regarding vaccines, I am confident that public authorities decide in the best interest of the community.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

confidence3 En ce qui concerne les vaccins, je suis convaincu que les autorités publiques décident dans le meilleur intérêt de la communauté.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

X→

complacency1 Vaccination is unnecessary because vaccine-preventable diseases are not common anymore.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

complacency1 La vaccination est inutile car les maladies évitables par la vaccination ne sont plus courantes.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

X→

complacency2 My immune system is so strong, it also protects me against diseases.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

complacency2 Mon système immunitaire est suffisamment fort pour me protéger des maladies.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

---

complacency3 Vaccine-preventable diseases are not so severe that I should get vaccinated.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

complacency3 Les maladies évitables par la vaccination ne sont pas graves au point de devoir me faire vacciner.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

X→

constraints1 Everyday stress prevents me from getting vaccinated.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

constraints1 Le stress quotidien m'empêche de me faire vacciner.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

---

constraints2 For me, it is inconvenient to receive vaccinations.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

constraints2 Pour moi, il est peu pratique de recevoir des vaccins.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

constraints3 Visiting the doctor’s makes me feel uncomfortable; this keeps me from getting vaccinated.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

constraints3 Visiter le médecin me met mal à l’aise; cela m'empêche de me faire vacciner.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

calculation1 When I think about getting vaccinated, I weigh benefits and risks to make the best decision possible.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

calculation1 Quand je pense à me faire vacciner, je mesure les avantages et les risques pour prendre la meilleure décision possible.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

---

calculation2 For each and every vaccination, I closely consider whether it is useful for me.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

calculation2 Pour chaque vaccin, j’évalue de près si cela m’est utile.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

X→

calculation3 It is important for me to fully understand the topic of vaccination, before I get vaccinated.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

calculation3 Il est important pour moi de bien comprendre tout ce qu’englobe la vaccination avant de me faire vacciner.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

X→

collectiverespon1 When everyone else is vaccinated, I don’t have to get vaccinated, too.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

collectiverespon1 Quand tous les autres sont vaccinés, je n’ai pas à me faire vacciner aussi.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

X→

collectiverespon2 I get vaccinated because I can also protect people with a weaker immune system.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

collectiverespon2 Je me fais vacciner parce que je peux aussi protéger les personnes dont le système immunitaire est plus faible.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

collectiverespon3 Vaccination is a collective action to prevent the spread of diseases.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

collectiverespon3 La vaccination est une action collective visant à prévenir la propagation des maladies

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

---

t_5Cscale Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_5Cscale Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: 5CScale

Start of Block: Emotion

emotion1 I am worried about getting ${e://Field/diseasedescEN}.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

emotion1 Je suis inquiet de contracter ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

emotion2 I am worried about people in my life (family, friends) getting ${e://Field/diseasedescEN}.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

emotion2 Je suis inquiet que mon entourage (famille, amis) contracte ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

emotion3 I am worried about vulnerable people in my community (babies, young children, older people, cancer patients) getting ${e://Field/diseasedescEN}.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

emotion3 Je suis inquiet que des gens vulnérables de ma communauté (bébés, jeunes enfants, personnes âgées, patients atteints du cancer) contractent ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

emotion4 I would feel guilty if someone in my life (a family member, a friend) got the ${e://Field/diseasedescEN} from me.

○ strongly disagree (1)

○ moderately disagree (2)

○ slightly disagree (3)

○ neutral (4)

○ slightly agree (5)

○ moderately agree (6)

○ strongly agree (7)

emotion4 Je me sentirais coupable si je transmettais à quelqu'un de mon entourage (famille, amis) ${e://Field/diseasedescFR-CA}.

○ fortement en désaccord (1)

○ modérément en désaccord (2)

○ légèrement en désaccord (3)

○ neutre (4)

○ légèrement en accord (5)

○ modérément en accord (6)

○ fortement en accord (7)

emotion5 I would feel guilty if a vulnerable person (a baby, a young child, an older person, a cancer patient) got ${e://Field/diseasedescEN} from me.

- ○ strongly disagree (1)
- ○ moderately disagree (2)
- ○ slightly disagree (3)
- ○ neutral (4)
- ○ slightly agree (5)
- ○ moderately agree (6)
- ○ strongly agree (7)

emotion5 Je me sentirais coupable si je transmettais à une personne vulnérable (un bébé, un jeune enfant, une personne âgée, un patient atteint du cancer) ${e://Field/diseasedescFR-CA}.

- ○ fortement en désaccord (1)
- ○ modérément en désaccord (2)
- ○ légèrement en désaccord (3)
- ○ neutre (4)
- ○ légèrement en accord (5)
- ○ modérément en accord (6)
- ○ fortement en accord (7)

---

t_emotion Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_emotion Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Emotion

Start of Block: Knowledge

X→

know01 When a person gets vaccinated, it can help protect: (check one)

- ○ The person (1)
- ○ The people around them (2)
- ○ Vulnerable people in their community (3)
- ○ All of the above (4)

know01 Quand une personne est vaccinée, cela peut aider à protéger : (cochez une réponse)

- ○ La personne elle-même (1)
- ○ Les gens autour d'elle (2)
- ○ Les personnes vulnérables dans sa communauté (3)
- ○ Toutes ces réponses (4)

X→

know02 Which people are considered more vulnerable to contagious diseases, either because they can catch them more easily, or the diseases can make them sicker? (check all that apply)

- ☐ People who are very old (1)
- ☐ People who are very young (babies, etc.) (2)
- ☐ People who are athletes (3)
- ☐ People who have cancer (4)
- ☐ People who are healthy adults (5)

know02 Quelles personnes peuvent être considérées comme étant plus vulnérables aux maladies contagieuses, soit parce qu’ils peuvent attraper les maladies plus facilement, et les maladies peuvent les rendre plus malades? (Cochez toutes les réponses qui s'appliquent)

- ☐ Les personnes très âgées (1)
- ☐ Les personnes très jeunes (bébés, etc.) (2)
- ☐ Les personnes qui sont des athlètes (3)
- ☐ Les personnes atteintes du cancer (4)
- ☐ Les adultes en bonne santé (5)

X→

statement For the following statements, please indicate whether they are true, false, or you don’t know:

	True (1)	False (0)	I don't know (77)
${e://Field/diseasedescEN_CAP} can spread from one person to another person. (know03)	○	○	○
When you come into contact with someone who is infected with ${e://Field/diseasedescEN}, there is a chance you could get the disease. (know04)	○	○	○
The only way to get ${e://Field/diseasedescEN} is to touch someone who has the disease. (know05)	○	○	○
${e://Field/diseasedescEN_CAP} always spread through groups of people at the same speed. (know07)	○	○	○
Some vaccines provide less protection than others. (know08)	○	○	○
Every vaccine provides full protection from a single dose (one needle/shot, one dose of drops in the mouth, or a one spray up the nose). (know10)	○	○	○
Some vaccines only provide partial protection after one dose and require multiple doses to get protected. (know11)	○	○	○
Some vaccines become less effective over time and need booster doses. (know12)	○	○	○
Community protection (or herd immunity) means everyone in the community has been vaccinated. (know13)	○	○	○
Unvaccinated people in a population can be protected from infections when enough people in their community are vaccinated. (know14)	○	○	○

Every person in a community (100%) must be vaccinated to achieve community protection. (know15)	○	○	○
The percentage of people (that is, how many members of a community) who must be vaccinated for the community to achieve community protection depends on the disease and the vaccine. (know16)	○	○	○
An individual's decision to get vaccinated or not affects only that individual. (know17)	○	○	○

statement Pour les énoncés suivants, veuillez indiquer s'ils sont vrais, faux ou si vous ne savez pas :

	Vrai (1)	Faux (0)	Je ne sais pas (77)
${e://Field/diseasedescFR-CA_CAP} peut se propager d'une personne à l'autre (know03)	○	○	○
Lorsque vous entrez en contact avec une personne infectée par ${e://Field/diseasedescFR-CA}, vous risquez de contracter cette maladie. (know04)	○	○	○
La seule façon de contracter ${e://Field/diseasedescFR-CA} est d'être en contact direct avec une personne infectée. (know05)	○	○	○
${e://Field/diseasedescFR-CA_CAP} se propage à travers une communauté à la même vitesse. (know07)	○	○	○
Certains vaccins offrent moins de protection que d'autres. (know08)	○	○	○
Chaque vaccin offre une protection complète avec une seule dose (une injection par aiguille, une dose de gouttes dans la bouche ou une seule pulvérisation dans le nez). (know10)	○	○	○
Certains vaccins ne procurent qu'une protection partielle après une dose et nécessitent plusieurs doses pour être protégé. (know11)	○	○	○
Certains vaccins deviennent moins efficaces avec le temps et nécessitent des doses de	○	○	○

rappel. (know12)			
La protection communautaire (ou l'immunité collective) signifie que tout le monde dans la communauté a été vacciné. (know13)	○	○	○
Les gens non-vaccinés d’une population peuvent être protégés des infections quand assez de gens de leur communauté sont vaccinés. (know14)	○	○	○
100% des membres d'une communauté doivent être vaccinés pour assurer la protection communautaire. (know15)	○	○	○
Le pourcentage de personnes (c'est-à-dire combien de membres d'une communauté) qui doivent être vaccinées pour que la communauté puisse assurer la protection communautaire dépend de la maladie et du vaccin. (know16)	○	○	○
La décision d’un individu de se faire vacciner ou non affecte seulement cet individu. (know17)	○	○	○

---

t_know Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_know Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Knowledge

Start of Block: Trust

X→

trustinfo ${e://Field/trustq_EN}

○ not at all trustworthy (1)

○ moderately untrustworthy (2)

○ slightly untrustworthy (3)

○ neutral (4)

○ slightly trustworthy (5)

○ moderately trustworthy (6)

○ strongly trustworthy (7)

trustinfo ${e://Field/trustq_FR-CA}

○ pas du tout digne de confiance (1)

○ modérément NON digne de confiance (2)

○ légèrement NON digne de confiance (3)

○ neutre (4)

○ légèrement digne de confiance (5)

○ moyennement digne de confiance (6)

○ fortement digne de confiance (7)

t_trustinfo Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_trustinfo Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Trust

Start of Block: Behavioural intentions

Display This Question:

If disease = 1

Or disease = 2

Or disease = 3

X→

immune To the best of your knowledge, are you already immune to ${e://Field/diseasedescEN}?

○ Yes (1)

○ No (0)

○ I don’t know (77)

immune A votre connaissance, êtes-vous déjà immunisé contre ${e://Field/diseasedescFR-CA}?

○ Oui (1)

○ Non (0)

○ Je ne sais pas (77)

t_immune Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_immune Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Display This Question:

If To the best of your knowledge, are you already immune to ${e://Field/diseasedescEN}? = Yes

Or If

disease = 0

Q For the next question, imagine you were **not** already immune to ${e://Field/diseasedescEN}.

Q Imaginez que vous **n'étiez pas** déjà immunisé contre ${e://Field/diseasedescFR-CA}.

vaxintention If you were eligible to receive a free vaccine against ${e://Field/diseasedescEN}, how likely would you be to get vaccinated?

	Extremely unlikely, I would definitely NOT be vaccinated	Extremely likely, I would definitely BE vaccinated
1 ()		

vaxintention Si vous étiez admissible à recevoir gratuitement un vaccin contre ${e://Field/diseasedescFR-CA}, quelle serait votre probabilité de recevoir le vaccin ?

	Extrêmement improbable, je ne serais certainement PAS vacciné	Extrêmement probable, je SERAIS certainement vacciné

t_vaxintention Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_vaxintention Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

X→

c19vax Have you already received a COVID-19 vaccine (between Dec 2020 and now)?

○ Yes (1)

○ No (0)

○ I don’t know (77)

c19vax Avez-vous déjà reçu un vaccin contre la COVID-19 (entre décembre 2020 et maintenant) ?

○ Oui (1)

○ Non (0)

○ Je ne sais pas (77)

---

t_c19vax Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_c19vax Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Page Break

Display This Question:

If Have you already received a COVID-19 vaccine (between Dec 2020 and now)? = No

Or Have you already received a COVID-19 vaccine (between Dec 2020 and now)? = I don't know

c19vaxintention Assuming you will be eligible to get a COVID-19 vaccine this year, how likely are you to get the vaccine?

	Extremely unlikely, I will definitely NOT get a COVID-19 vaccine	Extremely likely, I will definitely get a COVID-19 vaccine
1 ()		

c19vaxintention En supposant que vous soyez éligible pour recevoir un vaccin COVID-19 cette année, quelle est la probabilité que vous receviez le vaccin ?

	Extrêmement improbable (c.-à-d. je ne me ferai certainement PAS vacciner contre la COVID-19)	Extrêmement probable, (c.-à-d. je me ferai certainement vacciner contre la COVID-19)
1 ()		

---

t_c19vaxintention Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

t_c19vaxintention Durée
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

End of Block: Behavioural intentions

S3: Study statistal analysis plan

```r
# Project: herdimm
# Description: code for analysing the results of Hina Hakim's final PhD study

# Script developed by: Hina Hakim Jul 2019 - Dec 2020 with the collaboration of Anne-Sophie Julien
# (statistician) and Holly Witteman (PhD advisor). Reviewed externally by a statistician
# unaffiliated with the project in October 2020; report received 2020-10-30.

# Last run: 2021-09-20 with real data from 2021-07-05

#### RESEARCH QUESTIONS ####

# This project uses the herdimm intervention to convey the concept of community immunity or herd
# immunity. The herdimm intervention is described in detail here: DOI: 10.2196/20113

# Research question 1 (primary): Across 4 vaccine-preventable diseases, does the herdimm
# intervention influence risk perception compared to a control?

# Research question 2 (secondary): Across 4 vaccine-preventable diseases, does the herdimm
# intervention influence other outcomes (5C scale, emotions, trust, knowledge, behavioural
# intentions) compared to a control?

# Research question 3 (secondary): For any of the 4 diseases, do existing interventions already
# available online outperform the herdimm intervention?

# Research question 4 (add-on, may be reported separately from Hina Hakim's thesis): Are the effects
# of the herdimm intervention different when the narration uses a female or male voice?


#### LOAD LIBRARIES ####

# Libraries are required to run different packages for running the code below.
library(psych) # describeBy
library(emmeans) # emmeans
library(moments) # kurtosis, skewness
library(car) # Anova
library(MASS) # Box-Cox
library(olsrr) # testing collinearity
library(readr) # read files
library(tidyverse)
library(dplyr)
library(ggplot2) # for plots
library(BlandAltmanLeh) # for Bland Altman test
library(ltm) # to determine Cronbach alpha
library(mice) # for multiple imputations
library(naniar)
library(R2HTML) #CHANGELOG ASJ: to output results 2021-08-06.
library(mitml) #CHANGELOG ASJ: add global test after multiple imputation 21-09-20.

sessionInfo() # for citing packages in paper

#### IMPORT DATA ####

# Import CSV file of raw data, removing Qualtrics artifacts in rows 1-2 (unneeded variable headers)
# and columns 1-16 (unneeded data automatically collected by default by Qualtrics.)

datRaw = (
  read.csv(
    "./contag_phase3b_July 5, 2021_08.19.csv",
    header = T,
    sep = ","
  )[-c(1, 2)]
) #CHANGELOG Changed to new file name 2021-07-05
```

```r
# View(datRaw)

# Subset the data we need: people who started (Consent==1) and finished (Finished==1) the study.
#CHANGELOG ASJ: Replace subset by gc = 1 equivalent to Consent + Finished + Quality, 21-08-06
dat<-subset(datRaw,gc==1)
# View(dat)

#CHANGELOG ASJ: save and print results 2021-09-03.
HTMLStart(outdir=".\\",
          file = "Herdimm Descriptive210903",extension = "html",echo=TRUE, HTMLframe = TRUE)


#### CLEAN DATA ####

#### + Risk Perception ####

# First change the columns (variables) from character to numeric form so that we can work with them.

dat$riskpercep1_1 <- as.numeric(as.character(dat$riskpercep1_1))
dat$riskpercep2 <- as.numeric(as.character(dat$riskpercep2))
dat$riskpercep3 <- as.numeric(as.character(dat$riskpercep3))
dat$riskpercep4 <- as.numeric(as.character(dat$riskpercep4))
dat$riskpercep5 <- as.numeric(as.character(dat$riskpercep5))
dat$riskpercep6 <- as.numeric(as.character(dat$riskpercep6))

# Then recode reverse-coded items riskpercep2, riskpercep4, riskpercep5.

dat$riskpercep2rev <- 8-dat$riskpercep2
dat$riskpercep4rev <- 8-dat$riskpercep4
dat$riskpercep5rev <- 8-dat$riskpercep5

# To check if risk percep_1 measures the same construct as riskpercep_2-6, we rescale riskpercep_1
# so that it is measured on the same scale as the other items.

dat$score_riskpercep1_1 <- dat$riskpercep1_1*6/100+1
summary(dat$score_riskpercep1_1)

# To calculate the mean of the six other risk perception items (that is: riskpercep_2, riskpercep_3,
# riskpercep_4, riskpercep_5, riskpercep_6), first, we create a subset of the items we need.

dat_riskpercep2_6 <-
  subset(
    dat,
    select = c(
      riskpercep2rev,
      riskpercep3,
      riskpercep4rev,
      riskpercep5rev,
      riskpercep6
    )
  )

# Sum all items (riskpercept_2 to riskpercep_6) for each participant.

risk_sumrows <-rowSums(dat_riskpercep2_6,na.rm=T)

# Count the number of missing items (no response in items riskpercept_2 to riskpercep_6) for each
# participant.

risk_summissing <-
  is.na(dat_riskpercep2_6$riskpercep2rev) +
```

```r
  is.na(dat_riskpercep2_6$riskpercep3) +
  is.na(dat_riskpercep2_6$riskpercep4rev) +
  is.na(dat_riskpercep2_6$riskpercep5rev) +
  is.na(dat_riskpercep2_6$riskpercep6)

table(risk_summissing)
prop.table(table(risk_summissing))*100

# Calculate the mean value (mean risk perception) from riskpercept_2 to riskpercep_6 of all answered
# items.

dat$mean_riskperception2_6<-risk_sumrows/(5-risk_summissing)
summary(dat$mean_riskperception2_6)

# To see if the item risk percep_1 and risk percept_2 to 6 are all measuring the same thing, we need
# a Bland Altman plot. To create a Bland Altman plot, we need the difference between _1 and the mean
# of _2 to _6. So, for that purpose, we will create 2 new variables then create the plot.

dat$diff_riskp1_6 <-dat$score_riskpercep1_1-dat$mean_riskperception2_6
summary(dat$diff_riskp1_6)

dat$mean_riskp1_6<- (dat$score_riskpercep1_1+dat$mean_riskperception2_6)/2
summary(dat$mean_riskp1_6)

plot(dat$score_riskpercep1_1,dat$mean_riskperception2_6)

# We are using two = 1.96 (two standard deviations) rather than two = 2.0. If there are data points
# outside these limits, this raises concerns that these two measures may not, in fact, measure the
# same construct.

bland.altman.plot(
  dat$mean_riskperception2_6,
  dat$score_riskpercep1_1,
  two = 1.96,
  mode = 1,
  graph.sys = "base",
  conf.int = 0,
  silent = TRUE,
  sunflower = FALSE,
  geom_count = FALSE
)

# Then we formally test the hypothesis. If this test is statistically significant, it means the
# difference between the two new variables is not zero, which implies a lack of similarity between
# the two original measurement methods (i.e., riskpercep_1 and riskpercep_2 to _6.)

summary(lm(dat$mean_riskp1_6 ~ dat$diff_riskp1_6))

# If riskpercep_1 and riskpercep_2 to _6 are similar, we will combine them into a single measure if
# their combined Cronbach alpha is >=0.8 and will consider doing so if their combined Cronbach alpha
# can be raised to >=0.8 by removing inconsistent items. We will explore removing items because this
# measure is an ad hoc measure created specifically for this study.

# Create a subset for riskpercep_2 to_6 to determine Cronbach alpha more easily

dat_riskpercep2_6_alpha <-
  subset(
    dat,
    select = c(
```

```
      riskpercep2rev,
      riskpercep3,
      riskpercep4rev,
      riskpercep5rev,
      riskpercep6
    )
  )

# Cronbach alpha value for riskpercep_2-6

cronbach.alpha(
  dat_riskpercep2_6_alpha,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# Create  a subset for riskpercep_1 to_6 to determine Cronbach alpha more easily

dat_riskpercep1_6_alpha <-
  subset(
    dat,
    select = c(
      score_riskpercep1_1,
      riskpercep2rev,
      riskpercep3,
      riskpercep4rev,
      riskpercep5rev,
      riskpercep6
    )
  )

# Cronbach alpha value for riskpercep_1-6

cronbach.alpha(
  dat_riskpercep1_6_alpha,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# 1. If we can combine the measures (based on Bland Altman & Cronbach alpha) then we create a single
# measure using the same series of steps as above (subset, sum of all items, number of missing
# items, calculate mean). This is our preferred approach, so we are prioritizing this option in the
# pre-registered code.

# subset
dat_riskpercep <-
  subset(
    dat,
    select = c(
      score_riskpercep1_1,
      riskpercep2rev,
      riskpercep3,
      riskpercep4rev,
      riskpercep5rev,
      riskpercep6
    )
  )

# sum of all items for each person
risk_sumrows <-rowSums(dat_riskpercep,na.rm=T)
```

```r

# count of missing values
risk_summissing <-
  is.na(dat_riskpercep$score_riskpercep1_1) +
  is.na(dat_riskpercep$riskpercep2rev) +
  is.na(dat_riskpercep$riskpercep3) +
  is.na(dat_riskpercep$riskpercep4rev) +
  is.na(dat_riskpercep$riskpercep5rev) +
  is.na(dat_riskpercep$riskpercep6)

# mean
dat$mean_riskperception<-risk_sumrows/(6-risk_summissing)

# 2. If we cannot combine the measures, we will run the rest of the script once with risk perception
# being the first item only, and then re-run the rest of the script with risk perception being items
# 2 to 6. In that case, we will add a comment for which condition we are running the ANOVA each
# time. For now, we have commented these out, as we hope not to have to use them.

# dat$mean_riskperception<-(dat$score_riskpercep1_1)

dat$mean_riskperception<-(dat$mean_riskperception2_6)

# CHANGELOG: As mentioned above that if risk perception item 1 and 2 to 6 are measuring differently, we will use in our analysis ANOVA model separately
# CHANGELOG ASJ: Since riskperception 1 is asymetric, no ANOVA will be run on it, so the mean can be directly equal to 2-6, 2021-08-03

#### + Betsch et al 5C measure ####

# Variables are coming in as data type character. Force them to be numeric so we can work with them.

cols.num <- c(
  "confidence1",
  "confidence2",
  "confidence3",
  "complacency1",
  "complacency2",
  "complacency3",
  "constraints1",
  "constraints2",
  "constraints3",
  "calculation1",
  "calculation2",
  "calculation3",
  "collectiverespon1",
  "collectiverespon2",
  "collectiverespon3"
)

dat[cols.num] <- sapply(dat[cols.num],as.numeric)

# Recode reverse-coded item collectiverespon1

dat$collectiverespon1rev <- 8-dat$collectiverespon1

# Calculate mean (subset, sum of all items, number of missing items, calculate mean).

dat_5c <-
  subset(
    dat,
    select = c(
      confidence1,
      confidence2,
      confidence3,
      complacency1,
```

```
      complacency2,
      complacency3,
      constraints1,
      constraints2,
      constraints3,
      calculation1,
      calculation2,
      calculation3,
      collectiverespon1rev,
      collectiverespon2,
      collectiverespon3
    )
  )

C5_sumrows <- rowSums(dat_5c, na.rm = T)

C5_summissing <-
  is.na(dat_5c$confidence1) +
  is.na(dat_5c$confidence2) +
  is.na(dat_5c$confidence3) +
  is.na(dat_5c$complacency1) +
  is.na(dat_5c$complacency2) +
  is.na(dat_5c$complacency3) +
  is.na(dat_5c$constraints1) +
  is.na(dat_5c$constraints2) +
  is.na(dat_5c$constraints3) +
  is.na(dat_5c$calculation1) +
  is.na(dat_5c$calculation2) +
  is.na(dat_5c$calculation3) +
  is.na(dat_5c$collectiverespon1rev) +
  is.na(dat_5c$collectiverespon2) +
  is.na(dat_5c$collectiverespon3)

dat$mean_C5 <- C5_sumrows / (15 - C5_summissing)

# Determine Cronbach alpha of overall 5C scale.
#CHANGELOG ASJ: Overall consistency is low, while it's high for subscales. Analysis of subscales instead of total score, 2021-08-06
cronbach.alpha(dat_5c, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)# CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14


# Create subsets for each subscale .

dat_5c_confidence <-subset(dat,select=c(confidence1,confidence2,confidence3))
dat_5c_complacency<-subset(dat,select=c(complacency1,complacency2,complacency3))
dat_5c_constraints<-subset(dat,select=c(constraints1,constraints2,constraints3))
dat_5c_calculation<-subset(dat,select=c(calculation1,calculation2,calculation3))
dat_5c_collectiverespon<-subset(dat,select=c(collectiverespon1rev,collectiverespon2,collectiverespon3))

# Determine Cronbach alpha values for each of 5C subscales.

# Confidence
cronbach.alpha(
  dat_5c_confidence,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# Complacency
cronbach.alpha(
  dat_5c_complacency,
  standardized = FALSE,
  CI = FALSE,
```

```r
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# Constraints
cronbach.alpha(
  dat_5c_constraints,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# Calculation
cronbach.alpha(
  dat_5c_calculation,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

# Collective responsibility
cronbach.alpha(
  dat_5c_collectiverespon,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

#CHANGELOG ASJ: Calculate mean for each subscale because of low overall consistency, 2021-08-06.
C5conf_sumrows <- rowSums(dat_5c_confidence, na.rm = T)
C5comp_sumrows <- rowSums(dat_5c_complacency, na.rm = T)
C5cons_sumrows <- rowSums(dat_5c_constraints, na.rm = T)
C5calc_sumrows <- rowSums(dat_5c_calculation, na.rm = T)
C5coll_sumrows <- rowSums(dat_5c_collectiverespon, na.rm = T)


C5conf_summissing <-    is.na(dat_5c$confidence1) +  is.na(dat_5c$confidence2) + is.na(dat_5c$confidence3)
C5comp_summissing <-    is.na(dat_5c$complacency1) +  is.na(dat_5c$complacency2) + is.na(dat_5c$complacency3)
C5cons_summissing <-    is.na(dat_5c$constraints1) +  is.na(dat_5c$constraints2) + is.na(dat_5c$constraints3)
C5calc_summissing <-    is.na(dat_5c$calculation1) +  is.na(dat_5c$calculation2) + is.na(dat_5c$calculation3)
C5coll_summissing <-    is.na(dat_5c$collectiverespon1rev) + is.na(dat_5c$collectiverespon2) +  is.na(dat_5c$collectiverespon3)

dat$mean_C5_conf <- C5conf_sumrows / (3 - C5conf_summissing)
dat$mean_C5_comp <- C5comp_sumrows / (3 - C5comp_summissing)
dat$mean_C5_cons <- C5cons_sumrows / (3 - C5cons_summissing)
dat$mean_C5_calc <- C5calc_sumrows / (3 - C5calc_summissing)
dat$mean_C5_coll <- C5coll_sumrows / (3 - C5coll_summissing)

#CHANGELOG ASJ: Dichotomize the five subscales because they are highly asymetric, 21-09-03
hist(dat$mean_C5_conf)
hist(dat$mean_C5_comp)
hist(dat$mean_C5_cons)
hist(dat$mean_C5_calc)
hist(dat$mean_C5_coll)

