Title

Statistical Analysis Plan for the SITA trial: An RCT of a decision aid to support informed choices about taking aspirin to prevent colorectal cancer and other chronic diseases

Date and version

19/05/2022 Version 1.0

Trial registration

SITA was prospectively registered on the Australian and New Zealand Clinical Trials Registry (ACTRN12620001003965) in August 2020. Participant recruitment commenced in October 2020 and was completed in April 2021. Follow-up and process data collection was completed in November 2021, 7 months after the last participant was recruited, allowing for an additional month to follow-up late responders of the 6-month follow up questionnaires. https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12620001003965

Funding acknowledgement

This trial is funded by a dedicated grant from the Victorian Cancer Agency ID: CPSRG19011. The University of Melbourne is the sponsor of the present trial. We would also like to acknowledge the general practice staff and patients for their participation in the SITA trial. The authors would like to acknowledge Primary Care Collaborative Cancer Clinical Trials Group (PC4) for supporting this project. They also acknowledge support by the CanTest Collaborative (funded by Cancer Research UK, C8640/A23385) of which Sibel Saya is affiliated researcher, Jon Emery is an Associate Director and Fiona M. Walter is Director.

Protocol reference

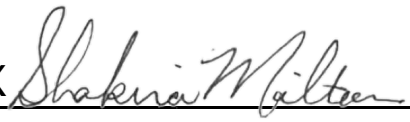
Milton, S., McIntosh, J., Macrae, F. et al. An RCT of a decision aid to support informed choices about taking aspirin to prevent colorectal cancer and other chronic diseases: a study protocol for the SITA (Should I Take Aspirin?) trial. Trials 22, 452 (2021). https://doi.org/10.1186/s13063-021-05365-8

SAP Revision History

Protocol version Updated	SAP version no.	Section number changed	Description and reason for change	Date changed
Version 1	1.0			

Roles and responsibilities

Signatory names and contributions

Name	Role	Affiliation	Signature
Shakira Milton*	Trial coordinator/ project manager/ data analyst	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia	X  Shakira Milton Data analyst
Patty Chondros	Senior statistician	Department of General Practice, University of Melbourne, Melbourne, Australia	X  Patty Chondros Senior statistician

Jon Emery	Chief investigator/ Project Lead	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia	X Jon Emery Chief investigator

*Contact person

Non-signatory names and contributions

Name	**Role**	**Affiliation**
Jennifer McIntosh	Investigator	Department of Software Systems & Cybersecurity, Monash University, Melbourne, Australia
Finlay Macrae	Investigator	Colorectal Medicine and Genetics, The Royal Melbourne Hospital, Melbourne, Australia and Department of Medicine, The University of Melbourne, Melbourne, Australia
Lyndal Trevena	Investigator	Faculty of Medicine and Health, School of Public Health, The University of Sydney, Sydney Australia
Mark Jenkins	Investigator	Melbourne School of Population and Global Health, University of Melbourne, Melbourne Australia
Fiona M. Walter	Investigator	Department of General Practice, University of Melbourne, Melbourne, Australia and The Primary Care Unit, University of Cambridge, Cambridge, United Kingdom
Natalie Taylor	Investigator	Behavioral and Implementation Research and Evaluation, Cancer Council NSW, New South Wales, Australia and Faculty of Medicine and Health, University of Sydney, Sydney, Australia
Lucy Boyd	Research assistant	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia
Sibel Saya	Investigator	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia
Rushani Wijesuriya	Statistician	Department of General Practice, University of Melbourne, Melbourne, Australia
Napin Karnchanachari	Research assistant	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia
Kitty Novy	Research assistant	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia
Carmody Forbes	Graphic designer	University of Melbourne, Melbourne, Australia
Javiera Martinez Gutierrez	Investigator	Department of General Practice, Centre for Cancer Research, University of Melbourne, Melbourne, Australia
Kate Broun	Investigator	Early Detection and Immunisation, Prevention Department, Cancer Council Victoria, Australia
Sara Whitburn	Investigator	Belmore Road Medical Clinic, Melbourne, Australia
Sarah McGill	Investigator	Cancer Screening and Prevention, Cancer Institute NSW, Australia
George Fishman	Consumer	Primary Care Collaborative Cancer Clinical Trials Group (PC4), Community Advisory Group, University of Melbourne, Australia
Julie Marker	Consumer	Primary Care Collaborative Cancer Clinical Trials Group (PC4), Community Advisory Group, University of Melbourne, Australia
Max Shub	Consumer	Primary Care Collaborative Cancer Clinical Trials Group (PC4), Community Advisory Group, University of Melbourne, Australia

SITA SAP

Contact Person:

Shakira Milton
Department of General Practice
Centre for Cancer Research
The University of Melbourne
email: Shakira.Milton@unimelb.edu.au

Abbreviations

RA – Research Assistant

EFT – Expected Frequency Trees

CRC – Colorectal cancer

GP – General practitioner

ANZCTR – Australia New Zealand Clinical Trial Registry

WHS – Women's Health Study

VCA – Victorian Cancer Agency

Keywords

Preventive medicine, General practice, Primary care, Cancer prevention, Bowel cancer, Colorectal cancer, Aspirin, Guideline implementation, Chemoprevention, Decision Aid, Informed decision making

Background

This trial, SITA, is a stratified individually randomised controlled trial (RCT) with general practice patients that aims to test the efficacy of a health consultation and use of a sex-specific decision aid, using an expected frequency tree (EFT) to present the benefits and harms of taking low dose aspirin, on informed decision-making at one month and uptake of aspirin at six-months. The decision aids convey the Cancer Council Australia aspirin guidelines which recommend that all people aged 50-70 years old actively consider taking daily low-dose aspirin (100–300mg per day) for 2.5 to 5 years to reduce their risk of colorectal cancer (CRC).(1) Control participants receive general information about modifiable risk factors for CRC prevention. The study rationale, and details of the study design, including setting, eligibility criteria, sample size calculations and statistical analysis are detailed in the published study trial protocol.(2) This document provides a detailed statistical analysis plan, to complement the study protocol and to expand on the secondary and sensitivity analyses.

Objectives

The two equally important objectives are to determine if the EFT-based decision aid, used in a health consultation compared with general CRC prevention information in general practice patients between 50 and 70 years old:

1. increases informed decision-making related to taking aspirin at one-month and
2. increases self-reported use of aspirin at six-months

Secondary objectives are to compare the novel EFT-based decision aid, used in a health consultation compared with general CRC prevention information in general practice patients between 50 and 70 years old with respect to:

1) self-reported use of aspirin at one-month

2) lower mean decisional conflict at one-month

4) self-reported changes in other behaviours to reduce the risk of CRC (e.g., dietary, quitting smoking, or having a screening test for CRC).

Primary hypotheses

There are two primary hypotheses:

1) The first null hypothesis is that there is no difference on informed decision-making at one-month for general practice patients between 50 and 70 years old who receive the EFT-based decision aid, used in health consultation and general CRC prevention information.

2) The second null hypothesis is that there is no difference in aspirin uptake at six-months for general practice patients between 50 and 70 years old who receive the EFT-based decision aid, used in health consultation and general CRC prevention information.

Trial methods

The teletrial methods included calling patients who were scheduled to see their general practitioner (GP) on the day or following day, and if interested, we checked their eligibility over the phone, and then invited them to participate in the trial either in the clinic via face-face or online via a Zoom appointment. Figure 1 shows the CONSORT diagram for the trial recruitment.

Sample size

For 80% power and a Bonferroni adjusted 2-sided alpha level of 2.5% to account for the two co-primary outcomes [29], we required 258 participants (129 per arm) to detect a minimum 20% difference, as decided on by the trial steering committee. Further justification for the sample size can be found in the study protocol.

Eligibility criteria

Participants were eligible if they were: i) aged between 50 and 70 years old and had an appointment with their GP on the day of recruitment or on the following day ii) were able to read and understand written English, and iii) competent to give informed consent.

General Practice clinics were recruited for the trial. The inclusion criteria for the clinic were that they had at least three full-time GPs and were not a COVID-19 testing clinic. The aim was to recruit a population of participants which were representative of the Victorian population in socio-economic status and education, so recruiting from regional Victoria was imperative. Detailed exclusion criteria and inclusion and exclusion criteria can be found in the protocol.

Consent and recruitment

All GPs and patient participants provided either written or electronic consent to participate in the trial. GPs consented to us approaching their patients while patients consented to being randomised into the trial and either received the intervention or control.

Two research assistants (RAs) at a time worked together to recruit the participants from six general practice clinics around Victoria, Australia. Participant recruitment commenced from 12th October 2020 was completed on 22nd April 2021. Participants were followed up after one and six-months which was completed on 26th May 2021 and 23rd November 2021, respectively. Participants received automatic reminders to complete the follow up questionnaires if they opted into receiving them via email. Follow up reminders were given to all participants over the phone by a third research assistant who was blinded to the intervention.