```

```
summary(dat$mean_C5_conf) #6
summary(dat$mean_C5_comp) #2.33
summary(dat$mean_C5_cons) #2
summary(dat$mean_C5_calc) #5.33
summary(dat$mean_C5_coll) #6.33

dat$mean_C5_conf2<-ifelse(dat$mean_C5_conf>=6,1,0)
dat$mean_C5_comp2<-ifelse(dat$mean_C5_comp<=2.34,1,0)
dat$mean_C5_cons2<-ifelse(dat$mean_C5_cons<=2,1,0)
dat$mean_C5_calc2<-ifelse(dat$mean_C5_calc>=5.33,1,0)
dat$mean_C5_coll2<-ifelse(dat$mean_C5_coll>=6.33,1,0)

table(dat$mean_C5_conf2,dat$mean_C5_conf)
table(dat$mean_C5_comp2,dat$mean_C5_comp)
table(dat$mean_C5_cons2,dat$mean_C5_cons)
table(dat$mean_C5_calc2,dat$mean_C5_calc)
table(dat$mean_C5_coll2,dat$mean_C5_coll)

table(dat$mean_C5_conf2)
table(dat$mean_C5_comp2)
table(dat$mean_C5_cons2)
table(dat$mean_C5_calc2)
table(dat$mean_C5_coll2)


#### + Emotions ####

# Variables are coming in as data type character. Force them to be numeric so we can work with them.

cols.num <- c("emotion1", "emotion2", "emotion3", "emotion4", "emotion5")
dat[cols.num] <- sapply(dat[cols.num],as.numeric)

# Calculate mean (subset, sum of all items, number of missing items, calculate mean).

dat_emotion <-
  subset(dat, select = c(emotion1, emotion2, emotion3, emotion4, emotion5))

emotion_sumrows <- rowSums(dat_emotion, na.rm = T)

emotion_summissing <-
  is.na(dat_emotion$emotion1) +
  is.na(dat_emotion$emotion2) +
  is.na(dat_emotion$emotion3) +
  is.na(dat_emotion$emotion4) +
  is.na(dat_emotion$emotion5)

dat$mean_emotion <- emotion_sumrows / (5 - emotion_summissing)

# Calculate Cronbach alpha for emotions questions.

cronbach.alpha(
  dat_emotion,
  standardized = FALSE,
  CI = FALSE,
  probs = c(0.025, 0.975),
  B = 1000,
  na.rm = TRUE
) # CHANGELOG Change na.rm = FALSE to TRUE 2021-07-14

#### + Knowledge ####

# Variables are coming in as data type character. Force them to be numeric so we can work with them.
# The original knowledge scale included slightly repetitive items know06 and know09 that were
# removed to shorten the overall survey.

cols.num <-
```

```r
  c(
    "know01",
    "know02_1",
    "know02_2",
    "know02_3",
    "know02_4",
    "know02_5",
    "know03",
    "know04",
    "know05",
    "know07",
    "know08",
    "know10",
    "know11",
    "know12",
    "know13",
    "know14",
    "know15",
    "know16",
    "know17"
  )

dat[cols.num] <- lapply(dat[cols.num],as.numeric)

# Calculate the knowledge score. If the participant selected the correct response, they get 1 point
# for that item. Otherwise they get 0 for that item. know01 and know02 are multiple choice.
# know03-17 are true/false. For know01, we required a single correct choice. For know02, we require
# all correct options selected. For know03-17, we require the correct choice (true or false.)

dat$know01_score <- ifelse(dat$know01 == 4, 1, 0)
dat$know02_score <- ifelse(dat$know02_1 == 1 &
                           dat$know02_2 == 1 &
                           dat$know02_3 == 0 &
                           dat$know02_4 == 1 &
                           dat$know02_5 == 0,1,0)
dat$know03_score <- ifelse(dat$know03 == 1, 1, 0)
dat$know04_score <- ifelse(dat$know04 == 1, 1, 0)
dat$know05_score <- ifelse(dat$know05 == 1, 1, 0)
dat$know07_score <- ifelse(dat$know07 == 0, 1, 0)
dat$know08_score <- ifelse(dat$know08 == 1, 1, 0)
dat$know10_score <- ifelse(dat$know10 == 0, 1, 0)
dat$know11_score <- ifelse(dat$know11 == 1, 1, 0)
dat$know12_score <- ifelse(dat$know12 == 1, 1, 0)
dat$know13_score <- ifelse(dat$know13 == 0, 1, 0)
dat$know14_score <- ifelse(dat$know14 == 1, 1, 0)
dat$know15_score <- ifelse(dat$know15 == 0, 1, 0)
dat$know16_score <- ifelse(dat$know16 == 1, 1, 0)
dat$know17_score <- ifelse(dat$know17 == 0, 1, 0)

# Subset the items so we can easily calculate total score.

dat_know <-
  subset(
    dat,
    select = c(
      know01_score,
      know02_score,
      know03_score,
      know04_score,
      know05_score,
      know07_score,
      know08_score,
      know10_score,
      know11_score,
      know12_score,
```

```
      know13_score,
      know14_score,
      know15_score,
      know16_score,
      know17_score
    )
  )

# Calculate sum of correct answers out of 15 total.

dat$sum_knowledge <- rowSums(dat_know,na.rm=T)

# If we want to know the number of NA's in the data we can run the line below. This would allow us
# to distinguish between people who selected the wrong answer and people who indicated they didn't
# know. We will do this if knowledge scores are very low across all conditions (e.g., mean scores
# less than about 5.) It is commented out for the preregistered code because we hope not to need it.

#CHANGELOG ASJ: the mean of sum_knowledge is 8.9, so no need for it, 2021-08-03
# dat$sum_knowledge <-rowSums(dat_know,na.rm=F)

#### + Trust ####

# Variable is coming in as data type character. Force it to be numeric so we can work with it.

dat$trustinfo <- as.numeric(dat$trustinfo)

#### + Behavourial Intentions ####

# Variables are coming in as data type character. Force them to be numeric so we can work with them.

cols.num <-
  c(
    "immune",
    "vaxintention_1",
    "c19vax",
    "c19vaxintention_1"
  )

dat[cols.num] <- sapply(dat[cols.num], as.numeric)

#### + Individualism and Collectivism Scale ####

# Variables are coming in as data type character. Force them to be numeric so we can work with them.

cols.num <-
  c(
    "ih1",
    "ih2",
    "ih3",
    "ih4",
    "iv1",
    "iv2",
    "iv3",
    "iv4",
    "ch1",
    "ch2",
    "ch3",
    "ch4",
    "cv1",
    "cv2",
    "cv3",
    "cv4"
```

```
  )

dat[cols.num] <- sapply(dat[cols.num],as.numeric)

# This scale may have up to 4 dimensions, so we need to check for collinearity. To do this, we
# calculate mean scores for each participant for each of the 4 subscales, using the same method as
# above (subset, sum, count missing, calculate mean scores.) Then we enter these in a linear model
# on our primary outcome and check VIF for indications of multicollinearity. If all the VIF are
# below 10, it means there is no multicollinearity. If some of the VIF values are more than 10, it
# means that variables are correlated. If there is collinearity then we can not use the 4
# dimensions in the model. In that case, we can use either 2 dimensions or just one or 3 dimensions.
# If there is no collinearity then we can use the 4 dimensions.

dat_indivhorz <-subset(dat,select=c(ih1,ih2,ih3,ih4))
dat_indivvertical <-subset(dat,select=c(iv1,iv2,iv3,iv4))
dat_collhorz <-subset(dat,select=c(ch1,ch2,ch3,ch4))
dat_collvertical <-subset(dat,select=c(cv1,cv2,cv3,cv4))
dat_indiv_coll <- subset(dat,select=c(ih1,ih2,ih3,ih4,iv1,iv2,iv3,iv4,ch1,ch2,ch3,ch4,cv1,cv2,cv3,cv4))

# Cronbach alpha for subscales (each dimension) of the scale.
# CHANGELOG Change below na.rm = FALSE to TRUE 2021-07-14

cronbach.alpha(dat_indivhorz, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)
cronbach.alpha(dat_indivvertical, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)
cronbach.alpha(dat_collhorz, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)
cronbach.alpha(dat_collvertical, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)
cronbach.alpha(dat_indiv_coll, standardized = FALSE, CI = FALSE,
               probs = c(0.025, 0.975), B = 1000, na.rm = TRUE)

# Sum all items

indivhorz_sumrows <- rowSums(dat_indivhorz,na.rm=T)
indivvertical_sumrows <- rowSums(dat_indivvertical,na.rm=T)
collhorz_sumrows <- rowSums(dat_collhorz,na.rm=T)
collvertical_sumrows <- rowSums(dat_collvertical,na.rm=T)

# Count missing items

indivhorz_summissing <-
  is.na(dat_indivhorz$ih1) + is.na(dat_indivhorz$ih2) + is.na(dat_indivhorz$ih3) +
  is.na(dat_indivhorz$ih4)

indivvertical_summissing <-
  is.na(dat_indivvertical$iv1) + is.na(dat_indivvertical$iv2) + is.na(dat_indivvertical$iv3) +
  is.na(dat_indivvertical$iv4)

collhorz_summissing <-
  is.na(dat_collhorz$ch1) + is.na(dat_collhorz$ch2) + is.na(dat_collhorz$ch3) +
  is.na(dat_collhorz$ch4)

collvertical_summissing <-
  is.na(dat_collvertical$cv1) + is.na(dat_collvertical$cv2) + is.na(dat_collvertical$cv3) +
  is.na(dat_collvertical$cv4)

# Calculate mean scores

```

```r
dat$mean_indivhorz <- indivhorz_sumrows/(4- indivhorz_summissing)
dat$mean_indivvertical <- indivvertical_sumrows/(4 -indivvertical_summissing)
dat$mean_collhorz <- collhorz_sumrows/(4 - collhorz_summissing)
dat$mean_collvertical <- collvertical_sumrows/(4 - collvertical_summissing)

# Test for collinearity
#CHANGELOG ASJ: no collineary, the 4 variables will be kept for analysis as covariates, 2021-08-03

model <-
  lm(
    mean_riskperception  ~ mean_indivhorz + mean_indivvertical + mean_collhorz + mean_collvertical ,
    data = dat
  )
vif(model)
ols_coll_diag(model)

#### + Other data cleaning for demographics ####

# Force year of birth as numeric and calculate age accordingly.

dat$yearofbirth <- as.numeric(dat$yearofbirth)
dat$yearofbirth[dat$yearofbirth==888] <- NA
#CHANGELOG ASJ: year of the survey was postponed to 2021 instead of 2020, 2021-08-03
dat$age <- 2021 - dat$yearofbirth


dat$agecat <- as.numeric(dat$agecat)

# Recode multiple answers for language and ethnicity so that people are appropriately counted in
# each group. These questions are set up so that people "check all that apply." Each of the
# questions also includes a single option to indicate "I prefer not to answer." For each language
# and ethnicity, we want to count the number of participants who affirmatively checked it
# (indicating 'yes' for that option), those who did not check it but did check other options in the
# list (indicating 'no' for that option), and those who checked that they prefer not to answer
# (indicating 'prefer not to answer' for each option in the list). To do this, we use a 2-step
# process in which we first recode the true 'no' values for each option as 0, meaning that all 'yes'
# answers are 1 and all 'no' answers are 0. Then we assign any 'prefer not to answer' to be 'NA'.

# Force variables to be numeric.

cols.num <-
  c(
    "bornincanada",
    "language_1",
    "language_2",
    "language_3",
    "language_888",
    "ethnicity_1",
    "ethnicity_2",
    "ethnicity_3",
    "ethnicity_4",
    "ethnicity_5",
    "ethnicity_6",
    "ethnicity_7",
    "ethnicity_8",
    "ethnicity_9",
    "ethnicity_10",
    "ethnicity_11",
    "ethnicity_12",
```

```
    "ethnicity_13",
    "ethnicity_14",
    "ethnicity_15",
    "ethnicity_16",
    "ethnicity_17",
    "ethnicity_888",
    "disability",
    "techdisability",
    "sexatbirth",
    "genderidentity",
    "supportedbyincome",
    "educationlevel",
    "studyarm",
    "viz",
    "disease",
    "immune",
    "c19vax"
  )

dat[cols.num] <- lapply(dat[cols.num],as.numeric)

# NB. language_1=English, language_2=French, language_3=Other

dat$language_1[is.na(dat$language_1)] <- 0
dat$language_2[is.na(dat$language_2)] <- 0
dat$language_3[is.na(dat$language_3)] <- 0

dat$language_1[dat$language_888==1&dat$language_1==0&dat$language_2==0&dat$language_3==0] <- NA
dat$language_2[dat$language_888==1&dat$language_1==0&dat$language_2==0&dat$language_3==0] <- NA
dat$language_3[dat$language_888==1&dat$language_1==0&dat$language_2==0&dat$language_3==0] <- NA

# ethnicity_1=Asian - East (e.g., Chinese, Japanese, Korean)
dat$ethnicity_1[is.na(dat$ethnicity_1)] <- 0
# ethnicity_2=Asian - Central (e.g., Kazakhstani, Uzbekistani)
dat$ethnicity_2[is.na(dat$ethnicity_2)] <- 0
# ethnicity_3=Asian - South (e.g., Indian, Pakistani, Sri Lankan)
dat$ethnicity_3[is.na(dat$ethnicity_3)] <- 0
# ethnicity_4=Asian - South-East (e.g., Malaysian, Filipino, Vietnamese)
dat$ethnicity_4[is.na(dat$ethnicity_4)] <- 0
# ethnicity_5=Indigenous or aboriginal person from outside of North America (e.g., Maori, Quechua)
dat$ethnicity_5[is.na(dat$ethnicity_5)] <- 0
# ethnicity_6=White / European (e.g., English, Italian, Portuguese, Russian)
dat$ethnicity_6[is.na(dat$ethnicity_6)] <- 0
# ethnicity_7=White / North American (e.g., Canadian, American)
dat$ethnicity_7[is.na(dat$ethnicity_7)] <- 0
# ethnicity_8=Inuit
dat$ethnicity_8[is.na(dat$ethnicity_8)] <- 0
# ethnicity_9=Latin American (e.g., Chilean, Mexican, Salvadorian)
dat$ethnicity_9[is.na(dat$ethnicity_9)] <- 0
# ethnicity_10=Metis (Metis Nation in Canada)
dat$ethnicity_10[is.na(dat$ethnicity_10)] <- 0
# ethnicity_11=North African (e.g., Moroccan, Tunisian)
dat$ethnicity_11[is.na(dat$ethnicity_11)] <- 0
# ethnicity_12=Middle Eastern (e.g., Egyptian, Iranian, Lebanese)
dat$ethnicity_12[is.na(dat$ethnicity_12)] <- 0
# ethnicity_13=Black - African (e.g., Ghanaian, Kenyan, Somali)
dat$ethnicity_13[is.na(dat$ethnicity_13)] <- 0
# ethnicity_14=Black - Caribbean (e.g., Barbadian, Jamaican)
dat$ethnicity_14[is.na(dat$ethnicity_14)] <- 0
# ethnicity_15=Black - North American (e.g., Canadian, American)
dat$ethnicity_15[is.na(dat$ethnicity_15)] <- 0
# ethnicity_16=First Nations
dat$ethnicity_16[is.na(dat$ethnicity_16)] <- 0
# ethnicity_17= Other
dat$ethnicity_17[is.na(dat$ethnicity_17)] <- 0
```

```r

# Subset ethnicity values

dat_ethnicity <-
  subset(
    dat,
    select = c(
      ethnicity_1,
      ethnicity_2,
      ethnicity_3,
      ethnicity_4,
      ethnicity_5,
      ethnicity_6,
      ethnicity_7,
      ethnicity_8,
      ethnicity_9,
      ethnicity_10,
      ethnicity_11,
      ethnicity_12,
      ethnicity_13,
      ethnicity_14,
      ethnicity_15,
      ethnicity_16,
      ethnicity_17
    )
  )

# Sum of the 17 ethnicity catogories checked

dat$sum_ethnicity <-rowSums(dat_ethnicity,na.rm=T)

# Assign NA when the participant selected prefer not to answer (_888) and no other option.

dat$ethnicity_1[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_2[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_3[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_4[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_5[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_6[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_7[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_8[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_9[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_10[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_11[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_12[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_13[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_14[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_15[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_16[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA
dat$ethnicity_17[dat$ethnicity_888==1&dat$sum_ethnicity==0] <- NA

# Regrouping ethnicities for later covariate analyses. E.g., if participant checked one or more of
# Asian options, they get a 1 for Asian_group. If they checked no Asian options, they get a 0 for
# Asian_group.

dat$Asian_group <-
  ifelse(
    dat$ethnicity_1 == 1 |
      dat$ethnicity_2 == 1 |
      dat$ethnicity_3 == 1 |
      dat$ethnicity_4 == 1,
    1,
    0
  )

dat$Black_group <-
```

```r
  ifelse(
    dat$ethnicity_13 == 1 |
      dat$ethnicity_14 == 1 |
      dat$ethnicity_15 == 1,
    1,
    0
  )

dat$Indigenous_group <-
  ifelse(
    dat$ethnicity_16 == 1 |
      dat$ethnicity_5 == 1 |
      dat$ethnicity_8 == 1 |
      dat$ethnicity_10 == 1,
    1,
    0
  )

dat$MaghrebMiddleEast_group <-
  ifelse(
    dat$ethnicity_11 == 1 |
      dat$ethnicity_12 == 1,
    1,
    0
  )

dat$White_group <-
  ifelse(
    dat$ethnicity_6 == 1 |
      dat$ethnicity_7 == 1,
    1,
    0
  )

# CHANGELOG: We are creating a new variable for high and low edcation to be later used in 2 way and one way ANOVA analysis 2021-07-28
#CHANGELOG ASJ: We need to exclude the do not know and prefer not to answer from the falses, 2021-08-03
dat$edhi <- ifelse(dat$educationlevel == 4 |
                     dat$educationlevel == 5|
                     dat$educationlevel == 6|
                     dat$educationlevel == 7, TRUE,
                   ifelse(dat$educationlevel == 8 | dat$educationlevel == 888, NA,FALSE))
table(dat$educationlevel,dat$edhi)

# Now we assign NA for 'prefer not to answer' responses for the rest of the socidemographic items.
# We will report the count of 'prefer not to answer' responses in the paper, but for our statistical
# analyses, we group these with the other responses indicating a lack of affirmative response.

dat$bornincanada[dat$bornincanada==888] <- NA
dat$disability[dat$disability==888] <- NA
dat$techdisability[dat$techdisability==888] <- NA
dat$sexatbirth[dat$sexatbirth==888] <- NA
dat$genderidentity[dat$genderidentity==888] <- NA
dat$income[dat$income==888] <- NA
dat$income[dat$income==5] <- NA
dat$supportedbyincome[dat$supportedbyincome==888] <- NA
dat$educationlevel[dat$educationlevel==888] <- NA
dat$educationlevel[dat$educationlevel==8] <- NA

#CHANGELOG ASJ: i prefer not to answer also put to NA, 2021-08-03
dat$agecat[dat$agecat==888 | dat$agecat==1] <- NA

# Now we force all categorical variables as factors and add labels for ease of interpretation later.

```

```
dat$bornincanada <-
  factor(dat$bornincanada,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$language_1 <-
  factor(dat$language_1,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$language_2 <-
  factor(dat$language_2,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$language_3 <-
  factor(dat$language_3,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_1 <-
  factor(dat$ethnicity_1,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_2 <-
  factor(dat$ethnicity_2,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_3 <-
  factor(dat$ethnicity_3,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_4 <-
  factor(dat$ethnicity_4,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_5 <-
  factor(dat$ethnicity_5,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_6 <-
  factor(dat$ethnicity_6,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_7 <-
  factor(dat$ethnicity_7,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_8 <-
  factor(dat$ethnicity_8,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_9 <-
  factor(dat$ethnicity_9,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_10 <-
  factor(dat$ethnicity_10,
         levels = c(1, 0),
         labels = c("Yes", "No"))
```

```r

dat$ethnicity_11 <-
  factor(dat$ethnicity_11,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_12 <-
  factor(dat$ethnicity_12,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_13 <-
  factor(dat$ethnicity_13,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_14 <-
  factor(dat$ethnicity_14,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_15 <-
  factor(dat$ethnicity_15,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_16 <-
  factor(dat$ethnicity_16,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$ethnicity_17 <-
  factor(dat$ethnicity_17,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$Asian_group <-
  factor(dat$Asian_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$MaghrebMiddleEast_group <-
  factor(dat$MaghrebMiddleEast_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$Black_group <-
  factor(dat$Black_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$White_group <-
  factor(dat$White_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$Indigenous_group <-
  factor(dat$Indigenous_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

dat$disability <-
  factor(
    dat$disability,
    levels = c(0, 1),
    labels = c("No disability", "yes,at least one")
  )

dat$techdisability <-
```

```r
  factor(dat$techdisability,
         levels = c(0, 1),
         labels = c("No", "Yes"))

dat$sexatbirth <-
  factor(dat$sexatbirth,
         levels = c(1, 2),
         labels = c("Female", "Male"))

dat$genderidentity <- factor(
  dat$genderidentity,
  levels = c(1, 2, 3, 4),
  labels = c(
    "Female",
    "Male",
    "Indigenous or other cultural gender minority identity",
    "Something else"
  )
)

dat$income <-
  factor(
    dat$income,
    levels = c(1, 2, 3, 4),
    labels = c("24 999 or less", "25 000 to 49 999",
               "50 000 to 99 999", "100 000 or more")
  )

dat$educationlevel <- factor(
  dat$educationlevel,
  levels = c(1, 2, 3, 4, 5, 6, 7),
  labels = c(
    "Some Elementary School",
    "High School Diploma",
    "Apprenticeship or trade certificate or diploma",
    "College or polytechnical school certificate or diploma",
    "University degree bachelor level or below",
    "University graduate degree Master level",
    "University graduate degree Doctorate level"
  )
)

dat$agecat<-
  factor(
    dat$agecat,
    levels = c(2, 3, 4),
    labels = c("18-34 years", "35-49 years", "50 years and more")
  )

dat$UserLanguage <- as.factor(dat$UserLanguage)

dat$studyarm <-
  factor (
    dat$studyarm,
    levels = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13),
    labels = c(
      "controlgeneric",
      "controlmeasles",
      "controlpertussis",
      "controlflu",
      "herdimmgeneric",
      "herdimmmeasles",
      "herdimmpertussis",
      "herdimmflu",
      "robertkochgeneric",
      "sbsnewsgeneric",
      "guardianmeasles",
      "theotheredmundmeasles",
      "publichealthagencycanadaflu"
```

```r
    )
  )

dat$viz <- factor (
  dat$viz,
  levels = c(0, 1, 2, 3, 5, 6, 7),
  labels = c(
    "no visualization",
    "herdimm",
    "robertkoch",
    "sbsnews",
    "guardian",
    "theotheredmund",
    "publichealthagencycanada"
  )
)

dat$disease <-
  factor (
    dat$disease,
    levels = c(0, 1, 2, 3),
    labels = c("Generic", "Measles", "Pertussis", "Flu")
  )


#Keep I don't know categories as is, not as NA, to distinguish them, for these special variables (COVID-19)
dat$immune <-
  factor(
    dat$immune,
    levels = c(1, 0, 77),
    labels = c("Yes", "No", "I don't know")
  )

dat$c19vax <-
  factor(
    dat$c19vax,
    levels = c(1, 0, 77),
    labels = c("Yes", "No", "I don't know")
  )

#CHANGELOG ASJ: All variables related to children were removed from the final survey and are deleted in the rest of the code (children, childimmune, childfluvax), 2021-08-06.


#Make restrictions on the dataset for analysis of patients kept for analysis vs not kept for analysis
#CHANGELOG ASJ: Variables were different than planned, subset done differently because of gc, 2021-08-06.
datConsent<- subset(datRaw,Consent==1)
datConsent$used<-ifelse(datConsent$gc==1 & datConsent$Finished ==1, 1,0) #used in our analysis (n = 5516 vs others with consent (not finished, no quality))
table(datConsent$used)


# Note re: possibly recoding low-frequency data after collecting data

# If we observe low frequencies (below 5%) among socio demographic variables, we will recode these
# groups. This is because small subgroups are unlikely to work well in our planned statistical
# models.

# Descriptive statistics of all original variables will be examined. The following steps will be
# used if a category has frequency smaller than 5% of the total sample. We anticipate this may be
# necessary for gender identity and education level.

```

```
# For gender identity, we will remove smaller categories from inclusion in analytical models since
# they cannot be combined with other category. We still still report these statistics in the
# sociodemographic statistics of the same and will otherwise use these participants' data as usual.
# This means that if categories 3 and 4 have frequencies smaller than 5 % than we will use this
# code:

#CHANGELOG ASJ: the variable is a factor, need to exclude with text instead of numbers, 2021-08-03
#CHANGELOG ASJ: this is removed from comments since <5% answered the two last categories
dat$genderidentity[dat$genderidentity == "Something else" |
                      dat$genderidentity == "Indigenous or other cultural gender minority identity"] <- NA
dat$genderidentity <- droplevels(dat$genderidentity)

# For education level, we will regroup smaller categories with the closest bigger categories with
# which it makes sense to group them. If, for example, there are small frequencies for 'some
# elementary school' we will regroup it with high school diploma. Potential code would be:

# dat$educationlevel[dat$educationlevel== 1] <- 2; dat$educationlevel <-
# factor(dat$educationlevel,levels = c(2,3,4,5,6,7), labels = c("Some Elementary School or High
# School Diploma", "Apprenticeship or trade certificate or diploma College or polytechnical school
# certificate or diploma", "University degree bachelor level or below","University graduate degree
# Master level", "University graduate degree Doctorate level", "Do not know"))

#CHANGELOG ASJ: Education was recoded as high vs low in edhi variable above, instead of using this code, 2021-08-03.

# The rest won't change unless we need to. If we observe another category with frequency below 5%,
# we will follow similar procedures.

#CHANGELOG ASJ: Several covariates have categories with <5% (language_3, MaghrebMiddleEast_group, Black_group, Indigenous_group, ethnicity_9, techdisability). However, they are dichotomous variable, so it is not possible to combine categories. They will simply be dropped from the models.
#CHANGELOG ASJ: However, for techdisability, it will be combined with disability, we don't want to loose that information. 2021-08-03.
dat$disability_any <-ifelse(dat$disability =="yes,at least one" | dat$techdisability == "Yes", 1, 0)
dat$disability_any <-
  factor(
    dat$disability_any,
    levels = c(1, 0),
    labels = c("Yes", "No")
  )

#CHANGELOG ASJ: As mentioned later on in the preregistered code, we are recoding here trust info because it is asymetric, 2021-08-03
dat$Htrustinfo <- ifelse(dat$trustinfo >= 5, 1, 0)
dat$Htrustinfo <-
  factor(
    dat$Htrustinfo,
    levels = c(1, 0),
    labels = c("Yes", "No")
  )

#CHANGELOG ASJ: Because 1st variable of risk perception does not have enough variance and is asymetric, recode as dichotomic variable, 2021-08-03
#CHANGELOG ASJ: Cutpoint = 80, keep people who correctly identified a high risk,
```

```r
2021-08-06.
hist(dat$riskpercep1_1)
dat$RP1_dicho <- ifelse(dat$riskpercep1_1 >= 80, 1, 0)
dat$RP1_dicho <-
  factor(
    dat$RP1_dicho,
    levels = c(1, 0),
    labels = c("High", "Low")
  )
table(dat$RP1_dicho)

#CHANGELOG ASJ: Because vaxintention does not have enough variance and is asymetric, recode as dichotomic variable, 2021-08-06.
#CHANGELOG ASJ: Cutpoint = 91, at or above median.
hist(dat$vaxintention_1)
summary(dat$vaxintention_1)
dat$vaxintent<-ifelse(dat$vaxintention_1>=91,1,0)

dat$vaxintent <-
  factor(
    dat$vaxintent,
    levels = c(1, 0),
    labels = c("High", "Low")
  )

table(dat$vaxintent)

#CHANGELOG ASJ: Because c19vaxintention does not have enough variance and is asymetric, recode as dichotomic variable, 2021-08-06.
#CHANGELOG ASJ: Cutpoint = 90, at or above median.
hist(dat$c19vaxintention_1)
summary(dat$c19vaxintention_1)
dat$c19vaxintent<-ifelse(dat$c19vaxintention_1>=90,1,0)

dat$c19vaxintent <-
  factor(
    dat$c19vaxintent,
    levels = c(1, 0),
    labels = c("High", "Low")
  )

table(dat$c19vaxintent)

#### + Subsets for primary and secondary analyses ####

# We create new subsets for primary and secondary analyses. The first subset (datMain) is for our
# primary research question. The other subsets are for the secondary question.

datMain <-
  subset(
    dat,
    dat$studyarm == "controlgeneric" |
      dat$studyarm == "controlmeasles" |
      dat$studyarm == "controlpertussis" |
      dat$studyarm == "controlflu" |
      dat$studyarm == "herdimmgeneric" |
      dat$studyarm == "herdimmmeasles" |
      dat$studyarm == "herdimmpertussis" |
      dat$studyarm == "herdimmflu"
  )

don<-subset(datMain,select=c(studyarm,mean_riskperception, viz, disease,mean_indivhorz,
                             mean_indivvertical ,
                             mean_collhorz ,
                             mean_collvertical ,
                             bornincanada ,
                             language_1 ,
                             language_2 ,
```

```
                             Asian_group ,
                            White_group,
                             disability_any ,
                             genderidentity ,
                             income ,
                             edhi,
                             age,
                            mean_riskperception2_6,RP1_dicho,vaxintent,mean_emotion,sum_knowledge,Htrustinfo,
                            mean_C5_conf2,mean_C5_comp2,mean_C5_cons2,mean_C5_calc2,mean_C5_coll2)) # CHANGELOG ASJ updated restrictions
                                    #Added the code to determine the missing data
                                  # for variables later used in ANOVA analysis, only those with >5%, 2021-08-03
                                  #remove c19vaxintent, too many missing values



datGeneric <-
  subset(
    dat,
    (
      dat$studyarm == "controlgeneric" |
        dat$studyarm == "herdimmgeneric" |
        dat$studyarm == "robertkochgeneric" |
        dat$studyarm == "sbsnewsgeneric"
    ) &
      UserLanguage == "EN"
  )

datMeasles <-
  subset(
    dat,
    dat$studyarm == "controlmeasles" |
      dat$studyarm == "herdimmmeasles" |
      dat$studyarm == "guardianmeasles" |
      dat$studyarm == "theotheredmundmeasles"
  )

datFlu <-
  subset(
    dat,
    dat$studyarm == "controlflu"  |
      dat$studyarm == "herdimmflu" |
      dat$studyarm == "publichealthagencycanadaflu"
  )

datMain$studyarm <- droplevels(datMain$studyarm)
datGeneric$studyarm <- droplevels(datGeneric$studyarm)
datMeasles$studyarm <- droplevels(datMeasles$studyarm)
datFlu$studyarm <- droplevels(datFlu$studyarm)

#CHANGELOG ASJ, save r dataset for analysis, 2021-09-03
write.csv(dat,file="./datMain.csv",row.names=T)
write.csv(dat,file="./datGeneric.csv",row.names=T)
write.csv(dat,file="./datMeasles.csv",row.names=T)
write.csv(dat,file="./datFlu.csv",row.names=T)

save(dat,file="dat.RData")
save(datMain,file="datMain.RData")
save(datGeneric,file="datGeneric.RData")
save(datMeasles,file="datMeasles.RData")
save(datFlu,file="datFlu.RData")
save(datConsent,file="datConsent.RData")

#### DESCRIPTIVE STATISTICS ####

```

```
# For each variable separately

### Nominal variables - independent variables to which people were randomized

# frequencies - study arm
table(dat$studyarm)
# proportion of each study arm that is
# 0 = controlgeneric
# 1 = controlmeasles
# 2 = controlpertussis
# 3 = controlflu
# 4 = herdimmgeneric
# 5 = herdimmmeasles
# 6 = herdimmpertussis
# 7 = herdimmflu
# 8 = robertkochgeneric
# 9 = sbsnewsgeneric
# 10 = nhsgeneric (no longer included because has been taken down)
# 11 = guardianmeasles
# 12 = theotheredmundmeasles
# 13 = publichealthagencycanadaflu
prop.table(table(dat$studyarm)) * 100

# frequencies - Visualization
table(dat$viz, exclude = NULL)
# proportion of each visualization that is
# 0 = no visualization
# 1 = herdimm
# 2 = robertkoch
# 3 = sbsnews
# 4 = nhs (no longer included because has been taken down)
# 5 = guardian
# 6 = theotheredmund
# 7 = publichealthagencycanada
prop.table(table(dat$viz, exclude = NULL)) * 100

# frequencies - Disease
table(dat$disease, exclude = NULL)
# proportion of each disease that is
# 0 = Generic
# 1 = Measles
# 2 = Pertussis
# 3 = Flu
prop.table(table(dat$disease, exclude = NULL)) * 100

### Nominal variables - in what language did the person complete the survey

# frequencies - User-language for survey
table(dat$UserLanguage, exclude = NULL)
# proportion of language that is
# English (EN)
# French (FR-CA)
prop.table(table(dat$UserLanguage, exclude = NULL)) * 100

### Nominal variables - Demographics # proportion of each demographic variable
# CHANGELOG exclude = NULL added in code below to detemine NA's in the variable 2021-05-10

# frequencies - place of birth (Canada or outside Canada)
table(dat$bornincanada, exclude = NULL)
prop.table(table(dat$bornincanada, exclude = NULL)) * 100

# frequencies - Language_English
table(dat$language_1, exclude = NULL)
prop.table(table(dat$language_1, exclude = NULL)) * 100

# frequencies - Language_French
table(dat$language_2,exclude = NULL)
prop.table(table(dat$language_2,exclude = NULL)) * 100
```

```r

# frequencies - Language_Other
table(dat$language_3,exclude = NULL)
prop.table(table(dat$language_3,exclude = NULL)) * 100

# frequencies - Ethnicity_Asian - East (e.g., Chinese, Japanese, Korean)
table(dat$ethnicity_1, exclude = NULL)
prop.table(table(dat$ethnicity_1, exclude = NULL)) * 100

# frequencies - Ethnicity_Asian - Central (e.g., Kazakhstani, Uzbekistani)
table(dat$ethnicity_2, exclude = NULL)
prop.table(table(dat$ethnicity_2,exclude = NULL)) * 100

# frequencies - Ethnicity_Asian - South (e.g., Indian, Pakistani, Sri Lankan)
table(dat$ethnicity_3,exclude = NULL)
prop.table(table(dat$ethnicity_3, exclude = NULL)) * 100

# frequencies - Ethnicity_Asian - South-East (e.g., Malaysian, Filipino, Vietnamese)
table(dat$ethnicity_4, exclude = NULL)
prop.table(table(dat$ethnicity_4, exclude = NULL)) * 100

# frequencies - Ethnicity_Indigenous or aboriginal person from outside of North America (e.g.,
# Maori, Quechua)
table(dat$ethnicity_5, exclude = NULL)
prop.table(table(dat$ethnicity_5, exclude = NULL)) * 100

# frequencies - Ethnicity_White / European (e.g., English, Italian, Portuguese, Russian)
table(dat$ethnicity_6, exclude = NULL)
prop.table(table(dat$ethnicity_6, exclude = NULL)) * 100

# frequencies - Ethnicity_White / North American (e.g., Canadian, American)
table(dat$ethnicity_7, exclude = NULL)
prop.table(table(dat$ethnicity_7, exclude = NULL)) * 100

# frequencies - Ethnicity_Inuit
table(dat$ethnicity_8,exclude = NULL)
prop.table(table(dat$ethnicity_8,exclude = NULL)) * 100

# frequencies - Ethnicity_Latin American (e.g., Chilean, Mexican, Salvadorian)
table(dat$ethnicity_9,exclude = NULL)
prop.table(table(dat$ethnicity_9,exclude = NULL)) * 100

# frequencies - Ethnicity_Metis (Metis Nation in Canada)
table(dat$ethnicity_10,exclude = NULL)
prop.table(table(dat$ethnicity_10,exclude = NULL)) * 100

# frequencies - Ethnicity_North African (e.g., Moroccan, Tunisian)
table(dat$ethnicity_11,exclude = NULL)
prop.table(table(dat$ethnicity_11,exclude = NULL)) * 100

# frequencies - Ethnicity_Middle Eastern (e.g., Egyptian, Iranian, Lebanese)
table(dat$ethnicity_12,exclude = NULL)
prop.table(table(dat$ethnicity_12,exclude = NULL)) * 100

# frequencies - Ethnicity_Black - African (e.g., Ghanaian, Kenyan, Somali)
table(dat$ethnicity_13,exclude = NULL)
prop.table(table(dat$ethnicity_13,exclude = NULL)) * 100

# frequencies - Ethnicity_Black - Caribbean (e.g., Barbadian, Jamaican)
table(dat$ethnicity_14,exclude = NULL)
prop.table(table(dat$ethnicity_14,exclude = NULL)) * 100

# frequencies - Ethnicity_Black - North American (e.g., Canadian, American)
table(dat$ethnicity_15,exclude = NULL)
prop.table(table(dat$ethnicity_15,exclude = NULL)) * 100

# frequencies - Ethnicity_First Nations
table(dat$ethnicity_16,exclude = NULL)
```

```
prop.table(table(dat$ethnicity_16,exclude = NULL)) * 100

# frequencies - Ethnicity_Other
table(dat$ethnicity_17,exclude = NULL)
prop.table(table(dat$ethnicity_17,exclude = NULL)) * 100

# frequencies - Asian_group
table(dat$Asian_group,exclude = NULL)
prop.table(table(dat$Asian_group,exclude = NULL)) * 100

# frequencies - Black_group
table(dat$Black_group, exclude = NULL)
prop.table(table(dat$Black_group, exclude = NULL)) * 100

# frequencies - Indigenous_group
table(dat$Indigenous_group,exclude = NULL)
prop.table(table(dat$Indigenous_group, exclude = NULL)) * 100

# frequencies - MaghrebMiddleEast_group
table(dat$MaghrebMiddleEast_group, exclude = NULL)
prop.table(table(dat$MaghrebMiddleEast_group, exclude = NULL)) * 100

# frequencies - White_group
table(dat$White_group, exclude = NULL)
prop.table(table(dat$White_group, exclude = NULL)) * 100

# frequencies - disability
table(dat$disability, exclude = NULL)
prop.table(table(dat$disability,exclude = NULL )) * 100

# frequencies - tech disability
table(dat$techdisability, exclude = NULL)
prop.table(table(dat$techdisability, exclude = NULL)) * 100

#CHANGELOG ASJ: frequencies - any disability, 2021-08-03
table(dat$disability_any, exclude = NULL)
prop.table(table(dat$disability_any, exclude = NULL)) * 100

# frequencies - sex at birth
table(dat$sexatbirth, exclude = NULL)
prop.table(table(dat$sexatbirth, exclude = NULL)) * 100

# frequencies - gender identity
table(dat$genderidentity, exclude = NULL)
prop.table(table(dat$genderidentity, exclude = NULL)) * 100

# frequencies - income
table(dat$income, exclude = NULL)
prop.table(table(dat$income, exclude = NULL)) * 100

# frequencies - supportedbyincome
table(dat$supportedbyincome, exclude = NULL)
prop.table(table(dat$supportedbyincome, exclude = NULL)) * 100

# frequencies - education level
table(dat$educationlevel, exclude = NULL)
prop.table(table(dat$educationlevel, exclude = NULL)) * 100

#CHANGELOG ASJ: frequencies - education level high vs low, 2021-08-03
table(dat$edhi, exclude = NULL)
prop.table(table(dat$edhi, exclude = NULL)) * 100

# frequencies - age category
table(dat$agecat,exclude = NULL)
prop.table(table(dat$agecat, exclude = NULL)) * 100

# frequencies - already immune
table(dat$immune, exclude = NULL)
prop.table(table(dat$immune, exclude = NULL)) * 100
```

```

# frequencies - COVID-19 vax
table(dat$c19vax, exclude = NULL)
prop.table(table(dat$c19vax, exclude = NULL)) * 100


### Nominal variables - Other variables
#CHANGELOG ASJ: Section added for new variables created, 2021-08-03
table(dat$Htrustinfo)
prop.table(table(dat$Htrustinfo)) * 100

table(dat$RP1_dicho)
prop.table(table(dat$RP1_dicho)) * 100

table(dat$c19vaxintent)
prop.table(table(dat$c19vaxintent)) * 100

table(dat$vaxintent)
prop.table(table(dat$vaxintent)) * 100

#CHANGELOG ASJ: Add new variables 21-09-03
table(dat$mean_C5_conf2)
prop.table(table(dat$mean_C5_conf2)) * 100

table(dat$mean_C5_comp2)
prop.table(table(dat$mean_C5_comp2)) * 100

table(dat$mean_C5_cons2)
prop.table(table(dat$mean_C5_cons2)) * 100

table(dat$mean_C5_calc2)
prop.table(table(dat$mean_C5_calc2)) * 100

table(dat$mean_C5_coll2)
prop.table(table(dat$mean_C5_coll2)) * 100

### Continuous variables - Demographics, scales, etc.

# Summary statistics for each continuous variable for all participants
summary(dat$mean_riskperception) #CHANGELOG ASJ: this is equivalent to riskperception2_6 2021-08-03
summary(dat$score_riskpercep1_1) # CHANGELOG Added separate analysis for risk perception item 1 2021-07-08
summary(dat$mean_riskperception2_6) # CHANGELOG Added separate analysis for risk perception item 2 to 6 2021-07-08
summary(dat$mean_C5) #CHANGELOG Added separate analysis for 5 subscales, 2021-08-06
summary(dat$mean_C5_conf)
summary(dat$mean_C5_comp)
summary(dat$mean_C5_cons)
summary(dat$mean_C5_calc)
summary(dat$mean_C5_coll)
summary(dat$mean_emotion)
summary(dat$sum_knowledge)
summary(dat$trustinfo)
summary(dat$vaxintention_1)
summary(dat$age)
summary(dat$c19vaxintention_1)
summary(dat$mean_indivhorz)
summary(dat$mean_indivvertical)
summary(dat$mean_collhorz)
summary(dat$mean_collvertical)

#### + Descriptive statistics per study arm ####

### Continuous variables

# The describeBy function in the psych package returns the number of observations, mean, median,
# trimmed means, minimum, maximum, range, skew, kurtosis, and standard error of the
```

```
mean for grouped
# data.

describeBy(dat$mean_riskperception,dat$studyarm) #CHANGELOG ASJ: this is equivalent to riskperception2_6 2021-08-03
describeBy(dat$riskpercep1_1,dat$studyarm) # CHANGELOG Added separate analysis for risk perception item 1 2021-07-08
describeBy(dat$mean_riskperception2_6,dat$studyarm) # CHANGELOG Added separate analysis for risk perception item 2 to 6 2021-07-08
describeBy(dat$mean_C5,dat$studyarm) #CHANGELOG Added separate analysis for 5 subscales, 2021-08-06
describeBy(dat$mean_C5_conf,dat$studyarm)
describeBy(dat$mean_C5_comp,dat$studyarm)
describeBy(dat$mean_C5_cons,dat$studyarm)
describeBy(dat$mean_C5_calc,dat$studyarm)
describeBy(dat$mean_C5_coll,dat$studyarm)
describeBy(dat$mean_emotion,dat$studyarm)
describeBy(dat$sum_knowledge,dat$studyarm)
describeBy(dat$age,dat$studyarm)
describeBy(dat$trustinfo,dat$studyarm)
describeBy(dat$vaxintention_1,dat$studyarm)
describeBy(dat$c19vaxintention_1,dat$studyarm)
describeBy(dat$mean_indivhorz, dat$studyarm)
describeBy(dat$mean_indivvertical, dat$studyarm)
describeBy(dat$mean_collhorz, dat$studyarm)
describeBy(dat$mean_collvertical, dat$studyarm)

### Nominal variables - frequencies and proportion per arm

t1<-data.frame(dat$bornincanada,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Language_English
t1<-data.frame(dat$language_1,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Language_French
t1<-data.frame(dat$language_2,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Language_Other
t1<-data.frame(dat$language_3,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Asian - East (e.g., Chinese, Japanese, Korean)
t1<-data.frame(dat$ethnicity_1,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Asian - Central (e.g., Kazakhstani, Uzbekistani)
t1<-data.frame(dat$ethnicity_2,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Asian - South (e.g., Indian, Pakistani, Sri Lankan)
t1<-data.frame(dat$ethnicity_3,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Asian - South-East (e.g., Malaysian, Filipino, Vietnamese)
t1<-data.frame(dat$ethnicity_4,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Indigenous or aboriginal person from outside of North America (e.g., Maori, Quechua)
```

```r
t1<-data.frame(dat$ethnicity_5,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# White / European (e.g., English, Italian, Portuguese, Russian)
t1<-data.frame(dat$ethnicity_6,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# White / North American (e.g., Canadian, American)
t1<-data.frame(dat$ethnicity_7,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Inuit
t1<-data.frame(dat$ethnicity_8,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Latin American (e.g., Chilean, Mexican, Salvadorian)
t1<-data.frame(dat$ethnicity_9,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Metis (Metis Nation in Canada)
t1<-data.frame(dat$ethnicity_10,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# North African (e.g., Moroccan, Tunisian)
t1<-data.frame(dat$ethnicity_11,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Middle Eastern (e.g., Egyptian, Iranian, Lebanese)
t1<-data.frame(dat$ethnicity_12,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Black - African (e.g., Ghanaian, Kenyan, Somali)
t1<-data.frame(dat$ethnicity_13,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Black - Caribbean (e.g., Barbadian, Jamaican)
t1<-data.frame(dat$ethnicity_14,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Black - North American (e.g., Canadian, American)
t1<-data.frame(dat$ethnicity_15,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# First Nations
t1<-data.frame(dat$ethnicity_16,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Other
t1<-data.frame(dat$ethnicity_17,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Asian_group
t1<-data.frame(dat$Asian_group,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100
```

```r
# MaghrebMiddleEast_group
t1<-data.frame(dat$MaghrebMiddleEast_group,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Black_group
t1<-data.frame(dat$Black_group,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# White_group
t1<-data.frame(dat$White_group,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# Indigenous_group
t1<-data.frame(dat$Indigenous_group,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# disability
t1<-data.frame(dat$disability,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# tech disability
t1<-data.frame(dat$techdisability,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#CHANGELOG ASJ: add any disability, 2021-08-03
t1<-data.frame(dat$disability_any,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# sex at birth
t1<-data.frame(dat$sexatbirth,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# gender identity
t1<-data.frame(dat$genderidentity,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# income
t1<-data.frame(dat$income,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# education level
t1<-data.frame(dat$educationlevel,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#CHANGELOG ASJ: add education level high/low, 2021-08-03
t1<-data.frame(dat$edhi,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# immune
t1<-data.frame(dat$immune,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

# COVID-19 vax
t1<-data.frame(dat$c19vax,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100
```

```


#CHANGELOG ASJ: add 4 new variables, 2021-08-06
#Trust Info
t1<-data.frame(dat$Htrustinfo,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#Risk Perception 1
t1<-data.frame(dat$RP1_dicho,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#VaxIntention
t1<-data.frame(dat$vaxintent,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#Covid19 VaxIntention
t1<-data.frame(dat$c19vaxintent,dat$studyarm)
table(t1)
prop.table(table(t1), 2) * 100

#### + Graphics per arm ####

# Boxplots (NB. find nicer colours for publication/thesis)

par(mar=c(12,5,1,1)) #CHANGELOG ASJ: change margins to see the labels properly, 2021-08-03

# Risk perception
# RECALL: Risk perception item 1 and item 2 to 6 are not measuring the same thing (item 1 will be analysed as dichotomous variable)
# Box plot of risk perception item 2 to 6 for overall data
boxplot(
  dat$mean_riskperception2_6 ~ dat$studyarm,las=2,
  xlab = "",
  ylab = "Risk Perception2_6",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue",
    "khaki",
    "gray75",
    "gray85",
    "firebrick3",
    "dodgerblue"
  )
) # CHANGELOG Added seperate analysis for risk perception item 2 to 6 2021-07-08



# Box plot of risk perception item 2 to 6 for subset Main data
boxplot(
  datMain$mean_riskperception2_6 ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "Risk Perception2_6",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
```

```r
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

# Box plot of risk perception item 2 to 6 for subset Generic data
boxplot(
  datGeneric$mean_riskperception2_6 ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "Risk Perception2_6",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
) # CHANGELOG Added seperate analysis for risk perception item 2 to 6 2021-07-08


par(mar=c(12,5,1,1)) #CHANGELOG ASJ: change margins to see the labels properly, 2021-08-03


# Box plot of risk perceptiom item 2 to 6 for subset Measles data
boxplot(
  datMeasles$mean_riskperception2_6 ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "Risk Perception2_6",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
) # CHANGELOG Added seperate analysis for risk perception item 2 to 6 2021-07-08


# Box plot of risk perception item 2 to 6 for subset Flu data
boxplot(
  datFlu$mean_riskperception2_6 ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "Risk Perception2_6",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
) # CHANGELOG Added seperate analysis for risk perception item 2 to 6 2021-07-08

# Betsch et al C5
#CHANGELOG ASJ: add boxplot for each subscales, 2021-08-06.


boxplot(
  datMain$mean_C5_conf ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "C5 Confidence",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datMain$mean_C5_comp ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "C5 complacency",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
```

```
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datMain$mean_C5_cons ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "C5 constraints",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datMain$mean_C5_calc ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "C5 calculation",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datMain$mean_C5_coll ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "C5 Collective Respons.",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)


boxplot(
  datGeneric$mean_C5_conf ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "C5 confidence",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