Baseline characteristics and outcomes

Screening and baseline data collection

At screening, the total number of participants approached, whether they were eligible, reasons for not meeting eligibility criteria, as well their age and sex were recorded. Participant demographic characteristics were captured at baseline. We asked for participants' age, gender (male, female, or variations of sex characteristics), home postcode, country of birth, education (never completed high school, high school only, TAFE or similar, or University degree or higher), how many medications they are taking, their living arrangements as whether they live alone (yes or no), and languages spoken at home. Participants' postcodes of residence at baseline will be linked with the Socio-Economic Indexes for Areas (SEIFA) Index of Relative Socio-economic Advantage and Disadvantage (IRSAD) from the Australian Bureau of Statistics (3) to describe the socio-economic status of the study sample. This IRSAD ranks all postcodes in Australia into index scores are based on an arbitrary numerical scale of both advantage and disadvantage, then divides them into deciles, one being the most disadvantaged in socio-economic status, 10 the most advantaged in socio-economic status. The Australian Bureau of Statistics defines relative socio-economic advantage and disadvantage in terms of people's access to material and social resources, and their ability to participate in society.(4) The IRSAD will be recoded, in STATA 17, from 10 to five deciles to show the diversity of socio-economic status of the sample. Country of birth will be dichotomised into either born in Australia or born overseas.

Participants' cardiovascular disease risk factors will be self-reported by answering the following questions (yes, no, or unsure): a family history of heart attack, angina, or stroke; a personal history of diabetes; medication for high blood pressure; personal history of high cholesterol; and a personal history of smoking cigarettes. Similarly, participants' CRC

familial risk will be self-reported by answering the following (yes, no, or unsure): a family history of CRC (parent, brother, sister, children) diagnosed before 55 years old, and more than one relative who had CRC at any age (parents, children, brothers, sister, grandparents, aunts, uncles, nieces, nephews, and grandchildren).

The Subjective numeracy scale (SNS) is a self-reported, validated (5) measure about preferences for numerical versus prose information and perceived ability to perform mathematical tasks. It is an eight-item scale, with four questions asking participants to assess their numerical ability and four questions asking them to state their preference for numerical or probabilistic information. Each item is rated on a six-point Likert scale. The eight items included in the subjective numeracy scale can be found in box 1. To calculate a total score, each item's score is summed then divided by eight for an average, the total number of questions (after reverse coding the "seventh question), with the total score range from one to six. A larger score indicates a higher subjective rating of numeracy abilities and preferences.

Box 1: Items of the numeracy scale (6)

For each of the following questions, please check the box that best reflects how good you are at doing the following things:

1. How good are you at working with fractions?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Not at all good					Extremely good

2. How good are you at working with percentages?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Not at all good					Extremely good

3. How good are you at calculating a 15% tip?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Not at all good					Extremely good

4. How good are you at figuring out how much a shirt will cost if it is 25% off?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Not at all good					Extremely good

For each of the following questions, please check the box that best reflects your answer:

1. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Always prefer words					Always prefer numbers

2. When people tell you the chance of something happening, do you prefer that they use words ("it rarely happens") or numbers ("there's a 1% chance")?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Always prefer words					Always prefer numbers

3. When you hear a weather forecast, do you prefer predictions using percentages (e.g., “there will be a 20% chance of rain today”) or predictions using only words (e.g., “there is a small chance of rain today”)?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Always prefer percentages					Always prefer words

4. How often do you find numerical information to be useful?

❑ 1	❑ 2	❑ 3	❑ 4	❑ 5	❑ 6
Never					Very often

Outcomes

Outcomes were measured at baseline before randomisation and one and six-months after randomisation.
Two co-primary outcomes will be assessed for the trial.

The **first co-primary outcome** is the difference in the proportion between the study arms of participants who made an informed decision about taking aspirin at one-month measured using the Multidimensional Measure of Informed Choice (MMIC). (7)

Multi-dimensional measure of informed choice

An informed choice is one where all the available information, as presented in the decision aids about aspirin chemoprevention is weighed up and used to inform the final decision; participants' choice should be consistent with their attitude.(8)

The MMIC consists of three domains: 1) the participant's knowledge of the aspirin advice covered that was delivered as part of the intervention at baseline; 2) their attitudes toward taking aspirin (positive or negative) and; 3) their behaviour as to whether they decided to take aspirin at one-month. Informed choices are those of participants who had sufficient knowledge and an attitude about taking aspirin which was consistent with their behaviour. Other choices, for example where participants show inadequate knowledge and/or their choice to take aspirin is not consistent with their attitude, are defined as uninformed.

1. ***Knowledge score***

The knowledge score consists of 12 items, 11 statements that require a true, false, or unsure response and one open ended item (see Box 2). Participants receive one point for every correct answer for the 11 items, and one point each (up to four points total) for each correct response to the open-ended question. All unsure responses of participants will be coded as an incorrect answer, which would be consistent with the participant having inadequate knowledge for the item. All responses left blank from the open-ended item will be coded as incorrect reflecting the participant having insufficient knowledge.

The items are summed to provide a total knowledge score ranging between zero to 15, with higher scores indicating greater knowledge. This total score will be dichotomised as sufficient knowledge for an informed choice or not based on a cut-off that will be set according to the Angoff (9) method. This method entails a panel of subject matter experts (from the authors: JM, FM, PC and JE) work through each knowledge item independently and decide a cut-off score for each.

The Angoff methods requires each subject matter expert to independently imagine 100 minimally competent individuals completing the 12 knowledge items and then estimate how many of these 100 individuals (*n)* would answer each item correctly. After the individual scoring of the knowledge items, the subject matter experts will then openly negotiate the scoring for each item and will have the opportunity to change their score if there is too much variation compared to the others' scoring. A minimum passing level (MPL) will be decided on for each knowledge item and the cut-off score for the overall scale will be decided by the methods outlined in the following example. *Subject matter expert A* independently estimates that, of the hypothetical 100 minimally competent individuals, 50 would answer item one correctly, 20 item two, 70 item three and so on for all 12 items. The MPL for *subject matter expert JM* (MPL_{JM}) = $(0.5 + 0.2 + 0.7 + \cdots x_{15})/15 \times 100$ = JM%. Similarly, for *subject matter expert* FM, PC, and JE, the MPLs are FM%, PC%,

JE% respectively. The MPL (cut-off score) for the examination = (JM% + FM% + PC% + JE%). See appendix 1 for the cut-off score calculations.

Box 2: Items for the knowledge domain of the MMIC

Now we would like to ask you some questions about taking aspirin to reduce your chances of getting various conditions. For the following statements, please state whether they are true, false, or unsure. The last question has four possible correct responses.

1. Taking aspirin daily can increase my risk of bleeding
2. Taking aspirin daily can increase my risk of dementia
3. Taking aspirin daily can reduce my risk of heart attacks and strokes
4. Taking aspirin daily can reduce my risk of bowel cancer
5. People who have had angina or a heart attack should consider taking aspirin
6. People who have had a stomach ulcer should consider taking aspirin
7. People who have several close relatives with bowel cancer should consider taking aspirin
8. Healthy people aged 50-70 years should consider taking aspirin
9. Aspirin reduces my chance of bowel cancer if I take it daily for at least a year
10. Aspirin reduces my chance of bowel cancer if I take it daily for at least 2 ½ years.
11. Aspirin doesn't have any effect on my chance of getting bowel cancer
12. The open-ended item is, what are the common side effects of aspirin? Please list as many as you can.

2. Attitude score

The attitude score consists of four items with responses in the format of a seven-point Likert scale. The total score will be dichotomised as either reflecting a positive or negative attitude towards taking aspirin. Participants are asked whether, for them, taking aspirin to reduce their risk of bowel cancer is a: beneficial or harmful, b: important or unimportant, c: a good thing or bad thing, and d: pleasant or unpleasant. The 7-point Likert scale spans across each dichotomous option for each item, e.g. for the first item, 1=very beneficial, 2=quite beneficial, 3=slightly beneficial, 4=neither beneficial nor harmful,5= slightly harmful, 6=quite harmful, or 7=very harmful. See Box 3, for a visual of attitude scale. Each item's response is summed to give a total score, ranging from four to 28, higher scores reflecting more negative attitudes. A positive attitude will be coded if the total score ranges from four to 15, and negative attitudes will be those ranging from 16 to 28. A score of 16 would reflect a neutral attitude and will be coded as a negative attitude for this study.(10)

3. Behaviour

Behaviour is based on the self-reported regular adherence to daily aspirin (i.e., taken five or more out of seven days in a week) at one month. Participants can answer with one of the following three responses (yes, I am currently taking aspirin, I started then stopped taking aspirin, and no, I haven't taken aspirin in the last month). Behaviour will be coded as binary response as either yes or no to whether they decided to take aspirin at one-month. Participants who respond "not taken aspirin in the last month" or "started and then stopped" will be coded as not having adhered to daily aspirin use.(11).

4. Combining the MMIC domains

Table 2 shows how knowledge and attitude scores from the MMIC and the participant's behaviour in taking aspirin or not are coded as informed and uninformed choices, adapted from Marteau et al. (12), Participants' choices to take aspirin

will be categorised into either informed or uninformed according to the following matrix and shown in table 2: if their knowledge, having a positive or negative attitude about aspirin, and behaviour to take aspirin or not, align. All domains are dichotomised for statistical analysis and data are triangulated as Participants can make an informed choice to not take aspirin as well if they have sufficient knowledge and their attitude is negative.