```

```
boxplot(
  datGeneric$mean_C5_comp ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "C5 complacency",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

boxplot(
  datGeneric$mean_C5_cons ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "C5 constraints",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

boxplot(
  datGeneric$mean_C5_calc ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "C5 calculation",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

boxplot(
  datGeneric$mean_C5_coll ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "C5 collective respons.",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)



boxplot(
  datMeasles$mean_C5_conf ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "C5 confidence",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datMeasles$mean_C5_comp ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "C5 complacency",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datMeasles$mean_C5_cons ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "C5 contraints",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datMeasles$mean_C5_calc ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "C5 calculation",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datMeasles$mean_C5_coll ~ datMeasles$studyarm,las=2,
  xlab = "",
```

```
  ylab = "C5 collective respons.",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)


boxplot(
  datFlu$mean_C5_conf ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "C5 confidence",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

boxplot(
  datFlu$mean_C5_comp ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "C5 complacency",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

boxplot(
  datFlu$mean_C5_cons ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "C5 constraints",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

boxplot(
  datFlu$mean_C5_calc ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "C5 calculation",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

boxplot(
  datFlu$mean_C5_coll ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "C5 collective respons.",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

# Emotion

boxplot(
  dat$mean_emotion ~ dat$studyarm,las=2,
  xlab = "",
  ylab = "emotion",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue",
    "khaki",
    "gray75",
    "gray85",
    "firebrick3",
    "firebrick1",
    "dodgerblue"
  )
```

```
)

boxplot(
  datMain$mean_emotion ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "emotion",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datGeneric$mean_emotion ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "emotion",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

boxplot(
  datMeasles$mean_emotion ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "emotion",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datFlu$mean_emotion ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "emotion",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)

# Knowledge

boxplot(
  dat$sum_knowledge ~ dat$studyarm,las=2,
  xlab = "",
  ylab = "knowledge",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue",
    "khaki",
    "gray75",
    "gray85",
    "firebrick3",
    "firebrick1",
    "dodgerblue"
  )
)

boxplot(
```

```r
  datMain$sum_knowledge ~ datMain$studyarm,las=2,
  xlab = "",
  ylab = "knowledge",
  main = "Boxplot",
  col = c(
    "white",
    "darksalmon",
    "darkolivegreen1",
    "lightcyan",
    "bisque3",
    "indianred4",
    "green4",
    "darkslateblue"
  )
)

boxplot(
  datGeneric$sum_knowledge ~ datGeneric$studyarm,las=2,
  xlab = "",
  ylab = "knowledge",
  main = "Boxplot",
  col = c("white", "bisque3", "khaki", "gray75", "gray85")
)

boxplot(
  datMeasles$sum_knowledge ~ datMeasles$studyarm,las=2,
  xlab = "",
  ylab = "knowledge",
  main = "Boxplot",
  col = c("darksalmon", "indianred4", "firebrick3", "firebrick1")
)

boxplot(
  datFlu$sum_knowledge ~ datFlu$studyarm,las=2,
  xlab = "",
  ylab = "knowledge",
  main = "Boxplot",
  col = c("lightcyan", "darkslateblue", "dodgerblue")
)


#CHANGELOG ASJ: Boxplot removed for variables that will be analysed as dichotomous variables (trust info & vax intention), 2021-08-06

#### MISSING DATA EXPLORATION ####

# CHANGELOG The code added to determine percentage of the missing data 2021-07-08

don %>%
  naniar::vis_miss(.,
                   sort_miss = TRUE,
                   cluster = TRUE)

don %>%
  naniar::gg_miss_upset(., nsets = 6) # CHANGELOG Descriptive data of missing 2021-07-08

don %>%
  summarise(n_complete = sum(stats::complete.cases(.)),
            pct_complete = 100 * naniar::prop_complete_case(.),
            n_missing = sum(!stats::complete.cases(.)),
            pct_missing = 100 * naniar::prop_miss_case(.))
don %>%
  naniar::miss_case_table() # CHANGELOG Missing numbers

don %>%
  naniar::miss_var_summary() %>%
  dplyr::filter(n_miss > 0) # CHANGELOG nvar missing

don %>%
```

```
  naniar::miss_prop_summary() # CHANGELOG percentage of missing overall, % de var avec missing, % obs avec missingdon %>% #résumé des missing stat + plot avec VIM


HTMLStop() #CHANGELOG ASJ: Stop output html 2021-08-06.


#CHANGELOG ASJ Add descriptive statistic per visualization, 21-09-20.
HTMLStart(outdir=".\\",
          file = "Herdimm Descriptive+ 210921",extension = "html",echo=TRUE, HTMLframe = TRUE)


describeBy(dat$mean_riskperception2_6,dat$viz)
describeBy(dat$mean_emotion,dat$viz)
describeBy(dat$sum_knowledge,dat$viz)
describeBy(dat$mean_C5_calc,dat$viz)
describeBy(dat$mean_C5_conf,dat$viz)
describeBy(dat$mean_C5_cons,dat$viz)
describeBy(dat$mean_C5_comp,dat$viz)
describeBy(dat$mean_C5_coll,dat$viz)
describeBy(dat$trustinfo,dat$viz)
describeBy(dat$vaxintention_1,dat$viz)
describeBy(dat$c19vaxintention_1,dat$viz)
describeBy(dat$riskpercep1_1,dat$viz)


t1<-data.frame(dat$RP1_dicho,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$Htrustinfo,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$vaxintent,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$c19vaxintent,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$mean_C5_calc2,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$mean_C5_conf2,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$mean_C5_comp2,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$mean_C5_cons2,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100

t1<-data.frame(dat$mean_C5_coll2,dat$viz)
table(t1)
prop.table(table(t1), 2) * 100
HTMLStop()

#### MULTIPLE IMPUTATION ####

# Create new data sets for imputed values to make analyses easier

#CHANGELOG ASJ: change list of variables for imputation according to above recoding,
```

```r
datMain4Imputation <-
  subset(
    dat,
    dat$studyarm == "controlgeneric" |
      dat$studyarm == "controlmeasles" |
      dat$studyarm == "controlpertussis" |
      dat$studyarm == "controlflu"|
      dat$studyarm == "herdimmgeneric" |
      dat$studyarm == "herdimmmeasles" |
      dat$studyarm == "herdimmpertussis" |
      dat$studyarm == "herdimmflu",
    select = c(
      viz,
      disease,
     mean_riskperception2_6,
      mean_C5_conf,
      mean_C5_comp,
      mean_C5_cons,
      mean_C5_calc,
      mean_C5_coll,
      mean_emotion,
      sum_knowledge,
      mean_indivhorz,
      mean_indivvertical,
      mean_collhorz,
      mean_collvertical,
      bornincanada,
      language_1,
      language_2,
      Asian_group,
      White_group,
     disability_any,
      genderidentity,
      income,
      edhi,
      age,
      Htrustinfo,
      RP1_dicho,
     vaxintent
    )
  )


datGeneric4Imputation <-
  subset(
    dat,
    dat$studyarm == "controlgeneric" |
      dat$studyarm == "herdimmgeneric" |
      dat$studyarm == "robertkochgeneric" |
      dat$studyarm == "sbsnewsgeneric" & UserLanguage == "EN",
    select = c(
      studyarm,
       UserLanguage,
      mean_riskperception2_6,
      mean_C5_conf,
      mean_C5_comp,
      mean_C5_cons,
      mean_C5_calc,
      mean_C5_coll,
      mean_emotion,
      sum_knowledge,
      mean_indivhorz,
      mean_indivvertical,
      mean_collhorz,
      mean_collvertical,
      bornincanada,
      language_1,
      language_2,
```

```
      Asian_group,
      White_group,
      disability_any,
      genderidentity,
      income,
      edhi,
      age,
      Htrustinfo,
      RP1_dicho,
      vaxintent
    )
  )

datMeasles4Imputation <-
  subset(
    dat,
    dat$studyarm == "controlmeasles" |
      dat$studyarm == "herdimmmeasles" |
      dat$studyarm == "guardianmeasles" |
      dat$studyarm == "theotheredmundmeasles",
    select = c(
      studyarm,
       UserLanguage,
      mean_riskperception2_6,
      mean_C5_conf,
      mean_C5_comp,
      mean_C5_cons,
      mean_C5_calc,
      mean_C5_coll,
      mean_emotion,
      sum_knowledge,
      mean_indivhorz,
      mean_indivvertical,
      mean_collhorz,
      mean_collvertical,
      bornincanada,
      language_1,
      language_2,
      Asian_group,
      White_group,
      disability_any,
      genderidentity,
      income,
      edhi,
      age,
      Htrustinfo,
      RP1_dicho,
      vaxintent
    )
  )

datFlu4Imputation <-
  subset(
    dat,
    dat$studyarm == "controlflu" |
      dat$studyarm == "herdimmflu" |
      dat$studyarm == "publichealthagencycanadaflu",
    select = c(
      studyarm,
       UserLanguage,
      mean_riskperception2_6,
      mean_C5_conf,
      mean_C5_comp,
      mean_C5_cons,
      mean_C5_calc,
      mean_C5_coll,
      mean_emotion,
      sum_knowledge,
      mean_indivhorz,
```

```
      mean_indivvertical,
      mean_collhorz,
      mean_collvertical,
      bornincanada,
      language_1,
      language_2,
      Asian_group,
      White_group,
      disability_any,
      genderidentity,
      income,
      edhi,
      age,
      Htrustinfo,
      RP1_dicho,
      vaxintent
    )
  )

# Drop any unused levels to clean up the dataset

datMain4Imputation$viz <- droplevels(datMain4Imputation$viz)

datGeneric4Imputation$studyarm <- droplevels(datGeneric4Imputation$studyarm)

datMeasles4Imputation$studyarm <- droplevels(datMeasles4Imputation$studyarm)

datFlu4Imputation$studyarm <- droplevels(datFlu4Imputation$studyarm)

# Impute the missing values, using parameters:
# m: Refers to 15 imputed data sets
# maxit: Refers to number of iterations taken to impute missing values: 15
#CHANGELOG ASJ: method removed from mice, use pmm, logreg or polyreg according to datatype, 21-09-20.

#CHANGELOG ASJ:  Set the seed for reproducibility 2021-08-03
seed_for_imp <- 20210920

datMainImputed <-
  mice(
    datMain4Imputation,
    m = 15,
    maxit = 15,

    seed = seed_for_imp,
  )
summary(datMainImputed)

#CHANGELOG ASJ: added to validate imputation. 2021-09-20.
plot(datMainImputed,c("bornincanada","language_1","Asian_group","White_group","disability_any",
                        "genderidentity","income","edhi","age","Htrustinfo","RP1_dicho","vaxintent"))
stripplot(datMainImputed,age)
bwplot(datMainImputed,age)
densityplot(datMainImputed,~age)
densityplot(datMainImputed,~bornincanada)
densityplot(datMainImputed,~language_1)
densityplot(datMainImputed,~Asian_group)
densityplot(datMainImputed,~White_group)
densityplot(datMainImputed,~disability_any)
densityplot(datMainImputed,~genderidentity)
densityplot(datMainImputed,~income)
densityplot(datMainImputed,~edhi)
densityplot(datMainImputed,~Htrustinfo) #only one missing obs
densityplot(datMainImputed,~RP1_dicho)
densityplot(datMainImputed,~vaxintent)
save(datMainImputed,file="datMainImputed.RData")
```

```r


datGenericImputed <-
  mice(
    datGeneric4Imputation,
    m = 15,
    maxit = 15,
    seed = seed_for_imp
  )
summary(datGenericImputed)

#CHANGELOG ASJ: added to validate imputation. 2021-09-20.
plot(datGenericImputed,c("bornincanada","language_1","Asian_group","White_group","disability_any",
                      "genderidentity","income","edhi","age","Htrustinfo","RP1_dicho","vaxintent"))
stripplot(datGenericImputed,age)
bwplot(datGenericImputed,age)
densityplot(datGenericImputed,~age)
densityplot(datGenericImputed,~bornincanada)
densityplot(datGenericImputed,~language_1)
densityplot(datGenericImputed,~Asian_group)
densityplot(datGenericImputed,~White_group)
densityplot(datGenericImputed,~disability_any)
densityplot(datGenericImputed,~genderidentity)
densityplot(datGenericImputed,~income)
densityplot(datGenericImputed,~edhi)
densityplot(datGenericImputed,~Htrustinfo) #only one missing obs
densityplot(datGenericImputed,~RP1_dicho)
densityplot(datGenericImputed,~vaxintent)
save(datGenericImputed,file="datGenericImputed.RData")



datMeaslesImputed <-
  mice(
    datMeasles4Imputation,
    m = 15,
    maxit = 15,
    seed = seed_for_imp
  )
summary(datMeaslesImputed)

#CHANGELOG ASJ: added to validate imputation. 2021-09-20.
plot(datMeaslesImputed,c("bornincanada","language_1","Asian_group","White_group","disability_any",
                         "genderidentity","income","edhi","age","RP1_dicho","vaxintent"))
stripplot(datMeaslesImputed,age)
bwplot(datMeaslesImputed,age)
densityplot(datMeaslesImputed,~age)
densityplot(datMeaslesImputed,~bornincanada)
densityplot(datMeaslesImputed,~language_1) #only 1 missing
densityplot(datMeaslesImputed,~Asian_group)
densityplot(datMeaslesImputed,~White_group)
densityplot(datMeaslesImputed,~disability_any)
densityplot(datMeaslesImputed,~genderidentity)
densityplot(datMeaslesImputed,~income)
densityplot(datMeaslesImputed,~edhi)
densityplot(datMeaslesImputed,~RP1_dicho)
densityplot(datMeaslesImputed,~vaxintent)
save(datMeaslesImputed,file="datMeaslesImputed.RData")

datFluImputed <-
  mice(
    datFlu4Imputation,
    m = 15,
    maxit = 15,
   seed = seed_for_imp
```

```
  )
summary(datFluImputed)

#CHANGELOG ASJ: added to validate imputation. 2021-09-20.
plot(datFluImputed,c("bornincanada","language_1","Asian_group","White_group","disability_any",
                          "genderidentity","income","edhi","age","RP1_dicho","vaxintent"))
stripplot(datFluImputed,age)
bwplot(datFluImputed,age)
densityplot(datFluImputed,~age)
densityplot(datFluImputed,~bornincanada)
densityplot(datFluImputed,~language_1) #only 1 missing
densityplot(datFluImputed,~Asian_group)
densityplot(datFluImputed,~White_group)
densityplot(datFluImputed,~disability_any)
densityplot(datFluImputed,~genderidentity)
densityplot(datFluImputed,~income)
densityplot(datFluImputed,~edhi)
densityplot(datFluImputed,~RP1_dicho)
densityplot(datFluImputed,~vaxintent)
save(datFluImputed,file="datFluImputed.RData")


#### ANALYSES ####

#### RESEARCH QUESTION 1 (PRIMARY): EFFECT OF HERDIMM ON RISK PERCEPTION ####

# + 2-way ANOVA (adjusted for covariates and with moderators) ####

# Recall: Research question 1 (primary): Across 4 vaccine-preventable diseases, does the herdimm
# intervention influence risk perception compared to a control?

#### + Model 1: Check for direct effects of factors without any covariates (Risk perception1_1) ####
# We have made a new variable with transformation. We want to know that when we look at model 1 to model 3
# what we get the significant results e.g., do we get significant results from interaction for all, or between which diseases, or viz,
# if it doesn't change then we use risk perception 2-6 without any transformation, and we will be good with that
# because yes, transformation makes residual better but it doesnt change the conclusion.
# Note: if we do not see changes in the model 1-3 for datMain, we do not use transform in imputation data,
# if there is change then we use transformation variable in the imputation data.

#CHANGELOG. Analysis of RiskPerception2-6 only in anova model, 2021-08-04
#CHANGELOG ASJ: After analysis of residuals, model also run with transformed variable
#CHANGELOG ASJ: results are similar and residuals not that improved, keep original variable
hist(datMain$mean_riskperception2_6)
datMain$mean_riskperception2_6trans<-datMain$mean_riskperception2_6 ** 2
hist(datMain$mean_riskperception2_6trans)
datMain$mean_riskperception2_6transinv<-log(8-datMain$mean_riskperception2_6)
hist(datMain$mean_riskperception2_6transinv)

#### + Model 1: Check for direct effects of factors without any covariates (for risk perception 2 to 6) ####

options(contrasts=c("contr.sum", "contr.poly")) # CHANGELOG: The statement below is added to calculate properly the type III sum of squares with lm 2021-07-27

mRP1_2_6 <- lm(mean_riskperception2_6  ~ viz * disease, data = datMain)
Anova(mRP1_2_6,type="III")
summary(mRP1_2_6)

# Model 2: Check for direct effects of factors with adjustment for other covariates
#CHANGELOG ASJ: change list of covariates after recoding and deletion for <5% categories, 2021-08-04
```

```
mRP2_2_6 <-
  lm(
    mean_riskperception2_6  ~ viz * disease +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMain
  )
Anova(mRP2_2_6, type = "III")
summary(mRP2_2_6)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other
# covariates

mRP3_2_6 <-
  lm(
    mean_riskperception2_6  ~
      viz * disease * mean_indivhorz +
      viz * disease * mean_indivvertical +
      viz * disease * mean_collhorz +
      viz * disease * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMain
  )
Anova(mRP3_2_6, type = "III")
summary(mRP3_2_6)

#### + Adding interaction for high and low education with herdimm vizsualization ####
# CHANGELOG We have added 2 new models to check the interaction between
# high and low education with herdimm vizualization 2021-07-22

#### + Model 4: Check for interaction viz*edhi without any covariates (for risk perception 2 to 6) ####

mRP4_2_6 <- lm(mean_riskperception2_6 ~ viz * disease + viz * edhi, data = datMain)
Anova(mRP4_2_6,type="III")
summary(mRP4_2_6)

# Model 5: Check for interaction viz*edhi with covariates

mRP5_2_6 <-
  lm(
    mean_riskperception2_6  ~ viz * disease +
      viz * edhi +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
```

```r
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      age,
    data = datMain
  )
Anova(mRP5_2_6, type = "III")
summary(mRP5_2_6)


#### + Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

# First, create an assumption-checking function that we can run on each model for efficiency.

checkassumptions <- function(modelname) {
  boxcox(modelname, lambda = seq(-2, 2, 0.1), plotit = T)
  # Extract the residuals from the model
  res <- studres(modelname)
  # Check their normality
  print(kurtosis(res) - 3)
  print(skewness(res))
  hist(res)
  qqnorm(res)
  qqline(res)
  print(ks.test(res, "pnorm", 0, 1))
  print(shapiro.test(res))
  # Homoscedasticity (homogeneity of variance)
  fitted <- fitted(modelname)
  plot(fitted, res)
  abline(h = c(-3, 0, 3))
}

# Then run the function to check the assumptions for each model.

checkassumptions(mRP1_2_6)
checkassumptions(mRP2_2_6)
checkassumptions(mRP3_2_6)
checkassumptions(mRP4_2_6) # CHANGELOG We have added check assumptions function for Model 4, 5 and 6 2021-07-22
checkassumptions(mRP5_2_6)


#### + Interpretation (Risk perception 2 to 6) ####

#CHANGELOG ASJ: only the appropriate section (with or without interaction) is kept for each model, 2021-08-09

# Interpretation of model 1 in the case of an interaction

mRP1_2_6_vd <- emmeans(mRP1_2_6,~viz|disease)
mRP1_2_6_vd
pairs(mRP1_2_6_vd) # Only do multiple comparisons if type 3 test is significant
plot(mRP1_2_6_vd) # Explore options for less ugly plot for publication/thesis

mRP1_2_6_dv <- emmeans(mRP1_2_6,~disease|viz)
mRP1_2_6_dv
pairs(mRP1_2_6_dv) # Only do multiple comparisons if type 3 test is significant
plot(mRP1_2_6_dv)
```

```


# Interpretation of model 2 in the case of an interaction

mRP2_2_6_vd <- emmeans(mRP2_2_6,~viz|disease)
mRP2_2_6_vd
pairs(mRP2_2_6_vd) # Only do multiple comparisons if type 3 test is significant
plot(mRP2_2_6_vd) # Explore options for less ugly plot for publication/thesis

mRP2_2_6_dv <- emmeans(mRP2_2_6,~disease|viz)
mRP2_2_6_dv
pairs(mRP2_2_6_dv) # Only do multiple comparisons if type 3 test is significant
plot(mRP2_2_6_dv)

# Interpretation of model 3 moderating effects
#CHANGELOG ASJ: change the focus of the interpretation 2021-08-11.

emmip(mRP3_2_6, disease ~ mean_indivhorz, cov.reduce = range)
emmip(mRP3_2_6, viz ~ mean_indivvertical, cov.reduce = range)
emmip(mRP3_2_6, viz ~ mean_collhorz, cov.reduce = range)

emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(2.5)))
pairs(emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(2.5))))
emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(5.0)))
pairs(emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(5.0))))
emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(7.5)))
pairs(emmeans(mRP3_2_6,  ~ viz, at=list(mean_indivvertical=(7.5))))

# Interpretation of model 4 in the case of an interaction viz*edhi
# CHANGELOG We have added interpretation of model 4 2021-07-22

mRP4_2_6_ve <- emmeans(mRP4_2_6,~viz|edhi)
mRP4_2_6_ve
pairs(mRP4_2_6_ve) # Only do multiple comparisons if type 3 test is significant
plot(mRP4_2_6_ve) # Explore options for less ugly plot for publication/thesis

mRP4_2_6_ev <- emmeans(mRP4_2_6,~edhi|viz)
mRP4_2_6_ev
pairs(mRP4_2_6_ev) # Only do multiple comparisons if type 3 test is significant
plot(mRP4_2_6_ev)


# Interpretation of model 5 in the case of an interaction viz*edhi
#No interaction


#### + Analyses on imputed data ####

#CHANGELOG ASJ: No missing data for mean_riskperception2_6, no model with imputation performed, 21-09-20.


#### RESEARCH QUESTION 2: EFFECT OF HERDIMM ON OTHER OUTCOMES ####

# Recall: Research question 2 (secondary): Across 4 vaccine-preventable diseases, does the herdimm
# intervention influence other outcomes (5C scale, emotions, trust, knowledge, behavioural
# intentions) compared to a control?

# The same analysis as above will be done for the secondary outcomes: mean_C5, mean_emotions,
# sum_knowledge, and behavioural intentions. For trustinfo the same analysis will be done if the
# distribution has enough variance (2-way ANOVA). However, if it's not the case the variable will be
# dichotomised into 1 to 4 (low trust) and 5 to 7 (high trust) and we will run a logistic
# regression.
```

```r


#### +++ Emotions ####

options(contrasts=c("contr.sum", "contr.poly")) # CHANGELOG: The statement below is added to calculate properly the type III sum of squares with lm 2021-07-27

#### + Model 1: Check for direct effects of factors without any covariates (Emotions) ####
mEmotions_1 <- lm(mean_emotion  ~ viz * disease, data = datMain)
Anova(mEmotions_1,type="III")
summary(mEmotions_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates

mEmotions_2 <-
  lm(
    mean_emotion  ~ viz * disease +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMain
  )
Anova(mEmotions_2, type = "III")
summary(mEmotions_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other
# covariates

mEmotions_3 <-
  lm(
    mean_emotion  ~
      viz * disease * mean_indivhorz +
      viz * disease * mean_indivvertical +
      viz * disease * mean_collhorz +
      viz * disease * mean_collvertical +
      bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMain
  )
Anova(mEmotions_3, type = "III")
summary(mEmotions_3)

# + Adding interaction for high and low education with herdimm vizsualization
# CHANGELOG We have added 2 new models to check the interaction between
# high and low education with herdimm vizualization 2021-07-26

#### + Model 4: Check for interaction edhi*viz without any covariates (for Emotions) ####

```

```r
mEmotions_4 <- lm(mean_emotion  ~ viz * disease + viz * edhi, data = datMain)
Anova(mEmotions_4,type="III")
summary(mEmotions_4)

# Model 5: Check for interaction edhi*viz with covariates

mEmotions_5 <-
  lm(
    mean_emotion  ~ viz * disease +
      viz * edhi +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      age,
    data = datMain
  )
Anova(mEmotions_5, type = "III")
summary(mEmotions_5)


#### + Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mEmotions_1)
checkassumptions(mEmotions_2)
checkassumptions(mEmotions_3)
checkassumptions(mEmotions_4)
checkassumptions(mEmotions_5)


#### + Interpretation (Emotions) ####

# Interpretation of model 1 in the case of no interaction

memory.size(max = TRUE)

mEmotions_1_v <- emmeans(mEmotions_1,~viz)
mEmotions_1_v
plot(mEmotions_1_v)

mEmotions_1_d <- emmeans(mEmotions_1,  ~ disease)
mEmotions_1_d
pairs(mEmotions_1_d) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_1_d)


# Interpretation of model 2 in the case of no interaction

mEmotions_2_v <- emmeans(mEmotions_2,~viz)
mEmotions_2_v
plot(mEmotions_2_v)

mEmotions_2_d <- emmeans(mEmotions_2,  ~ disease)
mEmotions_2_d
pairs(mEmotions_2_d) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_2_d)

```

```r


# Interpretation of model 3 in the case of an interaction with moderator
emmip(mEmotions_3, viz*disease ~ mean_collhorz, cov.reduce = range)

# Interpretation of model 4 in the case of an interaction viz edhi
# CHANGELOG We have added interpretation of model 4 2021-07-26

mEmotions_4_ve <- emmeans(mEmotions_4,~viz|edhi)
mEmotions_4_ve
pairs(mEmotions_4_ve) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_4_ve)

mEmotions_4_ev <- emmeans(mEmotions_4,~edhi|viz)
mEmotions_4_ev
pairs(mEmotions_4_ev) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_4_ev)

# Interpretation of model 5 in the case of an interaction viz edhi
# CHANGELOG We have added interpretation of model 5 2021-07-26

mEmotions_5_ve <- emmeans(mEmotions_5,~viz|edhi)
mEmotions_5_ve
pairs(mEmotions_5_ve) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_5_ve)

mEmotions_5_ev <- emmeans(mEmotions_5,~edhi|viz)
mEmotions_5_ev
pairs(mEmotions_5_ev) # Only do multiple comparisons if type 3 test is significant
plot(mEmotions_5_ev)


# Now we run the models on imputed data.

#CHANGELOG ASJ: No missing data for mean_emotion, no model with imputation performed, 21-09-20.


#### +++ Knowledge ####
#### + Model 1: Check for direct effects of factors without any covariates (Knowledge) ####

options(contrasts=c("contr.sum", "contr.poly")) # CHANGELOG: The statement below is added to calculate properly the type III sum of squares with lm 2021-07-27

mKnowledge_1 <- lm(sum_knowledge  ~ viz * disease, data = datMain)
Anova(mKnowledge_1,type="III")
summary(mKnowledge_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates

mKnowledge_2 <-
  lm(
    sum_knowledge  ~ viz * disease +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMain
```

```r
    )
  Anova(mKnowledge_2, type = "III")
  summary(mKnowledge_2)

  # Model 3: Check for moderating effects of individualism & collectivism with adjustment for other
  # covariates

  mKnowledge_3 <-
    lm(
      sum_knowledge  ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain
    )
  Anova(mKnowledge_3, type = "III")
  summary(mKnowledge_3)

  # + Adding interaction for high and low education with herdimm vizsualization
  # CHANGELOG We have added 2 new models to check the interaction between
  # high and low education with herdimm vizualization 2021-07-26

  #### + Model 4: Check for interaction viz*edhi without any covariates (for Knowledge) ####

  mKnowledge_4 <- lm(sum_knowledge  ~ viz * disease + viz * edhi, data = datMain)
  Anova(mKnowledge_4,type="III")
  summary(mKnowledge_4)

  # Model 5: Check for interaction disease*viz with covariates

  mKnowledge_5 <-
    lm(
      sum_knowledge  ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain
    )
  Anova(mKnowledge_5, type = "III")
  summary(mKnowledge_5)


  #### + Validating assumptions ####

  # Check to see if the assumptions of the models are respected. If not, we will
```

```
transform the
# outcome.

checkassumptions(mKnowledge_1)
checkassumptions(mKnowledge_2)
checkassumptions(mKnowledge_3)
checkassumptions(mKnowledge_4)
checkassumptions(mKnowledge_5)

#### + Interpretation (Knowledge) ####

# Interpretation of model 1 in the case of no interaction

memory.size(max = TRUE)

mKnowledge_1_v <- emmeans(mKnowledge_1,~viz)
mKnowledge_1_v
plot(mKnowledge_1_v)

mKnowledge_1_d <- emmeans(mKnowledge_1,  ~ disease)
mKnowledge_1_d
pairs(mKnowledge_1_d) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_1_d)


# Interpretation of model 2 in the case of an interaction

mKnowledge_2_vd <- emmeans(mKnowledge_2,~viz|disease)
mKnowledge_2_vd
pairs(mKnowledge_2_vd) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_2_vd)

mKnowledge_2_dv <- emmeans(mKnowledge_2,~disease|viz)
mKnowledge_2_dv
pairs(mKnowledge_2_dv) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_2_dv)

mKnowledge_3_d
pairs(mKnowledge_3_d) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_3_d)

# Interpretation of model 3
emmip(mKnowledge_3, viz ~ mean_indivvertical, cov.reduce = range)


# Interpretation of model 4 in the case of an interaction viz edhi
# CHANGELOG We have added interpretation of model 4 2021-07-26

mKnowledge_4_ve <- emmeans(mKnowledge_4,~viz|edhi)
mKnowledge_4_ve
pairs(mKnowledge_4_ve) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_4_ve)

mKnowledge_4_ev <- emmeans(mKnowledge_4,~edhi|viz)
mKnowledge_4_ev
pairs(mKnowledge_4_ev) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_4_ev)

# Interpretation of model 5 in case of an interaction viz*edhi
# CHANGELOG We have added interpretation of model 5 2021-07-26

mKnowledge_5_ve <- emmeans(mKnowledge_5,~viz|edhi)
mKnowledge_5_ve
pairs(mKnowledge_5_ve) # Only do multiple comparisons if type 3 test is significant
plot(mKnowledge_5_ve)

mKnowledge_5_ev <- emmeans(mKnowledge_5,~edhi|viz)
mKnowledge_5_ev
pairs(mKnowledge_5_ev) # Only do multiple comparisons if type 3 test is significant
```

```r
plot(mKnowledge_5_ev)

# Now we run the models on imputed data.


#CHANGELOG ASJ: No missing data for sum_knowledge, no model with imputation performed, 21-09-20.


#CHANGELOG ASJ: Add logistic regression models for dichotomous outcomes, 2021-08-05.



#### +++ Logistic regression for RP1 ####

#### + Model 1: Check for direct effects of factors without any covariates ####



options(contrasts=c("contr.sum", "contr.poly"))
rlog_RP1_1 <- glm(relevel(RP1_dicho,ref="Low")  ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_RP1_1,type="III")
summary(rlog_RP1_1)

# Interpretation of model 1
mRP1_v <- emmeans(rlog_RP1_1,~viz,type = "response") #proportions
mRP1_v
plot(mRP1_v)
confint(pairs(mRP1_v,reverse = TRUE)) #Odds ratios

mRP1_d <- emmeans(rlog_RP1_1,  ~ disease, type = "response") #proportions
mRP1_d
plot(mRP1_d)
confint(pairs(mRP1_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_RP1_2 <-
  glm(relevel(RP1_dicho,ref="Low")  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_RP1_2, type = "III")
summary(rlog_RP1_2)


# Interpretation of model 2
mRP1_2_v <- emmeans(rlog_RP1_2,~viz,type = "response")
mRP1_2_v
plot(mRP1_2_v)
confint(pairs(mRP1_2_v,reverse = TRUE))

mRP1_2_d <- emmeans(rlog_RP1_2,  ~ disease, type = "response")
mRP1_2_d
```

```r
plot(mRP1_2_d)
confint(pairs(mRP1_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_RP1_3 <-
  glm(relevel(RP1_dicho,ref="Low")  ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_RP1_3, type = "III")
summary(rlog_RP1_3)

## Interpretation of model 3 : no interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_RP1_4 <- glm(relevel(RP1_dicho,ref="Low")   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_RP1_4,type="III")
summary(rlog_RP1_4)

# Interpretation of model 4
mRP1_4_v <- emmeans(rlog_RP1_4,~viz|edhi,type = "response")
mRP1_4_v
confint(pairs(mRP1_4_v,reverse = TRUE))
pairs(mRP1_4_v,reverse = TRUE,type = "response")
plot(mRP1_4_v)

mRP1_4_e <- emmeans(rlog_RP1_4,  ~ edhi|viz, type = "response")
mRP1_4_e
confint(pairs(mRP1_4_e,reverse = TRUE))
pairs(mRP1_4_e,reverse = TRUE, type = "reponse")
plot(mRP1_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_RP1_5 <-
  glm(relevel(RP1_dicho,ref="Low")   ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
```

```r
  )
Anova(rlog_RP1_5, type = "III")
summary(rlog_RP1_5)

# Interpretation of model 5
mRP1_5_v <- emmeans(rlog_RP1_5,~viz|edhi,type = "response")
mRP1_5_v
confint(pairs(mRP1_5_v,reverse = TRUE))
pairs(mRP1_5_v,reverse = TRUE,type = "response")
plot(mRP1_5_v)

mRP1_5_e <- emmeans(rlog_RP1_5,  ~ edhi|viz, type = "response")
mRP1_5_e
confint(pairs(mRP1_5_e,reverse = TRUE))
pairs(mRP1_5_e,reverse = TRUE, type = "reponse")
plot(mRP1_5_e)


# Now we run the models on imputed data.

#### + Imputed Model 1: Check for direct effects of factors without any covariates ####

imprlog_RP1_1 <-
  with(data = datMainImputed,
       exp = glm(relevel(RP1_dicho,ref="Low") ~ viz * disease,family = binomial("logit")))
combineimprlog_RP1_1 <- pool(imprlog_RP1_1)
summary(combineimprlog_RP1_1)


#### +++ Logistic regression for C19 vax intention ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c19vi_1 <- glm(relevel(c19vaxintent,ref="Low")  ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c19vi_1,type="III")
summary(rlog_c19vi_1)

# Interpretation of model 1
mc19vi_1_v <- emmeans(rlog_c19vi_1,~viz,type = "response") #proportions
mc19vi_1_v
plot(mc19vi_1_v)
confint(pairs(mc19vi_1_v,reverse = TRUE)) #Odds ratios

mc19vi_1_d <- emmeans(rlog_c19vi_1,  ~ disease, type = "response") #proportions
mc19vi_1_d
plot(mc19vi_1_d)
confint(pairs(mc19vi_1_d,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_c19vi_2 <-
  glm(relevel(c19vaxintent,ref="Low")  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
```

```
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c19vi_2, type = "III")
summary(rlog_c19vi_2)


# Interpretation of model 2
mc19vi_2_v <- emmeans(rlog_c19vi_2,~viz,type = "response")
mc19vi_2_v
plot(mc19vi_2_v)
confint(pairs(mc19vi_2_v,reverse = TRUE))

mc19vi_2_d <- emmeans(rlog_c19vi_2,  ~ disease, type = "response")
mc19vi_2_d
plot(mc19vi_2_d)
confint(pairs(mc19vi_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c19vi_3 <-
  glm(relevel(c19vaxintent,ref="Low")  ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c19vi_3, type = "III")
summary(rlog_c19vi_3)

# Interpretation of model 3
emmip(rlog_c19vi_3, viz ~ mean_collvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c19vi_4 <- glm(relevel(c19vaxintent,ref="Low")   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_c19vi_4,type="III")
summary(rlog_c19vi_4)

# Interpretation of model 4  (no significant interaction viz edhi)
mc19vi_4_v <- emmeans(rlog_c19vi_4,~viz|edhi,type = "response")
mc19vi_4_v
confint(pairs(mc19vi_4_v,reverse = TRUE))
plot(mc19vi_4_v)

mc19vi_4_e <- emmeans(rlog_c19vi_4,  ~ edhi|viz, type = "response")
mc19vi_4_e
confint(pairs(mc19vi_4_e,reverse = TRUE))
plot(mc19vi_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c19vi_5 <-
  glm(relevel(c19vaxintent,ref="Low")   ~ viz * disease +
        viz * edhi +
```

```r
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c19vi_5, type = "III")
summary(rlog_c19vi_5)

# Interpretation of model 5
mc19vi_5_v <- emmeans(rlog_c19vi_5,~viz|edhi,type = "response")
mc19vi_5_v
confint(pairs(mc19vi_5_v,reverse = TRUE))
plot(mc19vi_5_v)

mc19vi_5_e <- emmeans(rlog_c19vi_5,  ~ edhi|viz, type = "response")
mc19vi_5_e
confint(pairs(mc19vi_5_e,reverse = TRUE))
plot(mc19vi_5_e)


# Now we run the models on imputed data.

#CHANGELOG ASJ: No imputation for c19vaxintention, since 51% missing values, 21-09-20.



#### +++ Logistic regression for vax intention ####

#In order for the pairs function to work properly, relevel function must be removed from the model.
datMain$vaxintentRel<-relevel(datMain$vaxintent,ref="Low")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_vi_1 <- glm(vaxintentRel  ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_vi_1,type="III")
summary(rlog_vi_1)

# Interpretation of model 1
mvi_1_v <- emmeans(rlog_vi_1,~viz,type = "response") #proportions
mvi_1_v
plot(mvi_1_v)
confint(pairs(mvi_1_v,reverse = TRUE)) #Odds ratios

mvi_1_d <- emmeans(rlog_vi_1,  ~ disease, type = "response") #proportions
mvi_1_d
plot(mvi_1_d)
pairs(mvi_1_d,reverse = TRUE,type = "response")
confint(pairs(mvi_1_d,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_vi_2 <-
  glm(vaxintentRel  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
```

```r
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_vi_2, type = "III")
summary(rlog_vi_2)


# Interpretation of model 2
mvi_2_v <- emmeans(rlog_vi_2,~viz,type = "response")
mvi_2_v
plot(mvi_2_v)
confint(pairs(mvi_2_v,reverse = TRUE))

mvi_2_d <- emmeans(rlog_vi_2,  ~ disease, type = "response")
mvi_2_d
plot(mvi_2_d)
confint(pairs(mvi_2_d,reverse = TRUE))
pairs(mvi_2_d,reverse=TRUE,type="response")

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_vi_3 <-
  glm(vaxintentRel  ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_vi_3, type = "III")
summary(rlog_vi_3)

# Interpretation of model 3
emmip(rlog_vi_3, disease ~ mean_indivvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_vi_4 <- glm(vaxintentRel   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_vi_4,type="III")
summary(rlog_vi_4)

# Interpretation of model 4 (no significant interaction viz edhi)
mvi_4_v <- emmeans(rlog_vi_4,~viz|edhi,type = "response")
mvi_4_v
confint(pairs(mvi_4_v,reverse = TRUE))
plot(mvi_4_v)

mvi_4_e <- emmeans(rlog_vi_4,  ~ edhi|viz, type = "response")
```

```
mvi_4_e
confint(pairs(mvi_4_e,reverse = TRUE))
plot(mvi_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_vi_5 <-
  glm(vaxintentRel   ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        borneincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_vi_5, type = "III")
summary(rlog_vi_5)

# Interpretation of model 5 ( no significant interaction)
mvi_5_v <- emmeans(rlog_vi_5,~viz|edhi,type = "response")
mvi_5_v
confint(pairs(mvi_5_v,reverse = TRUE))
plot(mvi_5_v)

mvi_5_e <- emmeans(rlog_vi_5,  ~ edhi|viz, type = "response")
mvi_5_e
confint(pairs(mvi_5_e,reverse = TRUE))
plot(mvi_5_e)


# Now we run the models on imputed data.

#### + Imputed Model 1: Check for direct effects of factors without any covariates ####

imprlog_vi_1 <-
  with(data = datMainImputed,
       exp = glm(relevel(vaxintent,ref="Low") ~ viz * disease,family =
       binomial("logit")))
combineimprlog_vi_1 <- pool(imprlog_vi_1)
summary(combineimprlog_vi_1)


#### +++ Logistic regression for Trust info ####

datMain$Htrustinfo<-relevel(datMain$Htrustinfo,ref="No")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_ti_1 <- glm(Htrustinfo  ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_ti_1,type="III")
summary(rlog_ti_1)

# Interpretation of model 1
mti_1_v <- emmeans(rlog_ti_1,~viz|disease,type = "response") #proportions
mti_1_v
plot(mti_1_v)
confint(pairs(mti_1_v,reverse = TRUE)) #Odds ratios
pairs(mti_1_v,reverse=TRUE, type = "response")
```

```

mti_1_d <- emmeans(rlog_ti_1,  ~ disease|viz, type = "response") #proportions
mti_1_d
plot(mti_1_d)
pairs(mti_1_d,reverse = TRUE,type = "response")
confint(pairs(mti_1_d,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_ti_2 <-
  glm(Htrustinfo  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_ti_2, type = "III")
summary(rlog_ti_2)


# Interpretation of model 2
mti_2_v <- emmeans(rlog_ti_2,~viz|disease,type = "response")
mti_2_v
plot(mti_2_v)
confint(pairs(mti_2_v,reverse = TRUE))
pairs(mti_2_v,reverse=TRUE, type = "response")

mti_2_d <- emmeans(rlog_ti_2,  ~ disease|viz, type = "response")
mti_2_d
plot(mti_2_d)
confint(pairs(mti_2_d,reverse = TRUE))
pairs(mti_2_d,reverse=TRUE,type="response")

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_ti_3 <-
  glm(Htrustinfo  ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_ti_3, type = "III")
summary(rlog_ti_3)

#No interaction with moderators

```

```

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_ti_4 <- glm(Htrustinfo   ~ viz * disease + viz * edhi, data =
datMain,family=binomial("logit"))
Anova(rlog_ti_4,type="III")
summary(rlog_ti_4)

# Interpretation of model 4 (no significant interaction viz edhi)
mti_4_v <- emmeans(rlog_ti_4,~viz|edhi,type = "response")
mti_4_v
confint(pairs(mti_4_v,reverse = TRUE))
plot(mti_4_v)

mti_4_e <- emmeans(rlog_ti_4,  ~ edhi|viz, type = "response")
confint(pairs(mti_4_e,reverse = TRUE))
plot(mti_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_ti_5 <-
  glm(Htrustinfo   ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_ti_5, type = "III")
summary(rlog_ti_5)

# Interpretation of model 5 ( no significant interaction)
mti_5_v <- emmeans(rlog_ti_5,~viz|edhi,type = "response")
mti_5_v
confint(pairs(mti_5_v,reverse = TRUE))
plot(mti_5_v)

mti_5_e <- emmeans(rlog_ti_5,  ~ edhi|viz, type = "response")
mti_5_e
confint(pairs(mti_5_e,reverse = TRUE))
plot(mti_5_e)


# Now we run the models on imputed data.

#### + Imputed Model 1: Check for direct effects of factors without any covariates ####


imprlog_ti_1 <-
  with(data = datMainImputed,
       exp = glm(relevel(Htrustinfo,ref="No") ~ viz * disease,family =
       binomial("logit")))
combineimprlog_ti_1 <- pool(imprlog_ti_1)
summary(combineimprlog_ti_1)

#### +++ Logistic regression for Mean C5 Confidence ####

datMain$mean_C5_conf2<-relevel(as.factor(datMain$mean_C5_conf2),ref="0")
```

```r

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c5conf_1 <- glm(mean_C5_conf2 ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c5conf_1,type="III")
summary(rlog_c5conf_1)

# Interpretation of model 1
mc5conf_v <- emmeans(rlog_c5conf_1,~viz,type = "response") #proportions
mc5conf_v
plot(mc5conf_v)
confint(pairs(mc5conf_v,reverse = TRUE)) #Odds ratios

mc5conf_d <- emmeans(rlog_c5conf_1,  ~ disease, type = "response") #proportions
mc5conf_d
plot(mc5conf_d)
confint(pairs(mc5conf_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_c5conf_2 <-
  glm(mean_C5_conf2  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5conf_2, type = "III")
summary(rlog_c5conf_2)


# Interpretation of model 2
mc5conf_2_v <- emmeans(rlog_c5conf_2,~viz,type = "response")
mc5conf_2_v
plot(mc5conf_2_v)
confint(pairs(mc5conf_2_v,reverse = TRUE))

mc5conf_2_d <- emmeans(rlog_c5conf_2,  ~ disease, type = "response")
mc5conf_2_d
plot(mc5conf_2_d)
confint(pairs(mc5conf_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c5conf_3 <-
  glm(mean_C5_conf2 ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
```

```r
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5conf_3, type = "III")
summary(rlog_c5conf_3)

## Interpretation of model 3
emmip(rlog_c5conf_3, viz ~ mean_collvertical , cov.reduce = range, type = "response")



#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c5conf_4 <- glm(mean_C5_conf2   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_c5conf_4,type="III")
summary(rlog_c5conf_4)

# Interpretation of model 4
mc5conf_4_v <- emmeans(rlog_c5conf_4,~viz|edhi,type = "response")
mc5conf_4_v
confint(pairs(mc5conf_4_v,reverse = TRUE))
pairs(mc5conf_4_v,reverse = TRUE,type = "response")
plot(mc5conf_4_v)

mc5conf_4_e <- emmeans(rlog_c5conf_4,  ~ edhi|viz, type = "response")
mc5conf_4_e
confint(pairs(mc5conf_4_e,reverse = TRUE))
pairs(mc5conf_4_e,reverse = TRUE, type = "reponse")
plot(mc5conf_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c5conf_5 <-
  glm(mean_C5_conf2  ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5conf_5, type = "III")
summary(rlog_c5conf_5)

# Interpretation of model 5
mc5conf_5_v <- emmeans(rlog_c5conf_5,~viz|edhi,type = "response")
mc5conf_5_v
confint(pairs(mc5conf_5_v,reverse = TRUE))
pairs(mc5conf_5_v,reverse = TRUE,type = "response")
plot(mc5conf_5_v)

mc5conf_5_e <- emmeans(rlog_c5conf_5,  ~ edhi|viz, type = "response")
mc5conf_5_e
```

```r
confint(pairs(mc5conf_5_e,reverse = TRUE))
pairs(mc5conf_5_e,reverse = TRUE, type = "reponse")
plot(mc5conf_5_e)


# Now we run the models on imputed data.
#CHANGELOG ASJ: No missing data for mean_C5_conf2, no model with imputation performed, 21-09-20.

#### +++ Logistic regression for Mean C5 Complacency ####

datMain$mean_C5_comp2<-relevel(as.factor(datMain$mean_C5_comp2),ref="0")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c5comp_1 <- glm(mean_C5_comp2 ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c5comp_1,type="III")
summary(rlog_c5comp_1)

# Interpretation of model 1
mc5comp_v <- emmeans(rlog_c5comp_1,~viz,type = "response") #proportions
mc5comp_v
plot(mc5comp_v)
confint(pairs(mc5comp_v,reverse = TRUE)) #Odds ratios

mc5comp_d <- emmeans(rlog_c5comp_1,  ~ disease, type = "response") #proportions
mc5comp_d
plot(mc5comp_d)
confint(pairs(mc5comp_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_c5comp_2 <-
  glm(mean_C5_comp2  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5comp_2, type = "III")
summary(rlog_c5comp_2)


# Interpretation of model 2
mc5comp_2_v <- emmeans(rlog_c5comp_2,~viz,type = "response")
mc5comp_2_v
plot(mc5comp_2_v)
confint(pairs(mc5comp_2_v,reverse = TRUE))

mc5comp_2_d <- emmeans(rlog_c5comp_2,  ~ disease, type = "response")
mc5comp_2_d
plot(mc5comp_2_d)
confint(pairs(mc5comp_2_d,reverse = TRUE))


```

```r
# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c5comp_3 <-
  glm(mean_C5_comp2 ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5comp_3, type = "III")
summary(rlog_c5comp_3)

## Interpretation of model 3 : no moderating effect to interpret



#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c5comp_4 <- glm(mean_C5_comp2   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_c5comp_4,type="III")
summary(rlog_c5comp_4)

# Interpretation of model 4
mc5comp_4_v <- emmeans(rlog_c5comp_4,~viz|edhi,type = "response")
mc5comp_4_v
confint(pairs(mc5comp_4_v,reverse = TRUE))
pairs(mc5comp_4_v,reverse = TRUE,type = "response")
plot(mc5comp_4_v)

mc5comp_4_e <- emmeans(rlog_c5comp_4,  ~ edhi|viz, type = "response")
mc5comp_4_e
confint(pairs(mc5comp_4_e,reverse = TRUE))
pairs(mc5comp_4_e,reverse = TRUE, type = "reponse")
plot(mc5comp_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c5comp_5 <-
  glm(mean_C5_comp2  ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5comp_5, type = "III")
```

```
summary(rlog_c5comp_5)

# Interpretation of model 5
mc5comp_5_v <- emmeans(rlog_c5comp_5,~viz|edhi,type = "response")
mc5comp_5_v
confint(pairs(mc5comp_5_v,reverse = TRUE))
plot(mc5comp_5_v)

mc5comp_5_e <- emmeans(rlog_c5comp_5,  ~ edhi|viz, type = "response")
mc5comp_5_e
confint(pairs(mc5comp_5_e,reverse = TRUE))
plot(mc5comp_5_e)


# Now we run the models on imputed data.
#CHANGELOG ASJ: No missing data for mean_C5_comp2, no model with imputation performed, 21-09-20.




#### +++ Logistic regression for Mean C5 Constraints ####

datMain$mean_C5_cons2<-relevel(as.factor(datMain$mean_C5_cons2),ref="0")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c5cons_1 <- glm(mean_C5_cons2 ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c5cons_1,type="III")
summary(rlog_c5cons_1)

# Interpretation of model 1
mc5cons_v <- emmeans(rlog_c5cons_1,~viz,type = "response") #proportions
mc5cons_v
plot(mc5cons_v)
confint(pairs(mc5cons_v,reverse = TRUE)) #Odds ratios

mc5cons_d <- emmeans(rlog_c5cons_1,  ~ disease, type = "response") #proportions
mc5cons_d
plot(mc5cons_d)
confint(pairs(mc5cons_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_c5cons_2 <-
  glm(mean_C5_cons2  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5cons_2, type = "III")
summary(rlog_c5cons_2)

```

```
# Interpretation of model 2
mc5cons_2_v <- emmeans(rlog_c5cons_2,~viz,type = "response")
mc5cons_2_v
plot(mc5cons_2_v)
confint(pairs(mc5cons_2_v,reverse = TRUE))

mc5cons_2_d <- emmeans(rlog_c5cons_2,  ~ disease, type = "response")
mc5cons_2_d
plot(mc5cons_2_d)
confint(pairs(mc5cons_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c5cons_3 <-
  glm(mean_C5_cons2 ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5cons_3, type = "III")
summary(rlog_c5cons_3)

## Interpretation of model 3
emmip(rlog_c5cons_3, viz ~ mean_indivvertical , cov.reduce = range, type = "response")



#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c5cons_4 <- glm(mean_C5_cons2   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_c5cons_4,type="III")
summary(rlog_c5cons_4)

# Interpretation of model 4
mc5cons_4_v <- emmeans(rlog_c5cons_4,~viz|edhi,type = "response")
mc5cons_4_v
confint(pairs(mc5cons_4_v,reverse = TRUE))
pairs(mc5cons_4_v,reverse = TRUE,type = "response")
plot(mc5cons_4_v)

mc5cons_4_e <- emmeans(rlog_c5cons_4,  ~ edhi|viz, type = "response")
mc5cons_4_e
confint(pairs(mc5cons_4_e,reverse = TRUE))
pairs(mc5cons_4_e,reverse = TRUE, type = "reponse")
plot(mc5cons_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c5cons_5 <-
  glm(mean_C5_cons2  ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
```

```
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5cons_5, type = "III")
summary(rlog_c5cons_5)

# Interpretation of model 5
mc5cons_5_v <- emmeans(rlog_c5cons_5,~viz|edhi,type = "response")
mc5cons_5_v
confint(pairs(mc5cons_5_v,reverse = TRUE))
pairs(mc5cons_5_v,reverse = TRUE,type = "response")
plot(mc5cons_5_v)

mc5cons_5_e <- emmeans(rlog_c5cons_5,  ~ edhi|viz, type = "response")
mc5cons_5_e
confint(pairs(mc5cons_5_e,reverse = TRUE))
pairs(mc5cons_5_e,reverse = TRUE, type = "reponse")
plot(mc5cons_5_e)


# Now we run the models on imputed data.
#CHANGELOG ASJ: No missing data for mean_C5_cons2, no model with imputation performed, 21-09-20.



#### +++ Logistic regression for Mean C5 Calculation ####

datMain$mean_C5_calc2<-relevel(as.factor(datMain$mean_C5_calc2),ref="0")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c5calc_1 <- glm(mean_C5_calc2 ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c5calc_1,type="III")
summary(rlog_c5calc_1)

# Interpretation of model 1
mc5calc_v <- emmeans(rlog_c5calc_1,~viz,type = "response") #proportions
mc5calc_v
plot(mc5calc_v)
confint(pairs(mc5calc_v,reverse = TRUE)) #Odds ratios

mc5calc_d <- emmeans(rlog_c5calc_1,  ~ disease, type = "response") #proportions
mc5calc_d
plot(mc5calc_d)
confint(pairs(mc5calc_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates

rlog_c5calc_2 <-
  glm(mean_C5_calc2  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
```

```r
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5calc_2, type = "III")
summary(rlog_c5calc_2)


# Interpretation of model 2
mc5calc_2_v <- emmeans(rlog_c5calc_2,~viz,type = "response")
mc5calc_2_v
plot(mc5calc_2_v)
confint(pairs(mc5calc_2_v,reverse = TRUE))

mc5calc_2_d <- emmeans(rlog_c5calc_2,  ~ disease, type = "response")
mc5calc_2_d
plot(mc5calc_2_d)
confint(pairs(mc5calc_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c5calc_3 <-
  glm(mean_C5_calc2 ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5calc_3, type = "III")
summary(rlog_c5calc_3)

## Interpretation of model 3 : no moderation effect to interpret



#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c5calc_4 <- glm(mean_C5_calc2   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
Anova(rlog_c5calc_4,type="III")
summary(rlog_c5calc_4)

# Interpretation of model 4
mc5calc_4_v <- emmeans(rlog_c5calc_4,~viz|edhi,type = "response")
mc5calc_4_v
confint(pairs(mc5calc_4_v,reverse = TRUE))
plot(mc5calc_4_v)

mc5calc_4_e <- emmeans(rlog_c5calc_4,  ~ edhi|viz, type = "response")
mc5calc_4_e
confint(pairs(mc5calc_4_e,reverse = TRUE))
```

```
plot(mc5calc_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c5calc_5 <-
 glm(mean_C5_calc2  ~ viz * disease +
    viz * edhi +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    genderidentity +
    income +
    age,
   data = datMain,family=binomial("logit")
 )
Anova(rlog_c5calc_5, type = "III")
summary(rlog_c5calc_5)

# Interpretation of model 5
mc5calc_5_v <- emmeans(rlog_c5calc_5,~viz|edhi,type = "response")
mc5calc_5_v
confint(pairs(mc5calc_5_v,reverse = TRUE))
plot(mc5calc_5_v)

mc5calc_5_e <- emmeans(rlog_c5calc_5,  ~ edhi|viz, type = "response")
mc5calc_5_e
confint(pairs(mc5calc_5_e,reverse = TRUE))
plot(mc5calc_5_e)


# Now we run the models on imputed data.
#CHANGELOG ASJ: No missing data for mean_C5_calc2, no model with imputation performed, 21-09-20.



#### +++ Logistic regression for Mean C5 Collective Response ####

datMain$mean_C5_coll2<-relevel(as.factor(datMain$mean_C5_coll2),ref="0")

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlog_c5coll_1 <- glm(mean_C5_coll2 ~ viz * disease, data = datMain,family=binomial("logit"))
Anova(rlog_c5coll_1,type="III")
summary(rlog_c5coll_1)

# Interpretation of model 1
mc5coll_v <- emmeans(rlog_c5coll_1,~viz,type = "response") #proportions
mc5coll_v
plot(mc5coll_v)
confint(pairs(mc5coll_v,reverse = TRUE)) #Odds ratios

mc5coll_d <- emmeans(rlog_c5coll_1,  ~ disease, type = "response") #proportions
mc5coll_d
plot(mc5coll_d)
confint(pairs(mc5coll_d,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates
```

```

rlog_c5coll_2 <-
  glm(mean_C5_coll2  ~ viz * disease +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5coll_2, type = "III")
summary(rlog_c5coll_2)


# Interpretation of model 2
mc5coll_2_v <- emmeans(rlog_c5coll_2,~viz,type = "response")
mc5coll_2_v
plot(mc5coll_2_v)
confint(pairs(mc5coll_2_v,reverse = TRUE))

mc5coll_2_d <- emmeans(rlog_c5coll_2,  ~ disease, type = "response")
mc5coll_2_d
pairs(mc5coll_2_d,reverse = TRUE)
plot(mc5coll_2_d)
confint(pairs(mc5coll_2_d,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlog_c5coll_3 <-
  glm(mean_C5_coll2 ~
        viz * disease * mean_indivhorz +
        viz * disease * mean_indivvertical +
        viz * disease * mean_collhorz +
        viz * disease * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5coll_3, type = "III")
summary(rlog_c5coll_3)

## Interpretation of model 3
emmip(rlog_c5coll_3, viz ~ mean_indivvertical , cov.reduce = range, type = "response")
emmip(rlog_c5coll_3, viz ~ mean_collhorz , cov.reduce = range, type = "response")



#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlog_c5coll_4 <- glm(mean_C5_coll2   ~ viz * disease + viz * edhi, data = datMain,family=binomial("logit"))
```

```
Anova(rlog_c5coll_4,type="III")
summary(rlog_c5coll_4)

# Interpretation of model 4
mc5coll_4_v <- emmeans(rlog_c5coll_4,~viz|edhi,type = "response")
mc5coll_4_v
confint(pairs(mc5coll_4_v,reverse = TRUE))
plot(mc5coll_4_v)

mc5coll_4_e <- emmeans(rlog_c5coll_4,  ~ edhi|viz, type = "response")
mc5coll_4_e
confint(pairs(mc5coll_4_e,reverse = TRUE))
plot(mc5coll_4_e)


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlog_c5coll_5 <-
  glm(mean_C5_coll2  ~ viz * disease +
        viz * edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMain,family=binomial("logit")
  )
Anova(rlog_c5coll_5, type = "III")
summary(rlog_c5coll_5)

# Interpretation of model 5
mc5coll_5_v <- emmeans(rlog_c5coll_5,~viz|edhi,type = "response")
mc5coll_5_v
confint(pairs(mc5coll_5_v,reverse = TRUE))
plot(mc5coll_5_v)

mc5coll_5_e <- emmeans(rlog_c5coll_5,  ~ edhi|viz, type = "response")
mc5coll_5_e
confint(pairs(mc5coll_5_e,reverse = TRUE))
plot(mc5coll_5_e)


# Now we run the models on imputed data.
#CHANGELOG ASJ: No missing data for mean_C5_coll2, no model with imputation performed, 21-09-20.



#### RESEARCH QUESTION 3: EXISTING INTERVENTIONS VS HERDIMM ####

# Recall: Research question 3 (secondary): For any of the 4 diseases, do existing interventions
# already available online outperform the herdimm intervention?

# 1-way ANOVA (adjusted for covariates and with moderators)

# Here, we test to see if there is an existing intervention to convey the concept of herd immunity
# that outperforms our intervention. Because the existing interventions each only represent one
# disease whereas ours was built to explicitly represent to the epidemiology of
```

```
different infectious
# diseases, we do this by comparing relevant study arms for one disease at a time. These models are
# therefore similar to our models for two-way analyses, but disease*viz collapses into a single
# variable (study arm).

#### +++ ONE WAY ANOVA FOR RISK PERCEPTION 2-6 ####
#### + Generic ####

#CHANGELOG ASJ: Mean risk perception replaced by MRP2_6 in the rest of the section, 2021-08-05.

# Model 1: Check for direct effects of factors without any covariates
options(contrasts=c("contr.sum", "contr.poly")) # CHANGELOG: Added to calculate properly the type III sum of squares with lm 2021-07-27
mG_RP1 <- lm(mean_riskperception2_6 ~ studyarm, data=datGeneric)
Anova(mG_RP1,type="III")
summary(mG_RP1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mG_RP2 <-
  lm(
    mean_riskperception2_6 ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +  #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_RP2,type="III")
summary(mG_RP2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mG_RP3 <-
  lm(
    mean_riskperception2_6 ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +  #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_RP3,type="III")
summary(mG_RP3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
```

```
# Model 4: Check for interaction studyarm*edhi
mG_RP4 <- lm(mean_riskperception2_6 ~ studyarm*edhi, data=datGeneric)
Anova(mG_RP4,type="III")
summary(mG_RP4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mG_RP5 <- lm(mean_riskperception2_6 ~ studyarm*edhi +
               mean_indivhorz +
               mean_indivvertical +
               mean_collhorz +
               mean_collvertical +
               bornincanada +
               language_1 +
               language_2 +
               Asian_group +
               White_group +
               disability_any +
               genderidentity +
               income +
               age, data=datGeneric)
Anova(mG_RP5,type="III")
summary(mG_RP5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mG_RP1)
checkassumptions(mG_RP2)
checkassumptions(mG_RP3)
checkassumptions(mG_RP4)
checkassumptions(mG_RP5)

#### ++ Interpretation ####

mG_RP1_arm <- emmeans(mG_RP1,~studyarm)
mG_RP1_arm
pairs(mG_RP1_arm)
plot(mG_RP1_arm)

mG_RP2_arm <- emmeans(mG_RP2,~studyarm)
mG_RP2_arm
pairs(mG_RP2_arm)
plot(mG_RP2_arm)

emmip(mG_RP3, studyarm ~ mean_indivhorz, cov.reduce = range)
emmip(mG_RP3, studyarm ~ mean_indivvertical, cov.reduce = range)
emmip(mG_RP3, studyarm ~ mean_collvertical, cov.reduce = range)

mG_RP4_v <- emmeans(mG_RP4,~studyarm|edhi)
mG_RP4_v
pairs(mG_RP4_v)
mG_RP4_e <- emmeans(mG_RP4,~edhi|studyarm)
mG_RP4_e
pairs(mG_RP4_e)

emmeans(mG_RP5,~studyarm|edhi)




# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_riskperception2_6, no model with imputation performed, 21-09-20.


#### + Measles ####
```

```r

# Model 1: Check for direct effects of factors without any covariates
mM_RP26_1 <- lm(mean_riskperception2_6 ~ studyarm, data=datMeasles)
Anova(mM_RP26_1,type="III")
summary(mM_RP26_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mM_RP26_2 <-
  lm(
    mean_riskperception2_6 ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +  #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_RP26_2,type="III")
summary(mM_RP26_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mM_RP26_3 <-
  lm(
    mean_riskperception2_6 ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +  #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_RP26_3,type="III")
summary(mM_RP26_3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mM_RP26_4 <- lm(mean_riskperception2_6 ~ studyarm*edhi, data=datMeasles)
Anova(mM_RP26_4,type="III")
summary(mM_RP26_4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mM_RP26_5 <- lm(mean_riskperception2_6 ~ studyarm*edhi +
                  mean_indivhorz +
                  mean_indivvertical +
                  mean_collhorz +
                  mean_collvertical +
                  bornincanada +
                  language_1 +
                  language_2 +
```

```
                  Asian_group +
                  White_group +
                  disability_any +
                  genderidentity +
                  income +
                  age, data=datMeasles)
Anova(mM_RP26_5,type="III")
summary(mM_RP26_5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mM_RP26_1)
checkassumptions(mM_RP26_2)
checkassumptions(mM_RP26_3)
checkassumptions(mM_RP26_4)
checkassumptions(mM_RP26_5)

#### ++ Interpretation ####
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests, 2021-08-11.
mM_RP26_1_arm <- emmeans(mM_RP26_1,~studyarm)
mM_RP26_1_arm
pairs(mM_RP26_1_arm)

emmeans(mM_RP26_2,~studyarm)

emmip(mM_RP26_3, studyarm ~ mean_indivvertical, cov.reduce = range)

emmeans(mM_RP26_4,~studyarm|edhi)

emmeans(mM_RP26_5,~studyarm|edhi)



# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_riskperception2_6, no model with imputation performed, 21-09-20.



#### + Flu ####

# Model 1: Check for direct effects of factors without any covariates
mF_RP26_1 <- lm(mean_riskperception2_6 ~ studyarm, data=datFlu)
Anova(mF_RP26_1,type="III")
summary(mF_RP26_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mF_RP26_2 <-
  lm(
    mean_riskperception2_6 ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +  #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
```

```
      age,
    data = datFlu
  )
Anova(mF_RP26_2,type="III")
summary(mF_RP26_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mF_RP26_3 <-
  lm(
    mean_riskperception2_6 ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datFlu
  )
Anova(mF_RP26_3,type="III")
summary(mF_RP26_3)


# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mF_RP26_4 <- lm(mean_riskperception2_6 ~ studyarm*edhi, data=datFlu)
Anova(mF_RP26_4,type="III")
summary(mF_RP26_4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mF_RP26_5 <- lm(mean_riskperception2_6 ~ studyarm*edhi +
                  mean_indivhorz +
                  mean_indivvertical +
                  mean_collhorz +
                  mean_collvertical +
                  bornincanada +
                  language_1 +
                  language_2 +
                  Asian_group +
                  White_group +
                  disability_any +
                  genderidentity +
                  income +
                  age, data=datFlu)
Anova(mF_RP26_5,type="III")
summary(mF_RP26_5)


#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mF_RP26_1)
checkassumptions(mF_RP26_2)
checkassumptions(mF_RP26_3)
checkassumptions(mF_RP26_4)
checkassumptions(mF_RP26_5)

#### ++ Interpretation ####
```

```r
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests, 2021-08-11.

mF_RP26_1_arm <- emmeans(mF_RP26_1,~studyarm)
mF_RP26_1_arm
pairs(mF_RP26_1_arm)

mF_RP26_2_arm <- emmeans(mF_RP26_2,~studyarm)
mF_RP26_2_arm
pairs(mF_RP26_2_arm)

emmip(mF_RP26_3, studyarm ~ mean_indivhorz, cov.reduce = range)
emmip(mF_RP26_3, studyarm ~ mean_indivvertical, cov.reduce = range)

emmeans(mF_RP26_4,~studyarm|edhi)

emmeans(mF_RP26_5,~studyarm|edhi)




# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_riskperception2_6, no model with imputation performed, 21-09-20.



#CHANGELOG ASJ: Added for secondary outcomes, 21-09-03.
#### +++ ONE WAY ANOVA FOR EMOTIONS ####
#### + Generic ####

# Model 1: Check for direct effects of factors without any covariates
options(contrasts=c("contr.sum", "contr.poly"))
mG_E1 <- lm(mean_emotion ~ studyarm, data=datGeneric)
Anova(mG_E1,type="III")
summary(mG_E1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mG_E2 <-
  lm(
    mean_emotion ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_E2,type="III")
summary(mG_E2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mG_E3 <-
  lm(
    mean_emotion ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
```

```
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_E3,type="III")
summary(mG_E3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mG_E4 <- lm(mean_emotion ~ studyarm*edhi, data=datGeneric)
Anova(mG_E4,type="III")
summary(mG_E4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mG_E5 <- lm(mean_emotion ~ studyarm*edhi +
              mean_indivhorz +
              mean_indivvertical +
              mean_collhorz +
              mean_collvertical +
              bornincanada +
              language_1 +
              language_2 +
              Asian_group +
              White_group +
              disability_any +
              genderidentity +
              income +
              age, data=datGeneric)
Anova(mG_E5,type="III")
summary(mG_E5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mG_E1)
checkassumptions(mG_E2)
checkassumptions(mG_E3)
checkassumptions(mG_E4)
checkassumptions(mG_E5)

#### ++ Interpretation ####

mG_E1_arm <- emmeans(mG_E1,~studyarm)
mG_E1_arm
plot(mG_E1_arm)

mG_E2_arm <- emmeans(mG_E2,~studyarm)
mG_E2_arm
plot(mG_E2_arm)

emmip(mG_E3, studyarm ~ mean_indivhorz, cov.reduce = range)
emmip(mG_E3, studyarm ~ mean_collhorz, cov.reduce = range)

emmeans(mG_E4,~studyarm|edhi)

emmeans(mG_E5,~studyarm|edhi)