Similar to the behaviour component of the MMIC mentioned above, the **second co-primary outcome** is the difference between the two-study arms in the proportion of participants who self-report regular adherence to daily aspirin (i.e., taken five or more out of seven days in a week) at six months. Participants can answer with one of the following three responses (yes, I am currently taking aspirin, I started then stopped taking aspirin, and no, I haven't taken aspirin in the last month). Participants who respond "not taken aspirin in the last month" or "started and then stopped" will be coded as not having adhered to daily aspirin use.(11)

Secondary outcomes include the difference between the study arms in:

1) Mean decisional conflict was measured using the Decisional Conflict Scale (DCS) (13) at one month. Participants were asked their preference out of four choices to reduce their risk of bowel cancer (change my diet, take aspirin, do the bowel cancer screening test or unsure) and answer the decisional conflict questions in response to their preference. The scale consists of 16 items, with three sub-domains: 1) participants' uncertainty about making a health-related decision; 2) factors that contribute to uncertainty and; 3) participants' perception of how well they came to their final decision.(10) The Decisional Conflict score (range from zero to 100), is calculated as the average of the 16 items scored on a five-point Likert scale (0=strongly disagree, 1=agree, 2=neither, 3=disagree and 4=strongly agree) and multiplied by 25, where 0 indicates no decisional conflict and 100 indicates extremely high decisional conflict. If two or more of the DCS items are left unanswered or are missing, the total will be missing for the participant. The DCS has been widely used in the evaluation of decision aids. (14) The test-retest correlation coefficient was 0.81. (15) Internal consistency was high, with alpha coefficients ranging from 0.78 to 0.92 for the total scale which shows that after administering the DCS twice over a period of time to a group of individuals their scores were similar at each timepoint.

2) Proportion for each of the following additional behaviours to reduce risk of CRC. At one- and six-months participants are asked whether they have done any of the following things to reduce their chances of getting bowel cancer since they joined the study: made changes to their diet, talked to their GP about quitting smoking, quit smoking, discussed with their GP screening for CRC by faecal occult blood test (FOBT) or colonoscopy, completed screening for CRC by FOBT or colonoscopy or, talked to their GP about taking aspirin. The response to each of the items will be coded as 1=Yes and 0=No and missing if they do not provide a response.

3) Proportion of participants who self-reported regular adherence to daily aspirin (i.e., taken five or more out of seven days in a week) at one month using the same measure as for the primary outcome at six-months. (Described above)

4) Proportion of participants who had a consultation with their general practitioner between baseline and six months. The information will be collected by researcher SM who is blinded to participant allocation from an audit of general practitioner medical records for each participant enrolled in the trial. The potential degree of contamination between the study arms will assessed by measuring general practitioner discussions about aspirin.

Other descriptive measures

1) Participants who answered “yes” or “started then stopped taking aspirin” to the questions about aspirin adherence were asked additional information about the dose of aspirin they were taking (100 mg/300 mg/other); their reasons for taking aspirin (reduce risk of heart attack, reduce risk of stroke, reduce my risk of bowel cancer), or other reasons for taking aspirin which they could give with an open-ended response. At six-months participants were asked the reasons why they did not take aspirin or why they stopped taking aspirin.

2) At six-months participants are asked whether they experienced any of the following side-effects from taking aspirin and could select one or more of the following: nausea, easy bruising, indigestion, bleeding, or any others. If participants select other, they typed or wrote other side-effects they experienced.

Data collection and management

Data management and workflow

Data will be prepared for analysis by the data analyst at the end of the six-month follow up period on 23rd November 2021. The data analyst, blinded to trial arm allocation, will export a de-identified CSV file from REDCap and then import it into Stata 17, (16), for data processing and statistical analysis. The senior trial statistician will ensure the data analyst remains blinded to the study arms by recoding and removing labels of the randomisation variable in the dataset. Data management tasks include checking that the values are within range and dichotomised or categorised when appropriate, renaming variables, re-labelling variables, creating composite variables when appropriate, and deleting any unnecessary variables. If any errors are found in the data, these will be corrected to as a part of the data cleaning process. Throughout the data processing and analysis, the trial statisticians will work closely with the data analyst to cross check the data, coding and analysis methods used. The data cleaning STATA 17 do file can be found in appendix 2. De-identified data will be stored on the University server for future use in accordance with the University of Melbourne’s Research Integrity and Misconduct Policy (MPF1318).(17)

Plans for assessment and collection of outcomes

The research assistant will capture in the REDCap database the number of participants who were ineligible, reasons for not meeting eligibility criteria, as well their age and sex. Participants complete a baseline questionnaire which is administered by a research assistant (LB or NK) prior to randomisation and entered directly into the REDCap trial database, in a private consultation room or via Zoom. One and six-month follow up questionnaires for the patient-reported outcome measures will be sent to each participant and completed by either text, email, over the phone by an RA who is not involved in recruitment or by receiving a paper copy in the post depending on their stated preference at baseline. Participants who opted to receive follow-up questionnaires by text or email will receive two automated text or email reminders to complete the questionnaires after three and six days and then a phone call reminder by a blinded RA after nine days. Participants who opted to receive follow up questionnaires by post will be reminded by phone to return them ten days after they are posted. If the participant does not have a phone number, we will repost the questionnaire with a reminder note attached, two after they are posted.

Statistical methods

Statistical analyses will be conducted at the end of the six-month data collection period and will commence after the Statistical Analysis Plan has been uploaded to the Australian and New Zealand Clinical Trials Registry. Blinded analysis will be start mid-May 2022. The preliminary results will be presented at an investigator meeting early July for a blinded review and interpretation. A draft report of the findings two co-primary outcomes will be submitted to the Victorian Cancer Agency end of July 2022. The results may be submitted to a peer-reviewed journal article for publication and presented at conferences both nationally and internationally, pending approval from the funding body, the Victorian Cancer Agency. All analyses will be conducted using Stata 17.(16)

Descriptive analysis

A flow chart will be created to show the flow of participants from screening to six-months of follow up see figure 1 for the template. The flowchart will show the number of participants approached in the general practices, the number who declined participation in the trial, the recruitment rate including the number of people screened, the attrition rates and the number of participants randomised into each study arm and follow up rates at both one and six-months.

Data collected at screening will be used to describe the number of participants who were ineligible, the reasons why they were not eligible for the trial, their age and sex.

Attrition rates and number of participants who completed the follow up questionnaires at one and six-months, by study arm, will also be reported. When such information is available, the reasons participants withdrew or lost to follow up will be reported by study arm.

Descriptive statistics will be used to compare baseline participant demographic characteristics between the two study arms and will be presented as frequencies and percentages shown in Table 1. These include participant demographic characteristics overall and by study arm, including their gender, socio-economic status, whether they were born in Australia or overseas, the aggregated number of medications they were taking, highest level of education attained, whether or not they were living alone, self-reported health measures for them and their family history of bowel cancer. Except for the subjective numeracy scale and age in years which will be presented as means with their standard deviation.

Counts and percentages for informed choices across all combinations of the three MMIC domains for all participants and by study arm will be presented, see Table 2.

To describe the missing data for the sample and for the co-primary outcomes, they will be summarised and presented as counts and percentages.

Primary analysis

All randomised participants will be included in the primary analysis in their assigned study arms in accordance with the intention-to-treat principle.(18)

Co-primary outcomes

1) The difference in proportions (absolute measure) and odd ratio (relative measure) of participants who are taking regular aspirin at six months between the two study arms will be estimated using a generalised linear model with the identity link function and binomial family (where appropriate) and logistic regression, respectively. Both regression models will be adjusted for GP clinic, brochure type based on sex (male or female) and mode of trial delivery (face-to-face or teletrial) included as covariates.

2) As above, the difference in the proportion and odds ratio of participants who make an informed choice about taking aspirin at one month between the two study arms will be estimated using generalised linear model with the identity link function and binomial family (where appropriate) and logistic regression, respectively. General practice, self-selected male or female decision aid and mode of trial delivery (face-to-face or teletrial) will be included as covariates in the regression models.

For the primary analysis multiple imputation will be used to handle incomplete data for the co-primary outcomes, as the co-primary outcome data are collected at one and six-months some responses in the questionnaires may be incomplete. We will impute 50 datasets for the co-primary outcomes using chained equations to generate imputed data. Datasets will be imputed at either the component or by each scale or measure, or at the composite level depending on the patterns of missing data, if they are missing at random or missing completely at random. In addition to the co-primary outcomes measured at other time points, the imputation model will include selected baseline variables (study arm status, age, face-to-face versus teletrial, brochure type (male/female), cardiovascular risk, family history of bowel cancer, number of medications, and Socio-Economic Indexes for Areas (SEIFA) based on participants' postcode of residence. The multiple imputation model will also include the secondary outcomes included in Table 3. Estimands of interest (that is, mean differences, odds ratios) and their standard errors will be combined using the methods originally outlined by Rubin (19).

Estimates of the between-arm difference in proportions and odds ratios for the co-primary outcomes will be reported with Bonferroni adjusted 95% confidence intervals, and the p-value estimated using the logistic model (see Table 3), together with the counts and percentages for each outcome by study arm.

Secondary outcomes

For the secondary binary outcomes presented in Table 3 and 4, we will use logistic regression to estimate the odd ratio, and (if appropriate) use generalised linear model with the identity link function and binomial family to estimate the between-arm difference in proportions for these outcomes. For the outcomes in Table 4, which are measured at two points (1 month and 6 months), we will use generalised estimating equation with robust standard errors to allow for the correlation of repeated outcomes on the same individual. The between-arm difference in means for the decisional conflict scale will be estimated using linear regression. All regression analyses will be adjusted for the randomisation stratification factors including general practice, sex, and mode of trial delivery (face-to-face or teletrial). The estimated intervention effect will be reported as the odds ratio between-arm difference in proportions for binary outcomes and the difference in means between the intervention and control arms for continuous outcomes. Missing values and incomplete data will be imputed as described above.

Estimates for secondary outcomes will be reported with respective 95% confidence intervals and p values with no adjustments for multiplicity(20).

Sensitivity analysis
We will perform two sensitivity analyses for the co-primary outcomes to assess the robustness of our results from the primary analysis and to account for missing data.

1. The first sensitivity analysis includes specifying a different method for the imputation model than what was used in the primary analysis. If we choose to impute data at the composite level, the data will be imputed at the component level for the sensitivity analysis, and vice versa.
2. The second sensitivity analysis will be conducted as a complete case analysis for the co-primary outcomes which would be done following the same methods outlined for the primary analysis above, but the cases with missing data will be excluded.

A sensitivity analysis will be performed on the primary and secondary outcomes to adjust for additional pre-specified baseline variables in the regression models. These include age in years, sex and family history of colorectal cancer, cardiovascular disease risk and subjective numeracy scores.

For the co-primary outcomes, the proportion of participants who are taking regular aspirin at six months and the proportion who have made an informed decision about taking aspirin at one-month, we will use a pattern mixture model to assess the robustness of the missing data assumption. The analysis to assess robustness of missing data assumption may be repeated, as appropriate, for the secondary outcomes.

Sub-group analysis
Exploratory sub-group analyses are planned to identify differences in intervention effects on the co-primary outcomes for participants by face-to-face versus teletrial, brochure type (male/female), cardiovascular risk, family history of bowel cancer, number of medications, and Socio-Economic Indexes for Areas (SEIFA) based on participants' postcode of residence and education (4). The estimated effects will be reported as the odds ratio for binary outcomes and the difference in means between the intervention and control arms for continuous outcomes. No corrections will be made for multiple testing.

Interim analysis
We do not plan to conduct an interim analysis for this trial.

Adherence adjusted analysis

In the protocol it was specified that an adherence adjusted analysis would be conducted for the two co-primary outcomes using a complier average casual effect (CACE) analysis. This analysis will no longer be required as there was no non-compliance as everyone received the intervention as intended. Participants were unable to discontinue the intervention as it was provided immediately post randomisation and for this study, we did not give them aspirin to take.

Tables and Figures

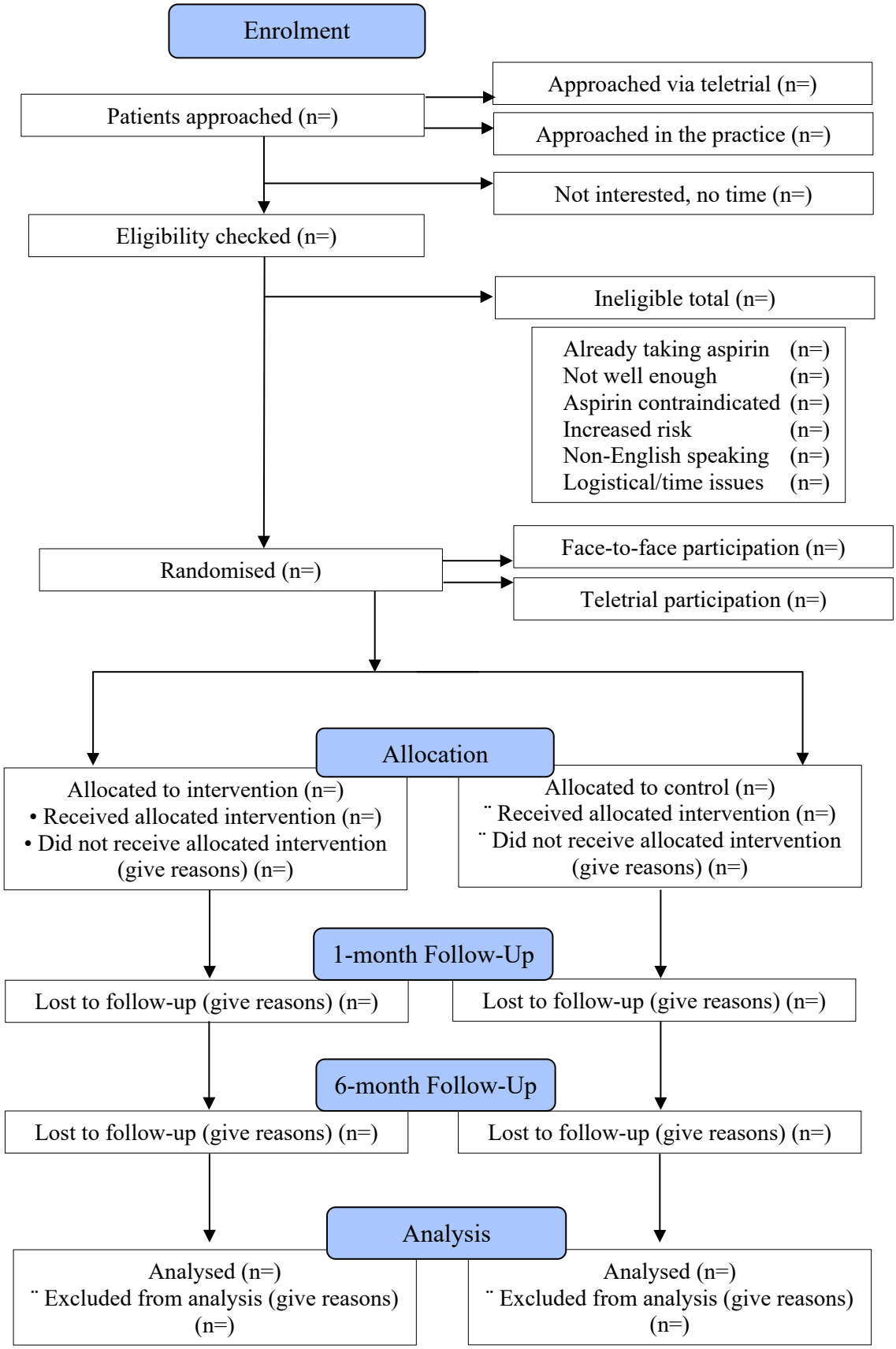


Figure 1. Baseline characteristics of participants according to study arm, in total and stratified by intervention and control arms.

Box 3. Snapshot from SITA trial participant 1-month follow up questionnaire attitude questions which is a part of the multi-dimensional measure of informed choice

For me, taking aspirin to reduce my bowel cancer risk is:

		VERY	QUITE	SLIGHTLY	NEITHER	SLIGHTLY	QUITE	VERY	
a)	Beneficial	1	2	3	4	5	6	7	Harmful
b)	Important	1	2	3	4	5	6	7	Unimportant
c)	Good thing	1	2	3	4	5	6	7	Bad thing
d)	Pleasant	1	2	3	4	5	6	7	Unpleasant

Table 1. Descriptive statistics of baseline characteristics for all participants and by study arm.

	All participants	Intervention	Control
Age (years), mean (SD)	mean (SD)	mean (SD)	mean (SD)
Gender			
Male	n (%)	n (%)	n (%)
Female	n (%)	n (%)	n (%)
Variations of sex characteristics	n (%)	n (%)	n (%)
*IRSAD Socio-Economic status			
Disadvantaged 1	n (%)	n (%)	n (%)
2	n (%)	n (%)	n (%)
3	n (%)	n (%)	n (%)
4	n (%)	n (%)	n (%)
Advantaged 5	n (%)	n (%)	n (%)
Country of birth			
Australia	n (%)	n (%)	n (%)
Overseas	n (%)	n (%)	n (%)
Current medications			
None	n (%)	n (%)	n (%)
One	n (%)	n (%)	n (%)
Two to three	n (%)	n (%)	n (%)
More than five	n (%)	n (%)	n (%)
Education			
Never completed high school	n (%)	n (%)	n (%)
Completed high school only	n (%)	n (%)	n (%)
TAFE qualification or similar	n (%)	n (%)	n (%)
University degree or higher	n (%)	n (%)	n (%)
Living alone			
Yes	n (%)	n (%)	n (%)
Languages spoken at home			
English	n (%)	n (%)	n (%)
Other	n (%)	n (%)	n (%)
Subjective numeracy score	mean (SD)	mean (SD)	mean (SD)
Cardiovascular disease risk			
Family history of heart attack or stroke			
Yes	n (%)	n (%)	n (%)
No	n (%)	n (%)	n (%)
Unsure	n (%)	n (%)	n (%)
Personal history of diabetes			
Yes	n (%)	n (%)	n (%)
Taking medication for high blood pressure			
Yes	n (%)	n (%)	n (%)
Personal history of high cholesterol			
Yes	n (%)	n (%)	n (%)
Bowel cancer risk			
Family history of bowel cancer			
Yes	n (%)	n (%)	n (%)

Notes: SD = Standard deviation, **s**ub-categories may be collapsed in final table published. *The Index of Relative Socio-economic Advantage and Disadvantage (IRSAD)

Table 2. Number and percentage of participants who had sufficient knowledge, a positive attitude and whether they decided to take aspirin or not, which are the three domains of the MMIC, in the SITA trial.