```

```
# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_emotion, no model with imputation performed,
21-09-20.



#### + Measles ####

# Model 1: Check for direct effects of factors without any covariates
mM_E_1 <- lm(mean_emotion ~ studyarm, data=datMeasles)
Anova(mM_E_1,type="III")
summary(mM_E_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mM_E_2 <-
  lm(
    mean_emotion ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_E_2,type="III")
summary(mM_E_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment
for other covariates
mM_E_3 <-
  lm(
    mean_emotion ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_E_3,type="III")
summary(mM_E_3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mM_E_4 <- lm(mean_emotion ~ studyarm*edhi, data=datMeasles)
Anova(mM_E_4,type="III")
summary(mM_E_4)
```

```
# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mM_E_5 <- lm(mean_emotion ~ studyarm*edhi +
               mean_indivhorz +
               mean_indivvertical +
               mean_collhorz +
               mean_collvertical +
               bornincanada +
               language_1 +
               language_2 +
               Asian_group +
               White_group +
               disability_any +
               genderidentity +
               income +
               age, data=datMeasles)
Anova(mM_E_5,type="III")
summary(mM_E_5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mM_E_1)
checkassumptions(mM_E_2)
checkassumptions(mM_E_3)
checkassumptions(mM_E_4)
checkassumptions(mM_E_5)

#### ++ Interpretation ####
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests, 2021-08-11.
mM_E_1_arm <- emmeans(mM_E_1,~studyarm)
mM_E_1_arm
pairs(mM_E_1_arm)

mM_E_2_arm <- emmeans(mM_E_2,~studyarm)
mM_E_2_arm
pairs(mM_E_2_arm)

emmeans(mM_E_4,~studyarm|edhi)

emmeans(mM_E_5,~studyarm|edhi)



# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_emotion, no model with imputation performed, 21-09-20.


#### + Flu ####

# Model 1: Check for direct effects of factors without any covariates
mF_E_1 <- lm(mean_emotion ~ studyarm, data=datFlu)
Anova(mF_E_1,type="III")
summary(mF_E_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mF_E_2 <-
  lm(
    mean_emotion ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
```

```
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datFlu
  )
Anova(mF_E_2,type="III")
summary(mF_E_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mF_E_3 <-
  lm(
    mean_emotion ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datFlu
  )
Anova(mF_E_3,type="III")
summary(mF_E_3)


# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mF_E_4 <- lm(mean_emotion ~ studyarm*edhi, data=datFlu)
Anova(mF_E_4,type="III")
summary(mF_E_4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mF_E_5 <- lm(mean_emotion ~ studyarm*edhi +
               mean_indivhorz +
               mean_indivvertical +
               mean_collhorz +
               mean_collvertical +
               bornincanada +
               language_1 +
               language_2 +
               Asian_group +
               White_group +
               disability_any +
               genderidentity +
               income +
               age, data=datFlu)
Anova(mF_E_5,type="III")
summary(mF_E_5)


#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
```

```
# outcome.

checkassumptions(mF_E_1)
checkassumptions(mF_E_2)
checkassumptions(mF_E_3)
checkassumptions(mF_E_4)
checkassumptions(mF_E_5)

#### ++ Interpretation ####
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests,
2021-08-11.

emmeans(mF_E_1,~studyarm)

emmeans(mF_E_2,~studyarm)

mF_E4_s<-emmeans(mF_E_4,~studyarm|edhi)
mF_E4_s
pairs(mF_E4_s)

mF_E4_e<-emmeans(mF_E_4,~edhi|studyarm)
pairs(mF_E4_e)

emmeans(mF_E_5,~studyarm|edhi)




# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for mean_emotion, no model with imputation performed,
21-09-20.


#### +++ ONE WAY ANOVA FOR KNOWLEDGE ####
#### + Generic ####

# Model 1: Check for direct effects of factors without any covariates
options(contrasts=c("contr.sum", "contr.poly"))
mG_K1 <- lm(sum_knowledge ~ studyarm, data=datGeneric)
Anova(mG_K1,type="III")
summary(mG_K1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mG_K2 <-
  lm(
    sum_knowledge ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_K2,type="III")
summary(mG_K2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment
for other covariates
mG_K3 <-
```

```
  lm(
    sum_knowledge ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datGeneric
  )
Anova(mG_K3,type="III")
summary(mG_K3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mG_K4 <- lm(sum_knowledge ~ studyarm*edhi, data=datGeneric)
Anova(mG_K4,type="III")
summary(mG_K4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mG_K5 <- lm(sum_knowledge ~ studyarm*edhi +
              mean_indivhorz +
              mean_indivvertical +
              mean_collhorz +
              mean_collvertical +
              bornincanada +
              language_1 +
              language_2 +
              Asian_group +
              White_group +
              disability_any +
              genderidentity +
              income +
              age, data=datGeneric)
Anova(mG_K5,type="III")
summary(mG_K5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mG_K1)
checkassumptions(mG_K2)
checkassumptions(mG_K3)
checkassumptions(mG_K4)
checkassumptions(mG_K5)

#### ++ Interpretation ####

mG_K1_arm <- emmeans(mG_K1,~studyarm)
mG_K1_arm
pairs(mG_K1_arm)

mG_K2_arm <- emmeans(mG_K2,~studyarm)
mG_K2_arm
pairs(mG_K2_arm)

emmeans(mG_K4,~studyarm|edhi)

```

```r
emmeans(mG_K5,~studyarm|edhi)


# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for sum_knowledge, no model with imputation performed, 21-09-20.



#### + Measles ####

# Model 1: Check for direct effects of factors without any covariates
mM_K_1 <- lm(sum_knowledge ~ studyarm, data=datMeasles)
Anova(mM_K_1,type="III")
summary(mM_K_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mM_K_2 <-
  lm(
    sum_knowledge ~
      studyarm +
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_K_2,type="III")
summary(mM_K_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mM_K_3 <-
  lm(
    sum_knowledge ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datMeasles
  )
Anova(mM_K_3,type="III")
summary(mM_K_3)

# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mM_K_4 <- lm(sum_knowledge ~ studyarm*edhi, data=datMeasles)
Anova(mM_K_4,type="III")
summary(mM_K_4)
```

```

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mM_K_5 <- lm(sum_knowledge ~ studyarm*edhi +
               mean_indivhorz +
               mean_indivvertical +
               mean_collhorz +
               mean_collvertical +
               bornincanada +
               language_1 +
               language_2 +
               Asian_group +
               White_group +
               disability_any +
               genderidentity +
               income +
               age, data=datMeasles)
Anova(mM_K_5,type="III")
summary(mM_K_5)

#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mM_K_1)
checkassumptions(mM_K_2)
checkassumptions(mM_K_3)
checkassumptions(mM_K_4)
checkassumptions(mM_K_5)

#### ++ Interpretation ####
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests, 2021-08-11.
mM_K_1_arm <- emmeans(mM_K_1,~studyarm)
mM_K_1_arm
pairs(mM_K_1_arm)

mM_K_2_arm <- emmeans(mM_K_2,~studyarm)
mM_K_2_arm
pairs(mM_K_2_arm)

emmip(mM_K_3, studyarm ~ mean_indivvertical, cov.reduce = range)

emmeans(mM_K_4,~studyarm|edhi)

emmeans(mM_K_5,~studyarm|edhi)



# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for sum_knowledge, no model with imputation performed, 21-09-20.




#### + Flu ####

# Model 1: Check for direct effects of factors without any covariates
mF_K_1 <- lm(sum_knowledge ~ studyarm, data=datFlu)
Anova(mF_K_1,type="III")
summary(mF_K_1)

# Model 2: Check for direct effects of factors with adjustment for other covariates
mF_K_2 <-
  lm(
    sum_knowledge ~
      studyarm +
```

```
      mean_indivhorz +
      mean_indivvertical +
      mean_collhorz +
      mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datFlu
  )
Anova(mF_K_2,type="III")
summary(mF_K_2)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
mF_K_3 <-
  lm(
    sum_knowledge ~
      studyarm * mean_indivhorz +
      studyarm * mean_indivvertical +
      studyarm * mean_collhorz +
      studyarm * mean_collvertical +
      bornincanada +
      language_1 +
      language_2 +
      Asian_group +
      White_group +
      disability_any +
      genderidentity +
      income +
      edhi +
      age,
    data = datFlu
  )
Anova(mF_K_3,type="III")
summary(mF_K_3)


# CHANGELOG: Model 4 + 5 : To check the interaction between the study arm and education
# Model 4: Check for interaction studyarm*edhi
mF_K_4 <- lm(sum_knowledge ~ studyarm*edhi, data=datFlu)
Anova(mF_K_4,type="III")
summary(mF_K_4)

# Model 5: Check for interaction studyarm*edhi with adjustment for other covariates
mF_K_5 <- lm(sum_knowledge ~ studyarm*edhi +
               mean_indivhorz +
               mean_indivvertical +
               mean_collhorz +
               mean_collvertical +
               bornincanada +
               language_1 +
               language_2 +
               Asian_group +
               White_group +
               disability_any +
               genderidentity +
               income +
               age, data=datFlu)
Anova(mF_K_5,type="III")
summary(mF_K_5)


```

```
#### ++ Validating assumptions ####

# Check to see if the assumptions of the models are respected. If not, we will transform the
# outcome.

checkassumptions(mF_K_1)
checkassumptions(mF_K_2)
checkassumptions(mF_K_3)
checkassumptions(mF_K_4)
checkassumptions(mF_K_5)

#### ++ Interpretation ####
#CHANGELOG ASJ: adapt relevant functions depending on results of global tests, 2021-08-11.

mF_K_1_arm <- emmeans(mF_K_1,~studyarm)
mF_K_1_arm
pairs(mF_K_1_arm)

mF_K_2_arm <- emmeans(mF_K_2,~studyarm)
mF_K_2_arm
pairs(mF_K_2_arm)

emmip(mF_K_3, studyarm ~ mean_indivvertical, cov.reduce = range)

emmeans(mF_K_4,~studyarm|edhi)

mF_K_5_arm <- emmeans(mF_K_5,~studyarm|edhi)
mF_K_5_arm
pairs(mF_K_5_arm)
pairs(emmeans(mF_K_5,~edhi|studyarm))



# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for sum_knowledge, no model with imputation performed, 21-09-20.







#CHANGELOG ASJ: Add logistic regression models for dichotomous outcomes, 2021-08-11.

#### +++ Logistic regression for RP1 ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$RP1_dicho <- relevel(datGeneric$RP1_dicho,ref = "Low")
datMeasles$RP1_dicho <- relevel(datMeasles$RP1_dicho,ref = "Low")
datFlu$RP1_dicho <- relevel(datFlu$RP1_dicho,ref = "Low")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_RP1_1 <- glm(RP1_dicho  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_RP1_1,type="III")
summary(rlogG_RP1_1)

mgRP1 <- emmeans(rlogG_RP1_1,~studyarm,type = "response") #proportions
mgRP1
confint(pairs(mgRP1,reverse = TRUE)) #Odds ratios
pairs(mgRP1,reverse = TRUE, type = "response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_RP1_2 <-
```

```
  glm(RP1_dicho  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_RP1_2, type = "III")
summary(rlogG_RP1_2)

mgRP1_2 <- emmeans(rlogG_RP1_2,~studyarm,type = "response")
mgRP1_2
confint(pairs(mgRP1_2,reverse = TRUE))
pairs(mgRP1_2,reverse = TRUE,type = "response")


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_RP1_3 <-
  glm(RP1_dicho  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_RP1_3, type = "III")
summary(rlogG_RP1_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_RP1_4 <- glm(RP1_dicho  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_RP1_4,type="III")
summary(rlogG_RP1_4)

mgRP1_4_s <- emmeans(rlogG_RP1_4,~studyarm|edhi,type = "response")
mgRP1_4_s
confint(pairs(mgRP1_4_s,reverse = TRUE))

mgRP1_4_e <- emmeans(rlogG_RP1_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgRP1_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_RP1_5 <-
```

```
  glm(RP1_dicho   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_RP1_5, type = "III")
summary(rlogG_RP1_5)

mgRP1_5_s <- emmeans(rlogG_RP1_5,~studyarm|edhi,type = "response")
mgRP1_5_s
confint(pairs(mgRP1_5_s,reverse = TRUE))

mgRP1_5_e <- emmeans(rlogG_RP1_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgRP1_5_e,reverse = TRUE))

# Now running models on imputed data.

imprlogG_RP1_1 <-
  with(data = datGenericImputed,
       exp = glm(relevel(RP1_dicho,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogG_RP1_1 <- pool(imprlogG_RP1_1)
summary(combineimprlogG_RP1_1)


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_RP1_1 <- glm(RP1_dicho  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_RP1_1,type="III")
summary(rlogM_RP1_1)

mmRP1 <- emmeans(rlogM_RP1_1,~studyarm,type = "response") #proportions
mmRP1
confint(pairs(mmRP1,reverse = TRUE)) #Odds ratios
pairs(mmRP1,reverse = TRUE, type = "response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_RP1_2 <-
  glm(RP1_dicho  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_RP1_2, type = "III")
```

```
summary(rlogM_RP1_2)

mmRP1_2 <- emmeans(rlogM_RP1_2,~studyarm,type = "response")
mmRP1_2
confint(pairs(mmRP1_2,reverse = TRUE))
pairs(mmRP1_2,reverse = TRUE,type = "response")


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_RP1_3 <-
  glm(RP1_dicho  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_RP1_3, type = "III")
summary(rlogM_RP1_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_RP1_4 <- glm(RP1_dicho  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_RP1_4,type="III")
summary(rlogM_RP1_4)

mmRP1_4_s <- emmeans(rlogM_RP1_4,~studyarm|edhi,type = "response")
mmRP1_4_s
confint(pairs(mmRP1_4_s,reverse = TRUE))

mmRP1_4_e <- emmeans(rlogM_RP1_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmRP1_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_RP1_5 <-
  glm(RP1_dicho   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_RP1_5, type = "III")
summary(rlogM_RP1_5)
```

```

mmRP1_5_s <- emmeans(rlogM_RP1_5,~studyarm|edhi,type = "response")
mmRP1_5_s
confint(pairs(mmRP1_5_s,reverse = TRUE))

mmRP1_5_e <- emmeans(rlogM_RP1_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmRP1_5_e,reverse = TRUE))

# Now running models on imputed data.

imprlogM_RP1_1 <-
  with(data = datMeaslesImputed,
       exp = glm(relevel(RP1_dicho,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogM_RP1_1 <- pool(imprlogM_RP1_1)
summary(combineimprlogM_RP1_1)


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_RP1_1 <- glm(RP1_dicho  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_RP1_1,type="III")
summary(rlogF_RP1_1)

mfRP1_1 <- emmeans(rlogF_RP1_1,~studyarm,type = "response") #proportions
mfRP1_1
confint(pairs(mfRP1_1,reverse = TRUE)) #Odds ratios
pairs(mfRP1_1,reverse = TRUE, type = "response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_RP1_2 <-
  glm(RP1_dicho  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_RP1_2, type = "III")
summary(rlogF_RP1_2)

mfRP1_2 <- emmeans(rlogF_RP1_2,~studyarm,type = "response")
mfRP1_2
confint(pairs(mfRP1_2,reverse = TRUE))
pairs(mfRP1_2,reverse = TRUE,type = "response")


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_RP1_3 <-
  glm(RP1_dicho  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
```

```r
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_RP1_3, type = "III")
summary(rlogF_RP1_3)

emmip(rlogF_RP1_3, studyarm ~ mean_collhorz, cov.reduce = range, type = "response")
emmip(rlogF_RP1_3, studyarm ~ mean_collvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_RP1_4 <- glm(RP1_dicho  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogF_RP1_4,type="III")
summary(rlogF_RP1_4)

mfRP1_4_s <- emmeans(rlogF_RP1_4,~studyarm|edhi,type = "response")
mfRP1_4_s
confint(pairs(mfRP1_4_s,reverse = TRUE))

mfRP1_4_e <- emmeans(rlogF_RP1_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfRP1_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_RP1_5 <-
  glm(RP1_dicho   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_RP1_5, type = "III")
summary(rlogF_RP1_5)

mfRP1_5_s <- emmeans(rlogF_RP1_5,~studyarm|edhi,type = "response")
mfRP1_5_s
confint(pairs(mfRP1_5_s,reverse = TRUE))

mfRP1_5_e <- emmeans(rlogF_RP1_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfRP1_5_e,reverse = TRUE))

# Now running models on imputed data.

imprlogF_RP1_1 <-
  with(data = datFluImputed,
       exp = glm(relevel(RP1_dicho,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogF_RP1_1 <- pool(imprlogF_RP1_1)
summary(combineimprlogF_RP1_1)