	Sufficient knowledge	Positive attitude	Taking aspirin	Overall	Intervention	Control
All possible informed choices						
1	✓	✓	✓	n (%)	n (%)	n (%)
2	✓	✗	✗	n (%)	n (%)	n (%)
All possible uninformed choices						
3	✓	✗	✓	n (%)	n (%)	n (%)
4	✓	✓	✗	n (%)	n (%)	n (%)
5	✗	✓	✓	n (%)	n (%)	n (%)
6	✗	✗	✓	n (%)	n (%)	n (%)
7	✗	✓	✗	n (%)	n (%)	n (%)
8	✗	✗	✗	n (%)	n (%)	n (%)

Notes: Difference in proportions between the two arms

Table 3. Co-primary outcomes and secondary outcomes by study arm for the SITA trial.

	Intervention	Control	Estimated effect size		
Number of participants	n	n			
Co-primary, self-reported daily aspirin at 1-month[1]	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[2]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[3]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[4]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Co-primary, informed choice about taking aspirin at 1-month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[2]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[3]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Sensitivity analysis[4]			Difference (95% CI)	Odds ratio (95% CI)	p-value
Secondary Outcomes					
Decisional conflict scale[1]	Mean (SD)	Mean (SD)	Mean Difference (95% CI)	Mean Difference (95% CI)	p-value
Self-reported daily aspirin at 6-months[1]	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
GP record audit, spoke to GP about taking aspirin[1]	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value

Notes: Difference – Difference in percentages between the arms; SD = Standard deviation; CI = Confidence interval.

1 Estimated using multiple imputation
2 Sensitivity analysis using alternative Missing Imputation model
3 Sensitivity analysis with complete cases only
4 Sensitivity analysis adjusted for age, family history of bowel cancer, subjective numeracy scores

Table 4. Participant self-reported changed behaviours at 1-month and 6-months by study arm in the SITA trial.

	Intervention	Control	Estimated effect size		
	n	n			
Behaviours to reduce bowel cancer risk					
Changes to their diet					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Spoke to GP about quitting smoking					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Quit smoking					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Spoke to GP about screening for bowel cancer by FOBT					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Completed FOBT test					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Spoke to GP about screening for bowel cancer by colonoscopy					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Had a colonoscopy					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
Spoke to GP about taking aspirin					
1 month	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value
6 months	n (%)	n (%)	Difference (95% CI)	Odds ratio (95% CI)	p-value

Acknowledgements

Authors' contributions

JE conceived of the study and developed the initial trial design. SM, JM, FM, PC, LT, MJ, FMW, NT, LB, SS, NK, KN, CF, JMG, KB, SW, SM, GF, JM, MS contributed to the study design. JE, MJ, LT, FW, FM, JM, SS, PC, and SM are the grant holders. PC and RW provided statistical expertise for the statistical analysis of the SAP. All authors contributed to refinement of the study protocol and approved the final manuscript

Additional information

Funding

This study is funded by a dedicated grant from the Victorian Cancer Agency ID: CPSRG19011. The University of Melbourne is the sponsor of the present trial. The sponsor had no role in the study design and will not have role or authority during the collection, management, analysis, and interpretation of data, and in the decision to submit the results for publication.

Ethics approval and consent to participate

We have obtained ethics approval through the University of Melbourne's Medicine and Dentistry Human Ethics Sub-Committee 2056513.

References

1. Macrae F, Trevor L, Clarke J, Emery J, Jenkins M, McNeil J, et al. Dietary and lifestyle strategies - Clinical Guidelines Wiki [Internet]. 2017 [cited 2020 Dec 9]. Available from: https://wiki.cancer.org.au/australia/Guidelines:Colorectal_cancer/Primary_prevention_dietary_and_lifestyle
2. Milton S, McIntosh J, Macrae F, Chondros P, Trevena L, Jenkins M, et al. An RCT of a decision aid to support informed choices about taking aspirin to prevent colorectal cancer and other chronic diseases: a study protocol for the SITA (Should I Take Aspirin?) trial. Trials 2021 221 [Internet]. 2021 Jul 15 [cited 2021 Aug 11];22(1):1–17. Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-021-05365-8
3. 2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia, 2016 [Internet]. [cited 2022 Mar 21]. Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/2033.0.55.001~2016~Main Features~SEIFA Basics~5
4. Australian Bureau of Statistics. Main Features - SOCIO-ECONOMIC INDEXES FOR AREAS (SEIFA) 2016. 2016 [cited 2021 Feb 22]; Available from: https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2033.0.55.001Main+Features12016?OpenDocument
5. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making [Internet]. 2007 Sep 14 [cited 2020 Mar 26];27(5):663–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17652180
6. Fagerlin A, Zikmund-Fisher BJ, Ubel P a, Jankovic A, Derry H a, Smith DM. Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A, Jankovic, A., Derry, H. A, & Smith, D. M. (2009). Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making [Internet]. 2009 Sep 14 [cited 2020 Mar 26];27(5):672–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17641137
7. Walker JG, Macrae F, Winship I, Oberoi J, Saya S, Milton S, et al. The use of a risk assessment and decision support tool (CRISP) compared with usual care in general practice to increase risk-stratified colorectal cancer screening: Study protocol for a randomised controlled trial. Trials. 2018;19(1).
8. Bekker H, Modell M, Denniss G, Silver A, Mathew C, Bobrow M, et al. Uptake of cystic fibrosis testing in primary care: supply push or demand pull? BMJ Br Med J [Internet]. 1993 [cited 2022 Mar 22];306(6892):1584. Available from: /pmc/articles/PMC1678014/?report=abstract
9. Bandaranayake RC. Setting and maintaining standards in multiple choice examinations: AMEE Guide No. 37. Med Teach [Internet]. 2008 [cited 2022 Feb 10];30(9–10):836–45. Available from: https://pubmed.ncbi.nlm.nih.gov/19117221/
10. Michie S, Dormandy E, Marteau TM. The multi-dimensional measure of informed choice: A validation study. Patient Educ Couns. 2002 Sep;48(1):87–91.

11. Sheridan SL, Draeger LB, Pignone MP, Keyserling TC, Simpson RJ, Rimer B, et al. A randomized trial of an intervention to improve use and adherence to effective coronary heart disease prevention strategies. Vol. 11, BMC Health Services Research. BMC Health Serv Res; 2011.
12. Marteau TM, Dormandy E, Michie S. A measure of informed choice. Health Expect [Internet]. 2001 [cited 2022 Mar 22];4(2):99–108. Available from: https://pubmed.ncbi.nlm.nih.gov/11359540/
13. Pozzar RA, Berry DL, Hong F. Item response theory analysis and properties of decisional conflict scales: findings from two multi-site trials of men with localized prostate cancer. BMC Med Inform Decis Mak [Internet]. 2019 Dec 4 [cited 2020 Mar 31];19(1):124. Available from: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-019-0853-5
14. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Vol. 2017, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2017.
15. O'connor AM. Validation of a Decisional Conflict Scale. Med Decis Mak [Internet]. 1995 Feb 2 [cited 2020 Jun 5];15(1):25–30. Available from: http://journals.sagepub.com/doi/10.1177/0272989X9501500105
16. StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC [Internet]. [cited 2022 Apr 1]. Available from: https://www.stata.com/stata17/
17. Research Integrity and Misconduct Policy (MPF1318) : Policy : The University of Melbourne [Internet]. [cited 2022 Apr 1]. Available from: https://policy.unimelb.edu.au/MPF1318
18. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: Lessons for intention-to-treat analysis. Clin Trials [Internet]. 2012 Aug [cited 2021 Feb 4];9(4):396–407. Available from: https://pubmed.ncbi.nlm.nih.gov/22752633/
19. Multiple Imputation for Nonresponse in Surveys - Donald B. Rubin - Google Books [Internet]. [cited 2022 Mar 31]. Available from: https://books.google.com.au/books?hl=en&lr=&id=bQBtw6rx_mUC&oi=fnd&pg=PR24&ots=8PrNcP3VgN&sig=SwrXw1b8IimvJE7n9cqYbkxhCgY&redir_esc=y#v=onepage&q&f=false
20. Li G, Taljaard M, Van Den Heuvel ER, Levine MA, Cook DJ, Wells GA, et al. An introduction to multiplicity issues in clinical trials: the what, why, when and how. Int J Epidemiol [Internet]. 2017 [cited 2021 Apr 1];46(2):746–55. Available from: https://academic.oup.com/ije/article/46/2/746/2741997

Appendices

Appendix 1. The Angoff Method

Angoff Method for knowledge items of MMIC (SITA trial)
Each of four judges considers 100 minimally competent individuals taking an examination of 11 items.
Can a person with minimal competence answer the item correctly?
Now we would like to ask you some questions about taking aspirin to reduce your chances of getting various conditions. For the following statements, please state whether they are true or false.

					JE	JM	FM	PC
1	Taking aspirin daily can increase my risk of bleeding	☐ True	☐ False	☐ Unsure	60	60	70	70
2	Taking aspirin daily can increase my risk of dementia	☐ True	☐ False	☐ Unsure	30	20	20	30
3	Taking aspirin daily can reduce my risk of heart attacks and strokes	☐ True	☐ False	☐ Unsure	70	70	80	70
4	Taking aspirin daily can reduce my risk of bowel cancer	☐ True	☐ False	☐ Unsure	70	80	70	60
5	People who have had angina or a heart attack should consider taking aspirin	☐ True	☐ False	☐ Unsure	50	65	80	70
6	People who have had a stomach ulcer should consider taking aspirin	☐ True	☐ False	☐ Unsure	50	65	80	60
7	People who have several close relatives with bowel cancer should consider taking aspirin	☐ True	☐ False	☐ Unsure	60	60	50	50
8	Healthy people aged 50-70 years should consider taking aspirin	☐ True	☐ False	☐ Unsure	60	80	60	50
9	Aspirin reduces my chance of bowel cancer if I take it daily for at least a year	☐ True	☐ False	☐ Unsure	65	50	70	50
10	Aspirin reduces my chance of bowel cancer if I take it daily for at least 2 ½ years.	☐ True	☐ False	☐ Unsure	40	50	40	40
11	Aspirin doesn't have any effect on my chance of getting bowel cancer	☐ True	☐ False	☐ Unsure	70	50	80	60
12	What are the common side effects of aspirin? Please list as many as you can: **Nausea, indigestion, easy bruising, bleeding**				50 50 50 50	65 65 65 65	50 50 50 50	50 50 50 50
	$(MPL_A) = (0.5 + 0.2 + 0.7 + \cdots xn)/n \times 100 = A\%$.				55	60.67	60	54
	$MPL_{JE} = (0.6 + 0.5 + 0.7 + 0.7 + 0.5 + 0.5 + 0.6 + 0.6 + 0.65 + 0.4 + 0.7 + 2)/15 \times 100 =$				(56.25% + 59.58% + 62.5% + 55%)/4			
	The MPL (cut-off score) for the examination = (A% + B% + C% + D% + E% + ⋯ N%)/N				57.41			
	Cut-off				15*0.5741 = 8.6			

```
**APPENDIX 2. STATA 17 - Data Cleaning Do-file codebook for the Should I Take Aspirin? SITA trial

//Opens the correct data file //
use "\\research-cifs.unimelb.edu.au\5850-Research\CTU\CANCER_SITA\Dataclean\SITA data in STATA\SITA Data 239 variables_includes non-participants.dta"

/*use command pwd tells me STATAs current working directory or where it searches in my files to open the dataset*/
/* change working directory in file menu*/

summarize

/* save data as efficiently as possible */
compress

/*drop uncecessary variables*/
drop mo_gp_data_extraction_complete datetime send_consent confirm_name ///
baseline_data_timestamp month_fup_study_admin_complete ///
month_follow_up_questionnaire_65 consent_pref middle_name middle_name ///
date_consent econsent_timestamp app_date app_day app_time paper_consent ///
econsent_complete randomisation_timestamp notes_1mo_ph_1 month_follow_up_questionnaire_ti ///
recording_consent_2 recording_consent crcn_consent folup_preference post_resi ///


/*drop participants who weren't randomised*/
drop if randomisation_complete==0

/* rename and variables syntax: rename old_varname new_varname*/
rename which_effect_6___0 se_nausea
rename which_effect_6___1 se_bruising
rename which_effect_6___2 se_indigestion
rename which_effect_6___3 se_bleeding
rename which_effect_6___4 se_other

rename exp_side_effects___0 exp_se_nausea
rename exp_side_effects___1 exp_se_bruising
rename exp_side_effects___2 exp_se_indigestion
rename exp_side_effects___3 exp_se_bleeding
rename exp_side_effects___4 exp_se_other

rename symp_exp_gpn___0 se_nausea_gpn
rename symp_exp_gpn___1 se_bruising_gpn
rename symp_exp_gpn___2 se_indigestion_gpn
rename symp_exp_gpn___3 se_bleeding_gpn
rename symp_exp_gpn___4 se_other_gpn

rename ppi_type___1 ppi_losec
rename ppi_type___2 ppi_nexium
rename ppi_type___3 ppi_pariet
rename ppi_type___4 ppi_somac
rename ppi_type___5 ppi_zoton_fastabs
```

```
50
51  /* re-label and variables lab var varname "label" */
52  lab var se_nausea "Nausea side-effect"
53  lab var se_bruising "Bruising side-effect"
54  lab var se_indigestion "Indegestion side-effect"
55  lab var se_bleeding "Bleeding side-effect"
56  lab var se_other "Other side-effect"
57
58  lab var exp_se_nausea "Nausea side-effect experienced"
59  lab var exp_se_bruising "Bruising side-effect experienced"
60  lab var exp_se_indigestion "Indegestion side-effect experienced"
61  lab var exp_se_bleeding "Bleeding side-effect experienced"
62  lab var exp_se_other "Other side-effect experienced"
63
64  lab var se_nausea_gpn "Nausea side-effect GP Notes"
65  lab var se_bruising_gpn "Bruising side-effect GP Notes"
66  lab var se_indigestion_gpn "Indegestion side-effect GP Notes"
67  lab var se_bleeding_gpn "Bleeding side-effect GP Notes"
68  lab var se_other_gpn "Other side-effect GP Notes"
69
70  lab var ppis "Taking PPI yes or now"
71  lab var ppi_losec "Lorsec PPI"
72  lab var ppi_nexium "Nexium PPI"
73  lab var ppi_pariet "Pariet PPI"
74  lab var ppi_somac "Somac PPI"
75  lab var ppi_zoton_fastabs "Zonton FastTabs PPI"
76
77  lab var randomisation_complete "0 incomplete 1 unverified 2 complete"
78
79  *BLINDED
80  label var randomise "Decision Aid or CRC borchure"
81  recode randomise 1=0 2=1
82  lab define randomise 0 Group_A 1 Group_B, replace
83  tab randomise
84
85  ********************************************************************************
    **************
86  /* postodes SEIFA continuous and dichotomised into advantaged and disadvantaged*/
87  ********************************************************************************
    **************
88  *change variable type to integer
89  rename study_id study_id_alph
90  destring study_id_alph, generate(study_id)
91  drop study_id_alph
92
93  *merge the SEIFA dataset with this one
94  merge 1:1 study_id using
    "\\research-cifs.unimelb.edu.au\5850-Research\CTU\CANCER_SITA\Dataclean\SITA data
    in STATA\SEIFA and postcodes.dta"
95
96  *change variable type to integer
97  recast int seifa
98
```

```
*close seifa so for the continuous addition to the table1 below
clonevar seifa_c = seifa
label variable seifa_c "Socio-economic status (SEIFA)"

*seifa from 10 to five quintiles
recode seifa (1 2 = 2)
recode seifa (3 4 = 4)
recode seifa (5 6 = 6)
recode seifa (7 8 = 8)
recode seifa (9 10 = 10)
lab var seifa "SEIFA IRSAD Quintiles"
label define seifa 2 "1 Disadvantaged" 4 "2" 6 "3" 8 "4" 10 "5 Advantaged", replace
label values seifa seifa
label list seifa
tab seifa

*******************************************************************************
*******************************
/*analyse Subjective Numeracy Scale, 6-point likert scales reported as odds ratios
with 95% ci & p-values
*******************************************************************************
*******************************
*Subjective numeracy scale
Response values increase left to right (1-6).
1= not at all good/ 6= extremely good
 Scoring is based on these values, except Question 7 is reverse coded (6-1) for
consistency.
SNS: Average rating across all 8 questions (w/ Q7 reverse coded) */

*re-order predictions_sns as it is reverse coded
tab predictions_sns
recode predictions_sns (0=5 "1. Always prefer percentages")(1=4 "2.")(2=3 "3.")(3=2
 "4.")(4=1 "5.")(5=0 "6. Always prefer words"), generate(weather_sns_r) label(
weather_sns_r) test
codebook weather_sns_r

*recode all sns variables from 0-5 to 1-6
recode fractions_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
label define fractions_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6
"6. Extremely good", replace
label values fractions_sns fractions_sns
codebook fractions_sns

recode percentage_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
label define percentage_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6
"6. Extremely good", replace
label values percentage_sns percentage_sns
codebook percentage_sns

recode tip_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
label define tip_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6 "6.
Extremely good", replace
label values tip_sns tip_sns
```

```
143  codebook tip_sns
144
145  recode shirt_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
146  label define shirt_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6 "6.
     Extremely good", replace
147  label values shirt_sns shirt_sns
148  codebook shirt_sns
149
150  recode news_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
151  label define news_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6 "6.
     Extremely good", replace
152  label values news_sns news_sns
153  codebook news_sns
154
155  recode words_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
156  label define words_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6 "6.
     Extremely good", replace
157  label values words_sns words_sns
158  codebook words_sns
159
160  recode weather_sns_r (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
161  label define weather_sns_r 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6