```

```r

#### +++ Logistic regression for VaxIntention ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$vaxintent <- relevel(datGeneric$vaxintent,ref = "Low")
datMeasles$vaxintent <- relevel(datMeasles$vaxintent,ref = "Low")
datFlu$vaxintent <- relevel(datFlu$vaxintent,ref = "Low")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_vi_1 <- glm(vaxintent ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_vi_1,type="III")
summary(rlogG_vi_1)

mgvi_1 <- emmeans(rlogG_vi_1,~studyarm,type = "response") #proportions
mgvi_1
confint(pairs(mgvi_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_vi_2 <-
  glm(vaxintent  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_vi_2, type = "III")
summary(rlogG_vi_2)

mgvi_2 <- emmeans(rlogG_vi_2,~studyarm,type = "response")
mgvi_2
confint(pairs(mgvi_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_vi_3 <-
  glm(vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_vi_3, type = "III")
```

```r
summary(rlogG_vi_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_vi_4 <- glm(vaxintent  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_vi_4,type="III")
summary(rlogG_vi_4)

mgvi_4_s <- emmeans(rlogG_vi_4,~studyarm|edhi,type = "response")
mgvi_4_s
confint(pairs(mgvi_4_s,reverse = TRUE))

mgvi_4_e <- emmeans(rlogG_vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgvi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_vi_5 <-
  glm(vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_vi_5, type = "III")
summary(rlogG_vi_5)

mgvi_5_s <- emmeans(rlogG_vi_5,~studyarm|edhi,type = "response")
mgvi_5_s
confint(pairs(mgvi_5_s,reverse = TRUE))

mgvi_5_e <- emmeans(rlogG_vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgvi_5_e,reverse = TRUE))

# Now running models on imputed data.

imprlogG_vi_1 <-
  with(data = datGenericImputed,
       exp = glm(relevel(vaxintent,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogG_vi_1 <- pool(imprlogG_vi_1)
summary(combineimprlogG_vi_1)

#CHANGELOG ASJ: add test to check if global test result changes or not, 21-09-20.
imprlogG_vi_vide <-
  with(data = datGenericImputed,
       exp = glm(relevel(vaxintent,ref="Low") ~ 1,family = binomial("logit"))) #model with intercept only
combineimprlogG_vi_vide <- pool(imprlogG_vi_vide)
summary(combineimprlogG_vi_vide)
test_sig<-D2(imprlogG_vi_1,imprlogG_vi_vide) #compare model with studyarm vs model with intercept only
summary(test_sig)
```

```
#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_vi_1 <- glm(vaxintent  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_vi_1,type="III")
summary(rlogM_vi_1)

mmvi <- emmeans(rlogM_vi_1,~studyarm,type = "response") #proportions
mmvi
confint(pairs(mmvi,reverse = TRUE)) #Odds ratios
pairs(mmvi,reverse = TRUE, type = "response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_vi_2 <-
  glm(vaxintent  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_vi_2, type = "III")
summary(rlogM_vi_2)

mmvi_2 <- emmeans(rlogM_vi_2,~studyarm,type = "response")
mmvi_2
confint(pairs(mmvi_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_vi_3 <-
  glm(vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_vi_3, type = "III")
summary(rlogM_vi_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####
```

```
rlogM_vi_4 <- glm(vaxintent  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_vi_4,type="III")
summary(rlogM_vi_4)

mmvi_4_s <- emmeans(rlogM_vi_4,~studyarm|edhi,type = "response")
mmvi_4_s
confint(pairs(mmvi_4_s,reverse = TRUE))

mmvi_4_e <- emmeans(rlogM_vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmvi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_vi_5 <-
  glm(vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_vi_5, type = "III")
summary(rlogM_vi_5)

mmvi_5_s <- emmeans(rlogM_vi_5,~studyarm|edhi,type = "response")
mmvi_5_s
confint(pairs(mmvi_5_s,reverse = TRUE))

mmvi_5_e <- emmeans(rlogM_vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmvi_5_e,reverse = TRUE))

# Now running models on imputed data.
imprlogM_vi_1 <-
  with(data = datMeaslesImputed,
       exp = glm(relevel(vaxintent,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogM_vi_1 <- pool(imprlogM_vi_1)
summary(combineimprlogM_vi_1)


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_vi_1 <- glm(vaxintent  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_vi_1,type="III")
summary(rlogF_vi_1)

mfvi_1 <- emmeans(rlogF_vi_1,~studyarm,type = "response") #proportions
mfvi_1
confint(pairs(mfvi_1,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_vi_2 <-
  glm(vaxintent  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
```

```
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_vi_2, type = "III")
summary(rlogF_vi_2)

mfvi_2 <- emmeans(rlogF_vi_2,~studyarm,type = "response")
mfvi_2
confint(pairs(mfvi_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_vi_3 <-
  glm(vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_vi_3, type = "III")
summary(rlogF_vi_3)

emmip(rlogF_vi_3, studyarm ~ mean_indivvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_vi_4 <- glm(vaxintent  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_vi_4,type="III")
summary(rlogF_vi_4)

mfvi_4_s <- emmeans(rlogF_vi_4,~studyarm|edhi,type = "response")
mfvi_4_s
confint(pairs(mfvi_4_s,reverse = TRUE))

mfvi_4_e <- emmeans(rlogF_vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfvi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_vi_5 <-
  glm(vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
```

```
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_vi_5, type = "III")
summary(rlogF_vi_5)

mfvi_5_s <- emmeans(rlogF_vi_5,~studyarm|edhi,type = "response")
mfvi_5_s
confint(pairs(mfvi_5_s,reverse = TRUE))

mfvi_5_e <- emmeans(rlogF_vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfvi_5_e,reverse = TRUE))

# Now running models on imputed data.
imprlogF_vi_1 <-
  with(data = datFluImputed,
       exp = glm(relevel(vaxintent,ref="Low") ~ studyarm,family = binomial("logit")))
combineimprlogF_vi_1 <- pool(imprlogF_vi_1)
summary(combineimprlogF_vi_1)

#### +++ Logistic regression for C19VaxIntention ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$c19vaxintent <- relevel(datGeneric$c19vaxintent,ref = "Low")
datMeasles$c19vaxintent <- relevel(datMeasles$c19vaxintent,ref = "Low")
datFlu$c19vaxintent <- relevel(datFlu$c19vaxintent,ref = "Low")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_c19vi_1 <- glm(c19vaxintent  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_c19vi_1,type="III")
summary(rlogG_c19vi_1)

mgc19vi_1 <- emmeans(rlogG_c19vi_1,~studyarm,type = "response") #proportions
mgc19vi_1
confint(pairs(mgc19vi_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_c19vi_2 <-
  glm(c19vaxintent  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
```

```r
Anova(rlogG_c19vi_2, type = "III")
summary(rlogG_c19vi_2)

mgc19vi_2 <- emmeans(rlogG_c19vi_2,~studyarm,type = "response")
mgc19vi_2
confint(pairs(mgc19vi_2,reverse = TRUE))
pairs(mgc19vi_2,reverse = TRUE, type = "response")

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_c19vi_3 <-
  glm(c19vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_c19vi_3, type = "III")
summary(rlogG_c19vi_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_c19vi_4 <- glm(c19vaxintent  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_c19vi_4,type="III")
summary(rlogG_c19vi_4)

mgc19vi_4_s <- emmeans(rlogG_c19vi_4,~studyarm|edhi,type = "response")
mgc19vi_4_s
confint(pairs(mgc19vi_4_s,reverse = TRUE))

mgc19vi_4_e <- emmeans(rlogG_c19vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgc19vi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_c19vi_5 <-
  glm(c19vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_c19vi_5, type = "III")
summary(rlogG_c19vi_5)
```

```
mgc19vi_5_s <- emmeans(rlogG_c19vi_5,~studyarm|edhi,type = "response")
mgc19vi_5_s
confint(pairs(mgc19vi_5_s,reverse = TRUE))

mgc19vi_5_e <- emmeans(rlogG_c19vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgc19vi_5_e,reverse = TRUE))

#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_c19vi_1 <- glm(c19vaxintent  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_c19vi_1,type="III")
summary(rlogM_c19vi_1)

mmc19vi_1 <- emmeans(rlogM_c19vi_1,~studyarm,type = "response") #proportions
mmc19vi_1
confint(pairs(mmc19vi_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_c19vi_2 <-
  glm(c19vaxintent  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_c19vi_2, type = "III")
summary(rlogM_c19vi_2)

mmc19vi_2 <- emmeans(rlogM_c19vi_2,~studyarm,type = "response")
mmc19vi_2
confint(pairs(mmc19vi_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_c19vi_3 <-
  glm(c19vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
```

```r
Anova(rlogM_c19vi_3, type = "III")
summary(rlogM_c19vi_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_c19vi_4 <- glm(c19vaxintent  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_c19vi_4,type="III")
summary(rlogM_c19vi_4)

mmc19vi_4_s <- emmeans(rlogM_c19vi_4,~studyarm|edhi,type = "response")
mmc19vi_4_s
confint(pairs(mmc19vi_4_s,reverse = TRUE))

mmc19vi_4_e <- emmeans(rlogM_c19vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmc19vi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_c19vi_5 <-
  glm(c19vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_c19vi_5, type = "III")
summary(rlogM_c19vi_5)

mmc19vi_5_s <- emmeans(rlogM_c19vi_5,~studyarm|edhi,type = "response")
mmc19vi_5_s
confint(pairs(mmc19vi_5_s,reverse = TRUE))

mmc19vi_5_e <- emmeans(rlogM_c19vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmc19vi_5_e,reverse = TRUE))



#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_c19vi_1 <- glm(c19vaxintent  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_c19vi_1,type="III")
summary(rlogF_c19vi_1)

mfc19vi_1 <- emmeans(rlogF_c19vi_1,~studyarm,type = "response") #proportions
mfc19vi_1
confint(pairs(mfc19vi_1,reverse = TRUE)) #Odds ratios


# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_c19vi_2 <-
  glm(c19vaxintent  ~ studyarm +
        mean_indivhorz +
```

```
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_c19vi_2, type = "III")
summary(rlogF_c19vi_2)

mfc19vi_2 <- emmeans(rlogF_c19vi_2,~studyarm,type = "response")
mfc19vi_2
confint(pairs(mfc19vi_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_c19vi_3 <-
  glm(c19vaxintent  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_c19vi_3, type = "III")
summary(rlogF_c19vi_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_c19vi_4 <- glm(c19vaxintent  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_c19vi_4,type="III")
summary(rlogF_c19vi_4)

mfc19vi_4_s <- emmeans(rlogF_c19vi_4,~studyarm|edhi,type = "response")
mfc19vi_4_s
confint(pairs(mfc19vi_4_s,reverse = TRUE))

mfc19vi_4_e <- emmeans(rlogF_c19vi_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfc19vi_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_c19vi_5 <-
  glm(c19vaxintent   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
```

```
          mean_collhorz +
          mean_collvertical +
          bornincanada +
          language_1 +
          language_2 +
          Asian_group +
          White_group +
          disability_any +
          genderidentity +
          income +
          age,
        data = datFlu,family=binomial("logit")
    )
Anova(rlogF_c19vi_5, type = "III")
summary(rlogF_c19vi_5)

mfc19vi_5_s <- emmeans(rlogF_c19vi_5,~studyarm|edhi,type = "response")
mfc19vi_5_s
confint(pairs(mfc19vi_5_s,reverse = TRUE))

mfc19vi_5_e <- emmeans(rlogF_c19vi_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfc19vi_5_e,reverse = TRUE))







#### +++ Logistic regression for Trust Info ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$Htrustinfo <- relevel(datGeneric$Htrustinfo,ref = "No")
datMeasles$Htrustinfo <- relevel(datMeasles$Htrustinfo,ref = "No")
datFlu$Htrustinfo <- relevel(datFlu$Htrustinfo,ref = "No")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_ti_1 <- glm(Htrustinfo  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_ti_1,type="III")
summary(rlogG_ti_1)

mgti_1 <- emmeans(rlogG_ti_1,~studyarm,type = "response") #proportions
mgti_1
confint(pairs(mgti_1,reverse = TRUE)) #Odds ratios
pairs(mgti_1,reverse = TRUE,type="response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_ti_2 <-
    glm(Htrustinfo  ~ studyarm +
          mean_indivhorz +
          mean_indivvertical +
          mean_collhorz +
          mean_collvertical +
          bornincanada +
          language_1 +
          language_2 +
          Asian_group +
          White_group +
          disability_any +
          genderidentity +
          income +
          edhi +
          age,
        data = datGeneric,family=binomial("logit")
    )
```

```
Anova(rlogG_ti_2, type = "III")
summary(rlogG_ti_2)

mgti_2 <- emmeans(rlogG_ti_2,~studyarm,type = "response")
mgti_2
confint(pairs(mgti_2,reverse = TRUE))
pairs(mgti_2,reverse = TRUE,type="response")

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_ti_3 <-
  glm(Htrustinfo  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_ti_3, type = "III")
summary(rlogG_ti_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_ti_4 <- glm(Htrustinfo  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_ti_4,type="III")
summary(rlogG_ti_4)

mgti_4_s <- emmeans(rlogG_ti_4,~studyarm|edhi,type = "response")
mgti_4_s
confint(pairs(mgti_4_s,reverse = TRUE))

mgti_4_e <- emmeans(rlogG_ti_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgti_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_ti_5 <-
  glm(Htrustinfo   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_ti_5, type = "III")
summary(rlogG_ti_5)
```

```r

mgti_5_s <- emmeans(rlogG_ti_5,~studyarm|edhi,type = "response")
mgti_5_s
confint(pairs(mgti_5_s,reverse = TRUE))

mgti_5_e <- emmeans(rlogG_ti_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgti_5_e,reverse = TRUE))

# Now running models on imputed data.
imprlogG_ti_1 <-
  with(data = datGenericImputed,
       exp = glm(relevel(Htrustinfo,ref="No") ~ studyarm,family = binomial("logit")))
combineimprlogG_ti_1 <- pool(imprlogG_ti_1)
summary(combineimprlogG_ti_1)


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_ti_1 <- glm(Htrustinfo  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_ti_1,type="III")
summary(rlogM_ti_1)

mmti_1 <- emmeans(rlogM_ti_1,~studyarm,type = "response") #proportions
mmti_1
confint(pairs(mmti_1,reverse = TRUE)) #Odds ratios
pairs(mmti_1,reverse = TRUE, type = "response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_ti_2 <-
  glm(Htrustinfo  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_ti_2, type = "III")
summary(rlogM_ti_2)

mmti_2 <- emmeans(rlogM_ti_2,~studyarm,type = "response")
mmti_2
confint(pairs(mmti_2,reverse = TRUE))
pairs(mmti_2,reverse = TRUE,type = "response")

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_ti_3 <-
  glm(Htrustinfo  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
```

```r
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_ti_3, type = "III")
summary(rlogM_ti_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_ti_4 <- glm(Htrustinfo  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_ti_4,type="III")
summary(rlogM_ti_4)

mmti_4_s <- emmeans(rlogM_ti_4,~studyarm|edhi,type = "response")
mmti_4_s
confint(pairs(mmti_4_s,reverse = TRUE))

mmti_4_e <- emmeans(rlogM_ti_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmti_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_ti_5 <-
  glm(Htrustinfo   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_ti_5, type = "III")
summary(rlogM_ti_5)

mmti_5_s <- emmeans(rlogM_ti_5,~studyarm|edhi,type = "response")
mmti_5_s
confint(pairs(mmti_5_s,reverse = TRUE))

mmti_5_e <- emmeans(rlogM_ti_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmti_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for htrustinfo in measles, no model with imputation performed, 21-09-20.



#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_ti_1 <- glm(Htrustinfo  ~ studyarm, data = datFlu,family=binomial("logit"))
```

```
Anova(rlogF_ti_1,type="III")
summary(rlogF_ti_1)

mfti_1 <- emmeans(rlogF_ti_1,~studyarm,type = "response") #proportions
mfti_1
confint(pairs(mfti_1,reverse = TRUE)) #Odds ratios
pairs(mfti_1,reverse = TRUE,type="response")

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_ti_2 <-
  glm(Htrustinfo  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_ti_2, type = "III")
summary(rlogF_ti_2)

mfti_2 <- emmeans(rlogF_ti_2,~studyarm,type = "response")
mfti_2
confint(pairs(mfti_2,reverse = TRUE))
pairs(mfti_2,reverse = TRUE,type="response")


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_ti_3 <-
  glm(Htrustinfo  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_ti_3, type = "III")
summary(rlogF_ti_3)

#No interaction with moderators to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_ti_4 <- glm(Htrustinfo  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_ti_4,type="III")
summary(rlogF_ti_4)

mfti_4_s <- emmeans(rlogF_ti_4,~studyarm|edhi,type = "response")
mfti_4_s
```

```
confint(pairs(mfti_4_s,reverse = TRUE))

mfti_4_e <- emmeans(rlogF_ti_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfti_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_ti_5 <-
  glm(Htrustinfo   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_ti_5, type = "III")
summary(rlogF_ti_5)

mfti_5_s <- emmeans(rlogF_ti_5,~studyarm|edhi,type = "response")
mfti_5_s
confint(pairs(mfti_5_s,reverse = TRUE))

mfti_5_e <- emmeans(rlogF_ti_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfti_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for htrustinfo in flu, no model with imputation performed, 21-09-20.



#### +++ Logistic regression for C5 Confidence ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$mean_C5_conf2 <- relevel(as.factor(datGeneric$mean_C5_conf2),ref = "0")
datMeasles$mean_C5_conf2 <- relevel(as.factor(datMeasles$mean_C5_conf2),ref = "0")
datFlu$mean_C5_conf2 <- relevel(as.factor(datFlu$mean_C5_conf2),ref = "0")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_C5conf_1 <- glm(mean_C5_conf2  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5conf_1,type="III")
summary(rlogG_C5conf_1)

mgC5conf <- emmeans(rlogG_C5conf_1,~studyarm,type = "response") #proportions
mgC5conf
confint(pairs(mgC5conf,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_C5conf_2 <-
  glm(mean_C5_conf2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
```

```
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5conf_2, type = "III")
summary(rlogG_C5conf_2)

mgC5conf_2 <- emmeans(rlogG_C5conf_2,~studyarm,type = "response")
mgC5conf_2
confint(pairs(mgC5conf_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_C5conf_3 <-
  glm(mean_C5_conf2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5conf_3, type = "III")
summary(rlogG_C5conf_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_C5conf_4 <- glm(mean_C5_conf2  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5conf_4,type="III")
summary(rlogG_C5conf_4)

mgC5conf_4_s <- emmeans(rlogG_C5conf_4,~studyarm|edhi,type = "response")
mgC5conf_4_s
confint(pairs(mgC5conf_4_s,reverse = TRUE))

mgC5conf_4_e <- emmeans(rlogG_C5conf_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5conf_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_C5conf_5 <-
  glm(mean_C5_conf2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
```

```
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5conf_5, type = "III")
summary(rlogG_C5conf_5)

mgC5conf_5_s <- emmeans(rlogG_C5conf_5,~studyarm|edhi,type = "response")
mgC5conf_5_s
confint(pairs(mgC5conf_5_s,reverse = TRUE))

mgC5conf_5_e <- emmeans(rlogG_C5conf_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5conf_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.



#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_C5conf_1 <- glm(mean_C5_conf2  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5conf_1,type="III")
summary(rlogM_C5conf_1)

mmC5conf <- emmeans(rlogM_C5conf_1,~studyarm,type = "response") #proportions
mmC5conf
confint(pairs(mmC5conf,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_C5conf_2 <-
  glm(mean_C5_conf2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5conf_2, type = "III")
summary(rlogM_C5conf_2)

mmC5conf_2 <- emmeans(rlogM_C5conf_2,~studyarm,type = "response")
mmC5conf_2
confint(pairs(mmC5conf_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_C5conf_3 <-
```

```
  glm(mean_C5_conf2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5conf_3, type = "III")
summary(rlogM_C5conf_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_C5conf_4 <- glm(mean_C5_conf2  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5conf_4,type="III")
summary(rlogM_C5conf_4)

mmC5conf_4_s <- emmeans(rlogM_C5conf_4,~studyarm|edhi,type = "response")
mmC5conf_4_s
confint(pairs(mmC5conf_4_s,reverse = TRUE))

mmC5conf_4_e <- emmeans(rlogM_C5conf_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5conf_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_C5conf_5 <-
  glm(mean_C5_conf2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5conf_5, type = "III")
summary(rlogM_C5conf_5)

mmC5conf_5_s <- emmeans(rlogM_C5conf_5,~studyarm|edhi,type = "response")
mmC5conf_5_s
confint(pairs(mmC5conf_5_s,reverse = TRUE))

mmC5conf_5_e <- emmeans(rlogM_C5conf_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5conf_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.
```

```
#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_C5conf_1 <- glm(mean_C5_conf2  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5conf_1,type="III")
summary(rlogF_C5conf_1)

mfC5conf_1 <- emmeans(rlogF_C5conf_1,~studyarm,type = "response") #proportions
mfC5conf_1
confint(pairs(mfC5conf_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_C5conf_2 <-
  glm(mean_C5_conf2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5conf_2, type = "III")
summary(rlogF_C5conf_2)

mfC5conf_2 <- emmeans(rlogF_C5conf_2,~studyarm,type = "response")
mfC5conf_2
confint(pairs(mfC5conf_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_C5conf_3 <-
  glm(mean_C5_conf2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5conf_3, type = "III")
summary(rlogF_C5conf_3)

emmip(rlogF_C5conf_3, studyarm ~ mean_indivvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####
```

```r

rlogF_C5conf_4 <- glm(mean_C5_conf2  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5conf_4,type="III")
summary(rlogF_C5conf_4)

mfC5conf_4_s <- emmeans(rlogF_C5conf_4,~studyarm|edhi,type = "response")
mfC5conf_4_s
confint(pairs(mfC5conf_4_s,reverse = TRUE))

mfC5conf_4_e <- emmeans(rlogF_C5conf_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5conf_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_C5conf_5 <-
  glm(mean_C5_conf2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5conf_5, type = "III")
summary(rlogF_C5conf_5)

mfC5conf_5_s <- emmeans(rlogF_C5conf_5,~studyarm|edhi,type = "response")
mfC5conf_5_s
confint(pairs(mfC5conf_5_s,reverse = TRUE))

mfC5conf_5_e <- emmeans(rlogF_C5conf_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5conf_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.

#### +++ Logistic regression for C5 Complacency ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$mean_C5_comp2 <- relevel(as.factor(datGeneric$mean_C5_comp2),ref = "0")
datMeasles$mean_C5_comp2 <- relevel(as.factor(datMeasles$mean_C5_comp2),ref = "0")
datFlu$mean_C5_comp2 <- relevel(as.factor(datFlu$mean_C5_comp2),ref = "0")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_C5comp_1 <- glm(mean_C5_comp2  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5comp_1,type="III")
summary(rlogG_C5comp_1)

mgC5comp_1 <- emmeans(rlogG_C5comp_1,~studyarm,type = "response") #proportions
mgC5comp_1
confint(pairs(mgC5comp_1,reverse = TRUE)) #Odds ratios
pairs(mgC5comp_1,reverse = TRUE)

# Model 2: Check for direct effects of factors with adjustment for other covariates
```

```r
rlogG_C5comp_2 <-
  glm(mean_C5_comp2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5comp_2, type = "III")
summary(rlogG_C5comp_2)

mgC5comp_2 <- emmeans(rlogG_C5comp_2,~studyarm,type = "response")
mgC5comp_2
confint(pairs(mgC5comp_2,reverse = TRUE))
pairs(mgC5comp_2,reverse = TRUE)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_C5comp_3 <-
  glm(mean_C5_comp2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5comp_3, type = "III")
summary(rlogG_C5comp_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_C5comp_4 <- glm(mean_C5_comp2  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5comp_4,type="III")
summary(rlogG_C5comp_4)

mgC5comp_4_s <- emmeans(rlogG_C5comp_4,~studyarm|edhi,type = "response")
mgC5comp_4_s
confint(pairs(mgC5comp_4_s,reverse = TRUE))
pairs(mgC5comp_4_s,reverse = TRUE)

mgC5comp_4_e <- emmeans(rlogG_C5comp_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5comp_4_e,reverse = TRUE))
pairs(mgC5comp_4_e,reverse = TRUE)

# Model 5: Check for interaction viz*edhi with adjustment for other covariates

```

```r
rlogG_C5comp_5 <-
  glm(mean_C5_comp2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5comp_5, type = "III")
summary(rlogG_C5comp_5)

mgC5comp_5_s <- emmeans(rlogG_C5comp_5,~studyarm|edhi,type = "response")
mgC5comp_5_s
confint(pairs(mgC5comp_5_s,reverse = TRUE))

mgC5comp_5_e <- emmeans(rlogG_C5comp_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5comp_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_C5comp_1 <- glm(mean_C5_comp2  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5comp_1,type="III")
summary(rlogM_C5comp_1)

mmC5comp_1 <- emmeans(rlogM_C5comp_1,~studyarm,type = "response") #proportions
mmC5comp_1
confint(pairs(mmC5comp_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_C5comp_2 <-
  glm(mean_C5_comp2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5comp_2, type = "III")
summary(rlogM_C5comp_2)

mmC5comp_2 <- emmeans(rlogM_C5comp_2,~studyarm,type = "response")
mmC5comp_2
```

```r
confint(pairs(mmC5comp_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_C5comp_3 <-
  glm(mean_C5_comp2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5comp_3, type = "III")
summary(rlogM_C5comp_3)

#No interaction with moderator to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_C5comp_4 <- glm(mean_C5_comp2  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5comp_4,type="III")
summary(rlogM_C5comp_4)

mmC5comp_4_s <- emmeans(rlogM_C5comp_4,~studyarm|edhi,type = "response")
mmC5comp_4_s
confint(pairs(mmC5comp_4_s,reverse = TRUE))

mmC5comp_4_e <- emmeans(rlogM_C5comp_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5comp_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_C5comp_5 <-
  glm(mean_C5_comp2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5comp_5, type = "III")
summary(rlogM_C5comp_5)

mmC5comp_5_s <- emmeans(rlogM_C5comp_5,~studyarm|edhi,type = "response")
mmC5comp_5_s
confint(pairs(mmC5comp_5_s,reverse = TRUE))

```

```r
mmC5comp_5_e <- emmeans(rlogM_C5comp_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5comp_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_C5comp_1 <- glm(mean_C5_comp2  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5comp_1,type="III")
summary(rlogF_C5comp_1)

mfC5comp_1 <- emmeans(rlogF_C5comp_1,~studyarm,type = "response") #proportions
mfC5comp_1
confint(pairs(mfC5comp_1,reverse = TRUE)) #Odds ratios
pairs(mfC5comp_1,reverse = TRUE)

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_C5comp_2 <-
  glm(mean_C5_comp2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5comp_2, type = "III")
summary(rlogF_C5comp_2)

mfC5comp_2 <- emmeans(rlogF_C5comp_2,~studyarm,type = "response")
mfC5comp_2
confint(pairs(mfC5comp_2,reverse = TRUE))
pairs(mfC5comp_2,reverse = TRUE)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_C5comp_3 <-
  glm(mean_C5_comp2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5comp_3, type = "III")
```

```r
summary(rlogF_C5comp_3)

emmip(rlogF_C5comp_3, studyarm ~ mean_collhorz, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_C5comp_4 <- glm(mean_C5_comp2  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5comp_4,type="III")
summary(rlogF_C5comp_4)

mfC5comp_4_s <- emmeans(rlogF_C5comp_4,~studyarm|edhi,type = "response")
mfC5comp_4_s
confint(pairs(mfC5comp_4_s,reverse = TRUE))

mfC5comp_4_e <- emmeans(rlogF_C5comp_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5comp_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_C5comp_5 <-
  glm(mean_C5_comp2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5comp_5, type = "III")
summary(rlogF_C5comp_5)

mfC5comp_5_s <- emmeans(rlogF_C5comp_5,~studyarm|edhi,type = "response")
mfC5comp_5_s
confint(pairs(mfC5comp_5_s,reverse = TRUE))

mfC5comp_5_e <- emmeans(rlogF_C5comp_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5comp_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### +++ Logistic regression for C5 Constraints ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$mean_C5_cons2 <- relevel(as.factor(datGeneric$mean_C5_cons2),ref = "0")
datMeasles$mean_C5_cons2 <- relevel(as.factor(datMeasles$mean_C5_cons2),ref = "0")
datFlu$mean_C5_cons2 <- relevel(as.factor(datFlu$mean_C5_cons2),ref = "0")

#### ++ Generic ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_C5cons_1 <- glm(mean_C5_cons2  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5cons_1,type="III")
summary(rlogG_C5cons_1)
```

```

mgC5cons_1 <- emmeans(rlogG_C5cons_1,~studyarm,type = "response") #proportions
mgC5cons_1
confint(pairs(mgC5cons_1,reverse = TRUE)) #Odds ratios
pairs(mgC5cons_1,reverse = TRUE)

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_C5cons_2 <-
  glm(mean_C5_cons2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")   )
Anova(rlogG_C5cons_2, type = "III")
summary(rlogG_C5cons_2)

mgC5cons_2 <- emmeans(rlogG_C5cons_2,~studyarm,type = "response")
mgC5cons_2
confint(pairs(mgC5cons_2,reverse = TRUE))
pairs(mgC5cons_2,reverse = TRUE)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_C5cons_3 <-
  glm(mean_C5_cons2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5cons_3, type = "III")
summary(rlogG_C5cons_3)

emmip(rlogG_C5cons_3, studyarm ~ mean_indivvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_C5cons_4 <- glm(mean_C5_cons2  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5cons_4,type="III")
summary(rlogG_C5cons_4)

mgC5cons_4_s <- emmeans(rlogG_C5cons_4,~studyarm|edhi,type = "response")
mgC5cons_4_s
confint(pairs(mgC5cons_4_s,reverse = TRUE))
```

```r

mgC5cons_4_e <- emmeans(rlogG_C5cons_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5cons_4_e,reverse = TRUE))

# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_C5cons_5 <-
  glm(mean_C5_cons2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5cons_5, type = "III")
summary(rlogG_C5cons_5)

mgC5cons_5_s <- emmeans(rlogG_C5cons_5,~studyarm|edhi,type = "response")
mgC5cons_5_s
confint(pairs(mgC5cons_5_s,reverse = TRUE))

mgC5cons_5_e <- emmeans(rlogG_C5cons_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5cons_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_C5cons_1 <- glm(mean_C5_cons2  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5cons_1,type="III")
summary(rlogM_C5cons_1)

mmC5cons_1 <- emmeans(rlogM_C5cons_1,~studyarm,type = "response") #proportions
mmC5cons_1
confint(pairs(mmC5cons_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_C5cons_2 <-
  glm(mean_C5_cons2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
```

```r
  )
Anova(rlogM_C5cons_2, type = "III")
summary(rlogM_C5cons_2)

mmC5cons_2 <- emmeans(rlogM_C5cons_2,~studyarm,type = "response")
mmC5cons_2
confint(pairs(mmC5cons_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_C5cons_3 <-
  glm(mean_C5_cons2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5cons_3, type = "III")
summary(rlogM_C5cons_3)

emmip(rlogM_C5cons_3, studyarm ~ mean_indivvertical, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_C5cons_4 <- glm(mean_C5_cons2  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5cons_4,type="III")
summary(rlogM_C5cons_4)

mmC5cons_4_s <- emmeans(rlogM_C5cons_4,~studyarm|edhi,type = "response")
mmC5cons_4_s
confint(pairs(mmC5cons_4_s,reverse = TRUE))

mmC5cons_4_e <- emmeans(rlogM_C5cons_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5cons_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_C5cons_5 <-
  glm(mean_C5_cons2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
```

```
  )
Anova(rlogM_C5cons_5, type = "III")
summary(rlogM_C5cons_5)

mmC5cons_5_s <- emmeans(rlogM_C5cons_5,~studyarm|edhi,type = "response")
mmC5cons_5_s
confint(pairs(mmC5cons_5_s,reverse = TRUE))

mmC5cons_5_e <- emmeans(rlogM_C5cons_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5cons_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_C5cons_1 <- glm(mean_C5_cons2  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5cons_1,type="III")
summary(rlogF_C5cons_1)

mfC5cons_1 <- emmeans(rlogF_C5cons_1,~studyarm,type = "response") #proportions
mfC5cons_1
confint(pairs(mfC5cons_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_C5cons_2 <-
  glm(mean_C5_cons2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5cons_2, type = "III")
summary(rlogF_C5cons_2)

mfC5cons_2 <- emmeans(rlogF_C5cons_2,~studyarm,type = "response")
mfC5cons_2
confint(pairs(mfC5cons_2,reverse = TRUE))
pairs(mfC5cons_2,reverse = TRUE)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_C5cons_3 <-
  glm(mean_C5_cons2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
```

```r
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5cons_3, type = "III")
summary(rlogF_C5cons_3)

emmip(rlogF_C5cons_3, studyarm ~ mean_collhorz, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_C5cons_4 <- glm(mean_C5_cons2  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5cons_4,type="III")
summary(rlogF_C5cons_4)

mfC5cons_4_s <- emmeans(rlogF_C5cons_4,~studyarm|edhi,type = "response")
mfC5cons_4_s
confint(pairs(mfC5cons_4_s,reverse = TRUE))

mfC5cons_4_e <- emmeans(rlogF_C5cons_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5cons_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_C5cons_5 <-
  glm(mean_C5_cons2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5cons_5, type = "III")
summary(rlogF_C5cons_5)

mfC5cons_5_s <- emmeans(rlogF_C5cons_5,~studyarm|edhi,type = "response")
mfC5cons_5_s
confint(pairs(mfC5cons_5_s,reverse = TRUE))

mfC5cons_5_e <- emmeans(rlogF_C5cons_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5cons_5_e,reverse = TRUE))


# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.

#### +++ Logistic regression for C5 Calculation ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$mean_C5_calc2 <- relevel(as.factor(datGeneric$mean_C5_calc2),ref = "0")
datMeasles$mean_C5_calc2 <- relevel(as.factor(datMeasles$mean_C5_calc2),ref = "0")
datFlu$mean_C5_calc2 <- relevel(as.factor(datFlu$mean_C5_calc2),ref = "0")

#### ++ Generic ####