     "6. Extremely good", replace
162  label values weather_sns_r weather_sns_r
163  codebook weather_sns_r
164
165  recode use_sns (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)
166  label define use_sns 1 "1. Not at all good" 2 "2." 3 "3." 4 "4." 5 "5." 6 "6.
     Extremely good", replace
167  label values use_sns use_sns
168  codebook use_sns
169
170  *gen newvariableneame = (fractions_sns + percentage_sns + tip_sns shirt_sns +
     news_sns + words_sns + weather_sns_r + use_sns)/8
171  gen sns = (fractions_sns + percentage_sns + tip_sns + shirt_sns + news_sns +
     words_sns + weather_sns_r + use_sns)/8
172  lab var sns "Subjective Numeracy Scale"
173  codebook sns
174  histogram sns //see the distribution
175
176  ********************************************************************************
     **************
177  /*AUTOMATED TABLE 1- BASELINE CHARACTERISTICS*/
178  ********************************************************************************
     **************
179  *Step 1 install the program *********
180  ssc install table1_mc
181
182  *Step 2 cob postcode oth_med #meds from string variables to numeric ********
183  destring postcode, generate(postcode_n)
184
185  *oth_med: contains nonnumeric characters, so had to make following edits to remove
     words
```

```
replace oth_med = "5" in 42
replace oth_med = "30" in 71
replace oth_med = "8" in 86
replace oth_med = "10" in 9
replace oth_med = "1" in 257
replace oth_med = "4" in 83
replace oth_med = "5" in 73
replace oth_med = "10" in 95
destring oth_med, generate(no_meds)
list no_meds

*Step 3 label the variables************/
label list fem_male_broch
tab fem_male_broch
label variable fem_male_broch "Male or Female Decision Aid"
tab fem_male_broch
revrs fem_male_broch //reverse code this variable so its the same as sex in the
table
tab revfem_male_broch

*dichotomizing language, then changing labels
label list language
recode language (2 3 4 5 6 7 8 9 10 11 12 13 = 1)
label list language
label define language 1 "Other" 0 "English", replace
tab language
label variable language "Language spoken at home"
tab language

* collapse countries, dichotomize, then change labels
encode cob, generate(cob_n)
label list cob_n
recode cob_n (1 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 ///
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45  = 1)
recode cob_n (2 46 = 3)
label define cob_n 1 "Born overseas" 3 "Born in Australia", replace
tab cob_n
tab cob_n, nol

label list living
label variable living "Living alone"
tab living
tab living, nol

label list education
tab education
label variable education "Education"
tab education
tab education, nol

tab no_meds, miss
label variable no_meds "Number of tablets taking, excluding vitamins"
recode no_meds .=0  1/1.5 = 1 2/3 = 2 4/5 = 3 5.5/max = 4, generate(med_cat)
```

```
tab med_cat
label define med_cat 0 "None" 1 "One" 2 "Two to three" 3 "Four to five" 4 "More than five", replace
label values med_cat med_cat
label variable med_cat "Number of tablets taking, excluding vitamins"
tab med_cat

codebook heart_attack
label variable heart_attack "Family history of heart attack or stroke"
tab heart_attack

codebook cholesterol
label variable cholesterol "Personal history of high cholesterol"
tab cholesterol

codebook blood_pressure
label variable blood_pressure "Taking medication for high blood pressure"
tab blood_pressure

codebook diab
label variable diab "Personal history of diabetes"
tab diab

codebook fdr
label variable fdr "Family history of bowel cancer"
tab fdr

codebook cig
label variable cig "Current or history of smoking cigarettes"
tab cig

*Step 4 read details of generating the baseline table1_mc
// example code: table1_mc, by(foreign) vars(price conts \ price contln %5.0f %4.2f \ weight contn %5.0f \ rep78 cate \ much_headroom bine)

table1_mc, by(randomise) vars( enrolage contn %5.1f \ sex_recruit cat %5.0f \ revfem_male_broch cat %5.0f \ ///
education cat %5.0f \ language cat 5.0f \ cob_n cat %5.0f \ seifa_c contn %5.1f \ seifa cat %5.0f ///
living bin %5.0f \ sns contn %5.1f \ ppis cat %5.0f \ med_cat cat %5.0f \ heart_attack cat %5.0f \ cholesterol cat %5.0f  ///
\ blood_pressure cat %5.0f \ cig cat %5.0f \ fdr cat %5.0f ) nospace percent_n onecol total(before) ///
saving ( "\\research-cifs.unimelb.edu.au\5850-Research\CTU\CANCER_SITA\Dataclean\SITA data in STATA\table 1.xls", replace)

*****************************************************************************************************
*First co-primary outcome Multi-dimensional measure of informed choice - MMIC
*****************************************************************************************************
* co-primary MMIC knowledge - 12 true/false/unsure 1-open ended - dichotomise
```

```
281  codebook aspirin_sidee bleed_risk dementia_risk heart_stroke_risk crc_risk
     heart_consider ulcer_consider fam_consider health_consider year_crc
     twohalf_year_crc no_chance_crc
282
283  rename bleed_risk knowledge1
284  rename dementia_risk knowledge2
285  rename heart_stroke_risk knowledge3
286  rename crc_risk knowledge4
287  rename heart_consider knowledge5
288  rename ulcer_consider knowledge6
289  rename fam_consider knowledge7
290  rename health_consider knowledge8
291  rename twohalf_year_crc knowledge9
292  rename year_crc knowledge10
293  rename no_chance_crc knowledge11
294  rename aspirin_sidee knowledge12
295
296  codebook knowledge1 knowledge2 knowledge3 knowledge4 knowledge5 knowledge6
     knowledge7 knowledge8 knowledge10 knowledge9 knowledge11
297
298  *knowledge12 is a string and needs to be recoded in excel- possible open-ended
     answers nausea, indigestion, easy bruising, bleeding
299  br knowledge12
300
301  *merge 1:1 study_id using
     "\\research-cifs.unimelb.edu.au\5850-Research\CTU\CANCER_SITA\Dataclean\SITA data
     in STATA\SEIFA and postcodes.dta"
302
303  rename _merge _merge2
304  merge 1:1 study_id using
     "\\research-cifs.unimelb.edu.au\5850-Research\CTU\CANCER_SITA\Dataclean\SITA data
     in STATA\knowledge12.dta"
305  codebook knowledge_nausea knowledge_indegestion knowledge_bruising
     knowledge_bleeding
306
307  rename knowledge12 knowledge12_text
308  generate knowledge12 = (knowledge_nausea + knowledge_indegestion +
     knowledge_bruising + knowledge_bleeding)
309  label variable knowledge12 "What are the common side effects of aspirin? Please
     list as many as you can:"
310
311  recode knowledge12 (0 = 1)(1 = 2)(2 = 3)(3 = 4)
312  recast byte knowledge12
313  label define knowledge12 1 "Bleeding" 2 "Indegestion" 3 "Easy bruising" 4 "Nausea",
      replace
314  lab val knowledge12 knowledge12
315  codebook knowledge12
316
317  *all correct answers for knowledge scale 1=True, 2=False, 3=True, 4=True, 5=true,
     6=False, 7=True, 8=True, 9=False, 10=True, 11=False
318  *recode unsure answers to be incorrect 0=True 1=False 2=Unsure
319  recode knowledge1 (2 = 1) //reverse
320  recode knowledge2 (2 = 0)
```

```
recode knowledge3 (2 = 1) //reverse
recode knowledge4 (2 = 1) //reverse
recode knowledge5 (2 = 1) //reverse
recode knowledge6 (2 = 0)
recode knowledge7 (2 = 1) //reverse
recode knowledge8 (2 = 1) //reverse
recode knowledge9 (2 = 0)
recode knowledge10 (2 = 1) //reverse
recode knowledge11 (2 = 0)

*reverse code knowledge items with true answers (knowledge1 knowledge3 knowledge4
knowledge5 knowledge7 knowledge8 knowledge10) so they get a point if they answer
the item correctly
*0=True 1=False
revrs knowledge1 knowledge3 knowledge4 knowledge5 knowledge7 knowledge8 knowledge10

*edit label values, from 1/2 to 0/1
recode revknowledge1 1=0 2=1
label define revknowledge1 0 "Incorrect" 1 "Correct", replace
lab val revknowledge1 revknowledge1
tab revknowledge1, nol

recode revknowledge3 1=0 2=1
label define revknowledge3 0 "Incorrect" 1 "Correct", replace
lab val revknowledge3 revknowledge3

recode revknowledge4 1=0 2=1
label define revknowledge4 0 "Incorrect" 1 "Correct", replace
lab val revknowledge4 revknowledge4

recode revknowledge5 1=0 2=1
label define revknowledge5 0 "Incorrect" 1 "Correct", replace
lab val revknowledge5 revknowledge5

recode revknowledge7 1=0 2=1
label define revknowledge7 0 "Incorrect" 1 "Correct", replace
lab val revknowledge7 revknowledge7

recode revknowledge8 1=0 2=1
label define revknowledge8 0 "Incorrect" 1 "Correct", replace
lab val revknowledge8 revknowledge8

recode revknowledge10 1=0 2=1
label define revknowledge10 0 "Incorrect" 1 "Correct", replace
lab val revknowledge10 revknowledge10

*edit value labels for other knowledge items that didn't need to be reverse coded
label define knowledge2 0 "Incorrect" 1 "Correct", replace
lab val knowledge2 knowledge2

label define knowledge6 0 "Incorrect" 1 "Correct", replace
lab val knowledge6 knowledge6