```

```r
#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_C5calc_1 <- glm(mean_C5_calc2  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5calc_1,type="III")
summary(rlogG_C5calc_1)

mgC5calc_1 <- emmeans(rlogG_C5calc_1,~studyarm,type = "response") #proportions
mgC5calc_1
confint(pairs(mgC5calc_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_C5calc_2 <-
  glm(mean_C5_calc2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")   )
Anova(rlogG_C5calc_2, type = "III")
summary(rlogG_C5calc_2)

mgC5calc_2 <- emmeans(rlogG_C5calc_2,~studyarm,type = "response")
mgC5calc_2
confint(pairs(mgC5calc_2,reverse = TRUE))

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_C5calc_3 <-
  glm(mean_C5_calc2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5calc_3, type = "III")
summary(rlogG_C5calc_3)

#No interaction to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogG_C5calc_4 <- glm(mean_C5_calc2  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5calc_4,type="III")
summary(rlogG_C5calc_4)
```

```
mgC5calc_4_s <- emmeans(rlogG_C5calc_4,~studyarm|edhi,type = "response")
mgC5calc_4_s
confint(pairs(mgC5calc_4_s,reverse = TRUE))

mgC5calc_4_e <- emmeans(rlogG_C5calc_4,   ~ edhi|studyarm, type = "response")
confint(pairs(mgC5calc_4_e,reverse = TRUE))

# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_C5calc_5 <-
  glm(mean_C5_calc2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5calc_5, type = "III")
summary(rlogG_C5calc_5)

mgC5calc_5_s <- emmeans(rlogG_C5calc_5,~studyarm|edhi,type = "response")
mgC5calc_5_s
confint(pairs(mgC5calc_5_s,reverse = TRUE))

mgC5calc_5_e <- emmeans(rlogG_C5calc_5,   ~ edhi|studyarm, type = "response")
confint(pairs(mgC5calc_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_C5calc_1 <- glm(mean_C5_calc2   ~ studyarm, data =
datMeasles,family=binomial("logit"))
Anova(rlogM_C5calc_1,type="III")
summary(rlogM_C5calc_1)

mmC5calc_1 <- emmeans(rlogM_C5calc_1,~studyarm,type = "response") #proportions
mmC5calc_1
confint(pairs(mmC5calc_1,reverse = TRUE)) #Odds ratios
pairs(mmC5calc_1,reverse = TRUE)

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_C5calc_2 <-
  glm(mean_C5_calc2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
```

```
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5calc_2, type = "III")
summary(rlogM_C5calc_2)

mmC5calc_2 <- emmeans(rlogM_C5calc_2,~studyarm,type = "response")
mmC5calc_2
confint(pairs(mmC5calc_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_C5calc_3 <-
  glm(mean_C5_calc2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5calc_3, type = "III")
summary(rlogM_C5calc_3)

#No interaction to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_C5calc_4 <- glm(mean_C5_calc2  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5calc_4,type="III")
summary(rlogM_C5calc_4)

mmC5calc_4_s <- emmeans(rlogM_C5calc_4,~studyarm|edhi,type = "response")
mmC5calc_4_s
confint(pairs(mmC5calc_4_s,reverse = TRUE))

mmC5calc_4_e <- emmeans(rlogM_C5calc_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5calc_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_C5calc_5 <-
  glm(mean_C5_calc2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
```

```
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5calc_5, type = "III")
summary(rlogM_C5calc_5)

mmC5calc_5_s <- emmeans(rlogM_C5calc_5,~studyarm|edhi,type = "response")
mmC5calc_5_s
confint(pairs(mmC5calc_5_s,reverse = TRUE))

mmC5calc_5_e <- emmeans(rlogM_C5calc_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5calc_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_C5calc_1 <- glm(mean_C5_calc2  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5calc_1,type="III")
summary(rlogF_C5calc_1)

mfC5calc_1 <- emmeans(rlogF_C5calc_1,~studyarm,type = "response") #proportions
mfC5calc_1
confint(pairs(mfC5calc_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_C5calc_2 <-
  glm(mean_C5_calc2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5calc_2, type = "III")
summary(rlogF_C5calc_2)

mfC5calc_2 <- emmeans(rlogF_C5calc_2,~studyarm,type = "response")
mfC5calc_2
confint(pairs(mfC5calc_2,reverse = TRUE))

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_C5calc_3 <-
  glm(mean_C5_calc2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
```

```
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5calc_3, type = "III")
summary(rlogF_C5calc_3)

#No interaction to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_C5calc_4 <- glm(mean_C5_calc2  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5calc_4,type="III")
summary(rlogF_C5calc_4)

mfC5calc_4_s <- emmeans(rlogF_C5calc_4,~studyarm|edhi,type = "response")
mfC5calc_4_s
confint(pairs(mfC5calc_4_s,reverse = TRUE))

mfC5calc_4_e <- emmeans(rlogF_C5calc_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5calc_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_C5calc_5 <-
  glm(mean_C5_calc2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5calc_5, type = "III")
summary(rlogF_C5calc_5)

mfC5calc_5_s <- emmeans(rlogF_C5calc_5,~studyarm|edhi,type = "response")
mfC5calc_5_s
confint(pairs(mfC5calc_5_s,reverse = TRUE))

mfC5calc_5_e <- emmeans(rlogF_C5calc_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5calc_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.



#### +++ Logistic regression for C5 Collective Response ####

#Create relevel before model to avoid problems with pairs function.
datGeneric$mean_C5_coll2 <- relevel(as.factor(datGeneric$mean_C5_coll2),ref = "0")
datMeasles$mean_C5_coll2 <- relevel(as.factor(datMeasles$mean_C5_coll2),ref = "0")
datFlu$mean_C5_coll2 <- relevel(as.factor(datFlu$mean_C5_coll2),ref = "0")

#### ++ Generic ####
```

```
#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogG_C5coll_1 <- glm(mean_C5_coll2  ~ studyarm, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5coll_1,type="III")
summary(rlogG_C5coll_1)

mgC5coll_1 <- emmeans(rlogG_C5coll_1,~studyarm,type = "response") #proportions
mgC5coll_1
confint(pairs(mgC5coll_1,reverse = TRUE)) #Odds ratios
pairs(mgC5coll_1,reverse = TRUE)


# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogG_C5coll_2 <-
  glm(mean_C5_coll2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")  )
Anova(rlogG_C5coll_2, type = "III")
summary(rlogG_C5coll_2)

mgC5coll_2 <- emmeans(rlogG_C5coll_2,~studyarm,type = "response")
mgC5coll_2
confint(pairs(mgC5coll_2,reverse = TRUE))
pairs(mgC5coll_2,reverse = TRUE)

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogG_C5coll_3 <-
  glm(mean_C5_coll2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5coll_3, type = "III")
summary(rlogG_C5coll_3)

emmip(rlogG_C5coll_3, studyarm ~ mean_indivvertical, cov.reduce = range, type = "response")
emmip(rlogG_C5coll_3, studyarm ~ mean_collhorz, cov.reduce = range, type = "response")

#### + Model 4: Check for interaction viz*edhi without any covariates  ####
```

```r

rlogG_C5coll_4 <- glm(mean_C5_coll2  ~ studyarm * edhi, data = datGeneric,family=binomial("logit"))
Anova(rlogG_C5coll_4,type="III")
summary(rlogG_C5coll_4)

mgC5coll_4_s <- emmeans(rlogG_C5coll_4,~studyarm|edhi,type = "response")
mgC5coll_4_s
confint(pairs(mgC5coll_4_s,reverse = TRUE))
pairs(mgC5coll_4_s,reverse = TRUE)

mgC5coll_4_e <- emmeans(rlogG_C5coll_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5coll_4_e,reverse = TRUE))
pairs(mgC5coll_4_e,reverse = TRUE)

# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogG_C5coll_5 <-
  glm(mean_C5_coll2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datGeneric,family=binomial("logit")
  )
Anova(rlogG_C5coll_5, type = "III")
summary(rlogG_C5coll_5)

mgC5coll_5_s <- emmeans(rlogG_C5coll_5,~studyarm|edhi,type = "response")
mgC5coll_5_s
confint(pairs(mgC5coll_5_s,reverse = TRUE))

mgC5coll_5_e <- emmeans(rlogG_C5coll_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mgC5coll_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Measles ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogM_C5coll_1 <- glm(mean_C5_coll2  ~ studyarm, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5coll_1,type="III")
summary(rlogM_C5coll_1)

mmC5coll_1 <- emmeans(rlogM_C5coll_1,~studyarm,type = "response") #proportions
mmC5coll_1
confint(pairs(mmC5coll_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogM_C5coll_2 <-
  glm(mean_C5_coll2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
```

```
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5coll_2, type = "III")
summary(rlogM_C5coll_2)

mmC5coll_2 <- emmeans(rlogM_C5coll_2,~studyarm,type = "response")
mmC5coll_2
confint(pairs(mmC5coll_2,reverse = TRUE))


# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogM_C5coll_3 <-
  glm(mean_C5_coll2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5coll_3, type = "III")
summary(rlogM_C5coll_3)

#No interaction to interpret

#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogM_C5coll_4 <- glm(mean_C5_coll2  ~ studyarm * edhi, data = datMeasles,family=binomial("logit"))
Anova(rlogM_C5coll_4,type="III")
summary(rlogM_C5coll_4)

mmC5coll_4_s <- emmeans(rlogM_C5coll_4,~studyarm|edhi,type = "response")
mmC5coll_4_s
confint(pairs(mmC5coll_4_s,reverse = TRUE))

mmC5coll_4_e <- emmeans(rlogM_C5coll_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5coll_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogM_C5coll_5 <-
  glm(mean_C5_coll2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
```

```
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datMeasles,family=binomial("logit")
  )
Anova(rlogM_C5coll_5, type = "III")
summary(rlogM_C5coll_5)

mmC5coll_5_s <- emmeans(rlogM_C5coll_5,~studyarm|edhi,type = "response")
mmC5coll_5_s
confint(pairs(mmC5coll_5_s,reverse = TRUE))

mmC5coll_5_e <- emmeans(rlogM_C5coll_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mmC5coll_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### ++ Flu ####

#### + Model 1: Check for direct effects of factors without any covariates ####

options(contrasts=c("contr.sum", "contr.poly"))
rlogF_C5coll_1 <- glm(mean_C5_coll2  ~ studyarm, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5coll_1,type="III")
summary(rlogF_C5coll_1)

mfC5coll_1 <- emmeans(rlogF_C5coll_1,~studyarm,type = "response") #proportions
mfC5coll_1
confint(pairs(mfC5coll_1,reverse = TRUE)) #Odds ratios

# Model 2: Check for direct effects of factors with adjustment for other covariates
rlogF_C5coll_2 <-
  glm(mean_C5_coll2  ~ studyarm +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5coll_2, type = "III")
summary(rlogF_C5coll_2)

mfC5coll_2 <- emmeans(rlogF_C5coll_2,~studyarm,type = "response")
mfC5coll_2
confint(pairs(mfC5coll_2,reverse = TRUE))

# Model 3: Check for moderating effects of individualism & collectivism with adjustment for other covariates
rlogF_C5coll_3 <-
  glm(mean_C5_coll2  ~
        studyarm * mean_indivhorz +
        studyarm * mean_indivvertical +
```

```
        studyarm * mean_collhorz +
        studyarm * mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        edhi +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5coll_3, type = "III")
summary(rlogF_C5coll_3)

emmip(rlogF_C5coll_3, studyarm ~ mean_collhorz, cov.reduce = range, type = "response")


#### + Model 4: Check for interaction viz*edhi without any covariates  ####

rlogF_C5coll_4 <- glm(mean_C5_coll2  ~ studyarm * edhi, data = datFlu,family=binomial("logit"))
Anova(rlogF_C5coll_4,type="III")
summary(rlogF_C5coll_4)

mfC5coll_4_s <- emmeans(rlogF_C5coll_4,~studyarm|edhi,type = "response")
mfC5coll_4_s
confint(pairs(mfC5coll_4_s,reverse = TRUE))

mfC5coll_4_e <- emmeans(rlogF_C5coll_4,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5coll_4_e,reverse = TRUE))


# Model 5: Check for interaction viz*edhi with adjustment for other covariates

rlogF_C5coll_5 <-
  glm(mean_C5_coll2   ~ studyarm*edhi +
        mean_indivhorz +
        mean_indivvertical +
        mean_collhorz +
        mean_collvertical +
        bornincanada +
        language_1 +
        language_2 +
        Asian_group +
        White_group +
        disability_any +
        genderidentity +
        income +
        age,
      data = datFlu,family=binomial("logit")
  )
Anova(rlogF_C5coll_5, type = "III")
summary(rlogF_C5coll_5)

mfC5coll_5_s <- emmeans(rlogF_C5coll_5,~studyarm|edhi,type = "response")
mfC5coll_5_s
confint(pairs(mfC5coll_5_s,reverse = TRUE))

mfC5coll_5_e <- emmeans(rlogF_C5coll_5,  ~ edhi|studyarm, type = "response")
confint(pairs(mfC5coll_5_e,reverse = TRUE))

# Now running models on imputed data.
#CHANGELOG ASJ: No missing data for c5conf, no model with imputation performed, 21-09-20.


#### COMPLETES VS INCOMPLETES ####
```

```r

# Here we test for differences between people who finished the survey (completes) and those who
# started but did not finish  (incompletes) to explore potential issues or limitations introduced by
# imputing missing data. To do this, we run Chi-squared tests on categorical variables and Wilcoxon
# tests on continuous variables, comparing people who did and did not complete the survey. We ran
# the sociodemographic panel of questions first to maximally enable this comparison but we won't be
# able to compare with people who never clicked on the survey or who started to read and didn't
# answer anything.

#CHANGELOG ASJ: Compare people analysed who finished with quality answer  VS (did not finished or finished with poor quality), 2021-08-06.
table(datConsent$used) #CHANGELOG ASJ: changed finished to used, and everywhere below 2021-08-06
prop.table(table(datConsent$used))*100

#CHANGELOG ASJ: Apply the same changes to the variables as in the dataset dat, 2021-09-01
datConsent$yearofbirth <- as.numeric(datConsent$yearofbirth)
datConsent$yearofbirth[datConsent$yearofbirth==888] <- NA
datConsent$age <- 2021 - datConsent$yearofbirth


cols.num <-
  c(
    "bornincanada",
    "language_1",
    "language_2",
    "language_3",
    "language_888",
    "ethnicity_1",
    "ethnicity_2",
    "ethnicity_3",
    "ethnicity_4",
    "ethnicity_5",
    "ethnicity_6",
    "ethnicity_7",
    "ethnicity_8",
    "ethnicity_9",
    "ethnicity_10",
    "ethnicity_11",
    "ethnicity_12",
    "ethnicity_13",
    "ethnicity_14",
    "ethnicity_15",
    "ethnicity_16",
    "ethnicity_17",
    "ethnicity_888",
    "disability",
    "techdisability",
    "sexatbirth",
    "genderidentity",
    "supportedbyincome",
    "educationlevel",
    "studyarm",
    "viz",
    "disease",
    "immune",
    "c19vax"
  )
datConsent[cols.num] <- lapply(datConsent[cols.num],as.numeric)

datConsent$language_1[is.na(datConsent$language_1)] <- 0
datConsent$language_2[is.na(datConsent$language_2)] <- 0
datConsent$language_3[is.na(datConsent$language_3)] <- 0
```

```
datConsent$language_1[datConsent$language_888==1&datConsent$language_1==0&datConsent$language_2==0&datConsent$language_3==0] <- NA
datConsent$language_2[datConsent$language_888==1&datConsent$language_1==0&datConsent$language_2==0&datConsent$language_3==0] <- NA
datConsent$language_3[datConsent$language_888==1&datConsent$language_1==0&datConsent$language_2==0&datConsent$language_3==0] <- NA

datConsent$ethnicity_1[is.na(datConsent$ethnicity_1)] <- 0
datConsent$ethnicity_2[is.na(datConsent$ethnicity_2)] <- 0
datConsent$ethnicity_3[is.na(datConsent$ethnicity_3)] <- 0
datConsent$ethnicity_4[is.na(datConsent$ethnicity_4)] <- 0
datConsent$ethnicity_5[is.na(datConsent$ethnicity_5)] <- 0
datConsent$ethnicity_6[is.na(datConsent$ethnicity_6)] <- 0
datConsent$ethnicity_7[is.na(datConsent$ethnicity_7)] <- 0
datConsent$ethnicity_8[is.na(datConsent$ethnicity_8)] <- 0
datConsent$ethnicity_9[is.na(datConsent$ethnicity_9)] <- 0
datConsent$ethnicity_10[is.na(datConsent$ethnicity_10)] <- 0
datConsent$ethnicity_11[is.na(datConsent$ethnicity_11)] <- 0
datConsent$ethnicity_12[is.na(datConsent$ethnicity_12)] <- 0
datConsent$ethnicity_13[is.na(datConsent$ethnicity_13)] <- 0
datConsent$ethnicity_14[is.na(datConsent$ethnicity_14)] <- 0
datConsent$ethnicity_15[is.na(datConsent$ethnicity_15)] <- 0
datConsent$ethnicity_16[is.na(datConsent$ethnicity_16)] <- 0
datConsent$ethnicity_17[is.na(datConsent$ethnicity_17)] <- 0


datConsent_ethnicity <-
  subset(
    datConsent,
    select = c(
      ethnicity_1,
      ethnicity_2,
      ethnicity_3,
      ethnicity_4,
      ethnicity_5,
      ethnicity_6,
      ethnicity_7,
      ethnicity_8,
      ethnicity_9,
      ethnicity_10,
      ethnicity_11,
      ethnicity_12,
      ethnicity_13,
      ethnicity_14,
      ethnicity_15,
      ethnicity_16,
      ethnicity_17
    )
  )

datConsent$sum_ethnicity <-rowSums(datConsent_ethnicity,na.rm=T)

datConsent$ethnicity_1[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_2[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_3[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_4[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_5[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_6[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_7[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_8[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_9[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_10[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_11[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_12[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_13[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_14[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_15[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
datConsent$ethnicity_16[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA
```

```r
datConsent$ethnicity_17[datConsent$ethnicity_888==1&datConsent$sum_ethnicity==0] <- NA

datConsent$Asian_group <-
  ifelse(
    datConsent$ethnicity_1 == 1 |
      datConsent$ethnicity_2 == 1 |
      datConsent$ethnicity_3 == 1 |
      datConsent$ethnicity_4 == 1,
    1,
    0
  )


datConsent$White_group <-
  ifelse(
    datConsent$ethnicity_6 == 1 |
      datConsent$ethnicity_7 == 1,
    1,
    0
  )

datConsent$edhi <- ifelse(datConsent$educationlevel == 4 |
                            datConsent$educationlevel == 5|
                            datConsent$educationlevel == 6|
                            datConsent$educationlevel == 7, TRUE,
                          ifelse(datConsent$educationlevel == 8 |
                          datConsent$educationlevel == 888, NA,FALSE))


datConsent$bornincanada[datConsent$bornincanada==888] <- NA
datConsent$disability[datConsent$disability==888] <- NA
datConsent$techdisability[datConsent$techdisability==888] <- NA
datConsent$sexatbirth[datConsent$sexatbirth==888] <- NA
datConsent$genderidentity[datConsent$genderidentity==888] <- NA
datConsent$income[datConsent$income==888] <- NA
datConsent$income[datConsent$income==5] <- NA
datConsent$supportedbyincome[datConsent$supportedbyincome==888] <- NA
datConsent$educationlevel[datConsent$educationlevel==888] <- NA
datConsent$educationlevel[datConsent$educationlevel==8] <- NA

datConsent$agecat[datConsent$agecat==888 | datConsent$agecat==1] <- NA


datConsent$bornincanada <-
  factor(datConsent$bornincanada,
         levels = c(1, 0),
         labels = c("Yes", "No"))

datConsent$language_1 <-
  factor(datConsent$language_1,
         levels = c(1, 0),
         labels = c("Yes", "No"))

datConsent$language_2 <-
  factor(datConsent$language_2,
         levels = c(1, 0),
         labels = c("Yes", "No"))

datConsent$Asian_group <-
  factor(datConsent$Asian_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))


datConsent$White_group <-
  factor(datConsent$White_group,
         levels = c(1, 0),
         labels = c("Yes", "No"))

```

```

datConsent$disability <-
  factor(
    datConsent$disability,
    levels = c(0, 1),
    labels = c("No disability", "yes,at least one")
  )

datConsent$techdisability <-
  factor(datConsent$techdisability,
         levels = c(0, 1),
         labels = c("No", "Yes"))

datConsent$sexatbirth <-
  factor(datConsent$sexatbirth,
         levels = c(1, 2),
         labels = c("Female", "Male"))

datConsent$genderidentity <- factor(
  datConsent$genderidentity,
  levels = c(1, 2, 3, 4),
  labels = c(
    "Female",
    "Male",
    "Indigenous or other cultural gender minority identity",
    "Something else"
  )
)

datConsent$income <-
  factor(
    datConsent$income,
    levels = c(1, 2, 3, 4),
    labels = c("24 999 or less", "25 000 to 49 999",
               "50 000 to 99 999", "100 000 or more")
  )

datConsent$educationlevel <- factor(
  datConsent$educationlevel,
  levels = c(1, 2, 3, 4, 5, 6, 7),
  labels = c(
    "Some Elementary School",
    "High School Diploma",
    "Apprenticeship or trade certificate or diploma",
    "College or polytechnical school certificate or diploma",
    "University degree bachelor level or below",
    "University graduate degree Master level",
    "University graduate degree Doctorate level"
  )
)

datConsent$agecat<-
  factor(
    datConsent$agecat,
    levels = c(2, 3, 4),
    labels = c("18-34 years", "35-49 years", "50 years and more")
  )

datConsent$UserLanguage <- as.factor(datConsent$UserLanguage)

datConsent$studyarm <-
  factor (
    datConsent$studyarm,
    levels = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13),
    labels = c(
      "controlgeneric",
      "controlmeasles",
      "controlpertussis",
      "controlflu",
```

```
      "herdimmgeneric",
      "herdimmmeasles",
      "herdimmpertussis",
      "herdimmflu",
      "robertkochgeneric",
      "sbsnewsgeneric",
      "guardianmeasles",
      "theotheredmundmeasles",
      "publichealthagencycanadaflu"
    )
  )

datConsent$viz <- factor (
  datConsent$viz,
  levels = c(0, 1, 2, 3, 5, 6, 7),
  labels = c(
    "no visualization",
    "herdimm",
    "robertkoch",
    "sbsnews",
    "guardian",
    "theotheredmund",
    "publichealthagencycanada"
  )
)

datConsent$disease <-
  factor (
    datConsent$disease,
    levels = c(0, 1, 2, 3),
    labels = c("Generic", "Measles", "Pertussis", "Flu")
  )


datConsent$genderidentity[datConsent$genderidentity == "Something else" |
                            datConsent$genderidentity == "Indigenous or other cultural
                            gender minority identity"] <- NA
datConsent$genderidentity <- droplevels(datConsent$genderidentity)


datConsent$disability_any <-ifelse(datConsent$disability =="yes,at least one" |
datConsent$techdisability == "Yes", 1, 0)
datConsent$disability_any <-
  factor(
    datConsent$disability_any,
    levels = c(1, 0),
    labels = c("Yes", "No")
  )

datConsent$agecat <- as.numeric(datConsent$agecat)

datConsent$used <-
  factor(
    datConsent$used,
    levels = c(1, 0),
    labels = c("Yes", "No")
  )

# For each categorical demographic to do along with the next step of chi square
#CHANGELOG ASJ: Add table for other demographic variables 21-08-31
t1<-data.frame(datConsent$used,datConsent$studyarm)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$viz)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$disease)
```

```
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$bornincanada)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$language_1)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$language_2)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$White_group)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$Asian_group)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$disability)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$techdisability)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$disability_any)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$sexatbirth)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$genderidentity)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$income)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$educationlevel)
table(t1)
prop.table(table(t1), 1) * 100

t1<-data.frame(datConsent$used,datConsent$edhi)
table(t1)
prop.table(table(t1), 1) * 100

#Comparison for continuous demographics
describeBy(datConsent$age,datConsent$used)
wilcox.test(datConsent$age~datConsent$used)

#CHANGELOG ASJ: Only keep variables analysed with >5% and dichotomous versions 2021-08-06
chisq.test(datConsent$used,datConsent$studyarm, correct=FALSE)
chisq.test(datConsent$used,datConsent$viz, correct=FALSE)
chisq.test(datConsent$used,datConsent$disease, correct=FALSE)
chisq.test(datConsent$used,datConsent$bornincanada, correct=FALSE)
chisq.test(datConsent$used,datConsent$language_1, correct=FALSE)
chisq.test(datConsent$used,datConsent$language_2, correct=FALSE)
chisq.test(datConsent$used,datConsent$Asian_group, correct=FALSE)
chisq.test(datConsent$used,datConsent$White_group, correct=FALSE)
chisq.test(datConsent$used,datConsent$disability, correct=FALSE)
```

```
chisq.test(datConsent$used,datConsent$techdisability, correct=FALSE)
chisq.test(datConsent$used,datConsent$disability_any, correct=FALSE)
chisq.test(datConsent$used,datConsent$sexatbirth, correct=FALSE)
chisq.test(datConsent$used,datConsent$genderidentity, correct=FALSE)
chisq.test(datConsent$used,datConsent$income, correct=FALSE)
chisq.test(datConsent$used,datConsent$educationlevel, correct=FALSE)
chisq.test(datConsent$used,datConsent$edhi, correct=FALSE)


#### RESEARCH QUESTION 4: Effects of male vs female voice ####

# Recall: Research question 4 (add-on, may be reported separately from Hina Hakim's thesis): Are the
# effects of the herdimm intervention different when the narration uses a female or male voice?

# Force this variable to be read as a factor and define the labels (this could be up in the data
# cleaning section but we are keeping it separate in case we don't include this in Hina's thesis.)

dat$voice <- factor(dat$voice,levels = c(1,2),labels = c("Male", "Female"))

# Create 2 new variables combining voice and study arm to facilitate analyses

dat$viz_voice <- as.factor(paste(dat$viz,dat$voice,sep="_"))
dat$studyarm_voice <- as.factor(paste(dat$studyarm,dat$voice,sep="_"))

# Effect of male versus female voice in herdimm narration

datHerdimm <-
  subset(
    dat,
    dat$studyarm == "herdimmgeneric" |
      dat$studyarm == "herdimmmeasles" |
      dat$studyarm == "herdimmpertussis" |
      dat$studyarm == "herdimmflu"
  )

# Analysis for risk perception, adjusted for covariates and moderators, with interaction of gender
# of respondents
options(contrasts=c("contr.sum", "contr.poly")) # CHANGELOG: Added to calculate properly the type III sum of squares with lm 2021-07-27
mHV_RP <- lm(
  mean_riskperception2_6 ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada + #CHANGELOG ASJ: covariate list shortened 2021-08-04
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm
)
Anova(mHV_RP,type="III")
summary(mHV_RP)


# The same analysis will be done for the other outcomes, and the model will be validated as for the
# main analyses. If there is an effect of voice in the above model, the 2-way ANOVA
```

```
model
# (viz*disease) and one-way anova model (studyarm) per disease will be done again. In these new
# models, Herdimm technique will be divided according to the voice gender. The factors for the
# two-way anova will be viz_voice and disease. The factor for the one-way anova will be
# studyarm_voice.

#CHANGELOG ASJ: Add this model for other continuous outcomes, 2021-08-05.

mHV_emot <- lm(
  mean_emotion ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm
)
Anova(mHV_emot,type="III")

mHV_know <- lm(
  sum_knowledge ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm
)
Anova(mHV_know,type="III")

#Logistic regressions
rlHV_RP1 <- glm(
  relevel(RP1_dicho,ref="Low") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
```

```
Anova(rlHV_RP1,type="III")

rlHV_c19vi <- glm(
  relevel(c19vaxintent,ref="Low") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_c19vi,type="III")


rlHV_vi <- glm(
  relevel(vaxintent,ref="Low") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_vi,type="III")

rlHV_Hti <- glm(
  relevel(Htrustinfo,ref="No") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_Hti,type="III")

#CHANGELOG ASJ: Add same analysis for C5 subscales, 21-09-20.
rlHV_C5conf <- glm(
  relevel(as.factor(mean_C5_conf2),ref="0") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
```

```
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_C5conf,type="III")

rlHV_C5cons <- glm(
  relevel(as.factor(mean_C5_cons2),ref="0") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_C5cons,type="III")

rlHV_C5comp <- glm(
  relevel(as.factor(mean_C5_comp2),ref="0") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_C5comp,type="III")


rlHV_C5calc <- glm(
  relevel(as.factor(mean_C5_calc2),ref="0") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
```

```
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_C5calc,type="III")


rlHV_C5coll <- glm(
  relevel(as.factor(mean_C5_coll2),ref="0") ~
    studyarm * voice * genderidentity +
    mean_indivhorz +
    mean_indivvertical +
    mean_collhorz +
    mean_collvertical +
    bornincanada +
    language_1 +
    language_2 +
    Asian_group +
    White_group +
    disability_any +
    income +
    edhi +
    age,
  data = datHerdimm,family=binomial("logit")
)
Anova(rlHV_C5coll,type="III")


#### END ####


```