```

```
label define knowledge9 0 "Incorrect" 1 "Correct", replace
lab val knowledge9 knowledge9

label define knowledge11 0 "Incorrect" 1 "Correct", replace
lab val knowledge11 knowledge11

codebook revknowledge1 revknowledge3 revknowledge4 revknowledge5 revknowledge7
revknowledge8 revknowledge10 knowledge2 knowledge6 knowledge9 knowledge11
knowledge12

generate knowledge_total_12 = (revknowledge1 + revknowledge3 + revknowledge4 +
revknowledge5 + revknowledge7 + revknowledge8 +  revknowledge10 +  knowledge2 +
knowledge6 + knowledge9 + knowledge11 + knowledge12)

recast byte knowledge_total_12
label variable knowledge_total_12 "MMIC Knowledge Item"
codebook knowledge_total_12

*Cut-off from Angoff Method was 57.41 meaning that participants had to get 57% of
the 15 items or 8.6 items correct to have sufficient knowledge

*if mmic_knowledge >8.60 = "sufficient knowledge"
sum knowledge_total_12 if knowledge_total_12>=8.60, detail

*if mmic_knowledge <8.6 = "insufficient knowledge"
sum knowledge_total_12 if knowledge_total_12<8.6, detail

*generate new variable for knowledge dichotomised
recode knowledge_total_12 0/8.599 = 0 8.6/max = 1, generate(knowledge_dich_12)
label variable knowledge_dich_12 "Knowledge dichotomised"
label define knowledge_dich_12 0 "Insufficient" 1 "Sufficient", replace
lab val knowledge_dich_12 knowledge_dich_12
tab knowledge_dich_12
tab randomise knowledge_dich_12

*a few more knowledge changes
label variable knowledge_total_12 "Total Knowledge"
rename knowledge_total_12 knowledge_total

**********************************************************************************
****************************************
* co-primary MMIC attititude logistic regression, 7-point likert scale* -
dichotomise
**********************************************************************************
****************************************
* 1=very beneficial, 2=quite beneficial, 3=slightly beneficial, 4=neither
beneficial nor harmful,5= slightly harmful, 6=quite harmful, or 7=very harmful

codebook asp_beneficial asp_important asp_bad asp_pleasant

rename asp_beneficial attitude1
rename asp_important attitude2
rename asp_bad attitude3
```

```
416   rename asp_pleasant attitude4
417
418   *recode all attitude variables from 0-6 to 1-7
419   recode attitude1 (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)(6=7)
420   label define attitude1 1 "very beneficial" 2 "quite beneficial" 3 "slightly
      beneficial" ///
421   4 "neither beneficial nor harmful" 5 "slightly harmful" 6 "quite harmful" 7 "very
      harmful", replace
422   label values attitude1 attitude1
423   codebook attitude1
424
425   *recode all attitude variables from 0-6 to 1-7
426   recode attitude2 (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)(6=7)
427   label define attitude2 1 "very important" 2 "quite important" 3 "slightly
      important" ///
428   4 "neither important nor unimportant" 5 "slightly unimportant" 6 "quite
      unimportant" 7 "very unimportant", replace
429   label values attitude2 attitude2
430   codebook attitude2
431
432   *recode all attitude variables from 0-6 to 1-7
433   recode attitude3 (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)(6=7)
434   label define attitude3 1 "very good" 2 "quite good" 3 "slightly good" ///
435   4 "neither good nor bad" 5 "slightly bad" 6 "quite bad" 7 "very bad", replace
436   label values attitude3 attitude3
437   codebook attitude3
438
439   *recode all attitude variables from 0-6 to 1-7
440   recode attitude4 (0=1)(1=2)(2=3)(3=4)(4=5)(5=6)(6=7)
441   label define attitude4 1 "very pleasant" 2 "quite pleasant" 3 "slightly pleasant"
      ///
442   4 "neither pleasant nor unpleasant" 5 "slightly unpleasant" 6 "quite unpleasant" 7
      "very unpleasant", replace
443   label values attitude4 attitude4
444   codebook attitude4
445
446   *Each item's response is summed to give a total score, ranging from four to 28,
      higher scores reflecting more negative attitudes.
447   gen attitude = (attitude1 + attitude2 + attitude3 + attitude4)
448   lab var attitude "Attitude Score MMIC"
449   codebook attitude
450   histogram attitude //see the distrubution
451
452   *generate new variable for attitude dichotomised
453   *A positive attitude 4 to 15, and negative attitudes 16 to 28. A score of 16 would
      reflect a neutral attitude and will be coded as a negative attitude for this study
454   recode attitude 4/15 = 0 16/max = 1, generate(attitude_dich)
455   label variable attitude_dich "Attitude dichotomised"
456   label define attitude_dich 0 "Positive attitude" 1 "Negative attitude", replace
457   lab val attitude_dich attitude_dich
458   tab attitude_dich
459   tab randomise attitude_dich
460
```

```
461  ********************************************************************************
     ****************************************
462  *co-primary self-reported daily adherence to aspirin 5/7 days, logistic
     regression, adjusted for GP clinic
463  *same coding as first co-primary outcome but at 1-month
464  ********************************************************************************
     ****************************************
465  label variable taking_aspirin  "Aspirin uptake 1-mo"
466  codebook taking_aspirin
467  tab taking_aspirin
468  tab randomise taking_aspirin
469
470  ********************************************************************************
     ****************************************
471  *MMIC triagulate the three, knowledge, attitude and uptake into informed or
     uninformed choices
472  ********************************************************************************
     ****************************************
473  rename knowledge_dich_12 knowledge_total_d
474  codebook knowledge_total_d
475  codebook attitude_dich
476  codebook taking_aspirin
477
478  /*combine the no, taking aspirin and started then stopped aspirin:
479  2 participants started then stopped taking aspirin after 1-mo and they were added
     to the taking aspirin group  */
480  recode taking_aspirin (1 = 0)
481  recode taking_aspirin (0 2 = 1)
482
483  label define taking_aspirin 0 "Yes", modify
484  label define taking_aspirin 1 "No and start stopped", modify //only 2 participants
     started then stopped aspirin at 1-mo
485  lab val taking_aspirin taking_aspirin
486
487  recast int knowledge_total_d
488  recast int attitude_dich
489  recast int taking_aspirin
490
491  *generate variable for mmic with all missing values
492  gen mmic=.
493  label variable mmic "Multi-dimensional measure of informed choice"
494  lab define mmic 0 "Uniformed choice" 1 "Informed choice" 2 "Informed choice" 3
     "Uninformed choice" 4 "Uninformed choice" 5 "Uninformed choice" ///
495  6 "Uninformed choice" 7 "Uninformed choice" 8 "Uninformed choice", replace
496  lab values mmic mmic
497  tab mmic randomise, miss
498
499  *all possible combinations of informed choices
500  replace mmic=1 if (knowledge_total_d==1 & attitude_dich==0 & taking_aspirin==0)
501  replace mmic=2 if (knowledge_total_d==1 & attitude_dich==1 & taking_aspirin==1)
502
503  *all possible combinations of UNinformed choices
504  replace mmic=3 if (knowledge_total_d==1 & attitude_dich==1 & taking_aspirin==0)
```

```
505   replace mmic=4 if (knowledge_total_d==1 & attitude_dich==0 & taking_aspirin==1)
506   replace mmic=5 if (knowledge_total_d==0 & attitude_dich==0 & taking_aspirin==0)
507   replace mmic=6 if (knowledge_total_d==0 & attitude_dich==1 & taking_aspirin==0)
508   replace mmic=7 if (knowledge_total_d==0 & attitude_dich==0 & taking_aspirin==1)
509   replace mmic=8 if (knowledge_total_d==0 & attitude_dich==1 & taking_aspirin==1)
510
511   tab mmic
512   codebook mmic
513
514   ********************************************************************************
      ****************************************
515   /*Second co-primary outcome: SELF-REPORT REGULAR ADHERENCE TO DAILY ASPIRIN
516   (i.e., taken 5 or more out of 7 days in a week) at 6 months */
517   ********************************************************************************
      ****************************************
518   codebook taking_aspirin_6
519   label variable taking_aspirin_6  "Aspirin uptake 6-mo"
520   codebook taking_aspirin_6
521   tab taking_aspirin_6
522   tab randomise taking_aspirin_6
523
524   ********************************************************************************
      ****************************************
525   /* Decisional conflict scale - Linear regression to estimate the mean difference
      between the two arms
526   ********************************************************************************
      ****************************************
527   (scores range 0-100)
528   scoring:16-items total [1-16 inclusive]
529   a) sum the 16 items
530   b) divided by 16
531   c) multiplied by 25
532   0 = no decisional conflict
533   100 = extremely high decisional conflict*/
534   codebook reduce_crc_prefer avail_option benefit_option risk_option benefit_me
      risk_me benefit_risk_me choice_support choice_pressure advice_choice best_me_choice
       sure_choice easy_choice informed_choice import_me_choice stick_w_decision
      satified_decision
535
536   *rename the variables to dcs 1-16
537   rename reduce_crc_prefer prefer_dcs
538   tab prefer_dcs
539   rename avail_option dcs1
540   rename benefit_option dcs2
541   rename risk_option dcs3
542   rename benefit_me dcs4
543   rename risk_me dcs5
544   rename benefit_risk_me dcs6
545   rename choice_support dcs7
546   rename choice_pressure dcs8
547   rename advice_choice dcs9
548   rename best_me_choice dcs10
549   rename sure_choice dcs11
```

```
550   rename easy_choice dcs12
551   rename informed_choice dcs13
552   rename import_me_choice dcs14
553   rename stick_w_decision dcs15
554   rename satified_decision dcs16
555
556   codebook dcs1 dcs2 dcs3 dcs4 dcs5 dcs6 dcs7 dcs8 dcs9 dcs10 dcs11 dcs12 dcs13 dcs14
       dcs15 dcs16
557
558   gen dcs = ([dcs1 + dcs2 + dcs3 + dcs4 + dcs5 + dcs6 + dcs7 + dcs8 + dcs9 + dcs10 +
      dcs11 + dcs12 + dcs13 + dcs14 + dcs15 + dcs16] / 16)*25
559   lab var dcs "Decisional conflict score"
560   label val dcs dcs
561   univar dcs
562   histogram dcs //see the distribution
563
```