

1

Combining vitamin C and carotenoid biomarkers better predicts fruit and vegetable intake [than individual biomarkers](#) in a dietary intervention studies

Alanna J McGrath¹, Lesley L Hamill¹, Chris R Cardwell¹, Claire R Draffin¹, Charlotte E Neville¹, Katherine M Appleton², Jane McEneny¹, Michelle C McKinley¹, Ian S Young¹, Jayne V Woodside^{1*}

¹Centre for Public Health, Queen's University Belfast, UK (AJM, LLH, CRC, CRD, CEN, JM, MCM, ISY, JVW)

²Department of Psychology, Bournemouth University, UK (KMA)

Running title: Biomarkers of fruit and vegetable intake

Author list for indexing: McGrath, Hamill, Cardwell, Draffin, Neville, Appleton, McEneny, McKinley, Young, Woodside

Supported by the Medical Research Council as part of grant G0901793. Clinical trial registration number NCT01591057 (www.clinicaltrials.gov)

*corresponding author; Jayne Woodside; Centre for Public Health, First Floor, Institute of Clinical Science B, Grosvenor Road, Belfast, BT12 6BJ; tel 00442890 632585; fax 00442590 235900; j.woodside@qub.ac.uk

Abbreviations used: FV – fruit and vegetables

1

2 *Abstract*

3 *Purpose:* The aim of this study was to determine whether combining potential biomarkers of fruit and
4 vegetables is better at predicting FV intake within FV intervention studies than single biomarkers.

5 *Design:* Data from a tightly controlled randomised FV intervention study (BIOFA V; all food provided
6 and two meals/d on weekdays consumed under supervision) were used. A total of 30 participants
7 were randomised to either 2, 5 or 8 portions FV/d for four weeks, and blood samples were collected at
8 baseline and four weeks for plasma vitamin C and serum carotenoid analysis. The combined
9 biomarker approach was also tested in three further FV intervention studies [conducted by the same](#)
10 [research team](#), with less strict dietary control (FV provided and no supervised meals).

11 *Results:* The combined model containing all carotenoids and vitamin C was a better fit than either the
12 vitamin C only ($P<0.001$) model or the lutein only ($P=0.006$) model in the BIOFAV study. The C-
13 statistic was slightly lower in the lutein only model (0.85) and in the model based upon factor analysis
14 (0.88), and much lower in the vitamin C model (0.68) compared with the full model (0.95). Results for
15 the other studies were similar, although the differences between the models were less marked.

16 *Conclusions:* Although there was some variation between studies, which may relate to the level of
17 dietary control or participant characteristics, a combined biomarker approach to assess overall FV
18 consumption may more accurately predict FV intake within intervention studies than the use of a
19 single biomarker. [The generalisability of these findings to other populations and study designs](#)
20 [remains to be tested.](#)

21 *Keywords:* fruit; vegetables; dietary intake; biomarkers; methodology

22 *Introduction*

23 Increased intake of fruit and vegetables (FV) has consistently been associated with reduced chronic
24 disease risk in observational studies [1], which have been subjected to meta-analysis [2,3]. Such
25 observational evidence has been supplemented by recent randomised controlled trials testing the effect
26 of increased FV intake on clinically relevant endpoints [4-8].

27
28 In both epidemiological research and interventional studies, the accurate measurement of FV
29 consumption is crucial. Traditional self-reported questionnaire-based approaches to the measurement
30 of FV consumption, for example, food frequency questionnaires, 24 hour dietary recalls or food
31 diaries, have well-reported inaccuracies [9-11]. Given such constraints, alternative, objective measures
32 of dietary intake of FV would be valuable. Nutritional biomarkers in biological samples, such as blood
33 and urine, may offer an objective indicator of FV intake [9]. The use of biomarkers would allow a
34 more accurate assessment of the association between FV intake and disease risk, allow population FV
35 intakes to be confirmed, as well as facilitating the measurement of compliance within FV intervention
36 studies [12].

37
38 In order for a biomarker to be an accurate and valid indicator of FV intake, there are a number of
39 requirements that must be satisfied. Biomarkers of FV intake need to be minimally invasive to
40 participants, have the ability to discriminate between different FV intakes, be easy to measure,
41 reproducible, and be highly responsive to any change in FV intake [13-15].

42
43 Suggested possible biomarkers of FV intake include plasma vitamin C, carotenoids and flavonoids
44 [16-20]. In a recent systematic review [15] vitamin C and carotenoids were the two biomarkers that
45 were most frequently measured and consistently responsive within dietary FV interventions. Some
46 single biomarkers have been shown to be strong indicators of specific single FV, for example the
47 carotenoid lycopene is a good predictor of tomato intake [21]. However, while some of these proposed
48 biomarkers have been associated with a specific fruit or vegetable, or FV class, they have been less
49 reliably associated with overall FV consumption [17-19, 22-26]. This is likely to be due, at least in

50 part, to FV being a complex food group with variability in the bioactive compounds contained within
51 individual FV. Therefore, measuring a panel of potential biomarkers of FV intake within FV
52 intervention studies has been recommended [15]. Examining this panel in an integrated way may
53 more reliably predict overall FV consumption than a single individual biomarker or panel of individual
54 biomarkers.

55

56 This paper examines the effect of increased FV intake on a panel of biomarkers of FV consumption
57 ([vitamin C and six carotenoids](#)), considered both singly and in combination. Using data from a strictly
58 controlled randomised FV intervention study (BIOFA V) designed for this purpose, [i.e. with strict](#)
59 [dietary control to ensure compliance](#), and three further FV intervention studies, we sought to
60 determine whether a combined biomarker approach was better at predicting overall FV intake within
61 FV intervention studies than single biomarkers.

62

63 *Subjects and Methods*

64 The Biomarkers of Fruit and Vegetables (BIOFAV) study was a randomised controlled FV feeding
65 study (2, 5 or 8 portions of FV per day for 4 weeks) in healthy volunteers (n=30). The study was
66 approved by the School of Medicine, Dentistry and Biomedical Sciences research ethics committee of
67 Queen's University Belfast and participants gave informed written consent. The study was registered
68 at clinicaltrials.gov as NCT01591057. [The study duration was chosen to be long enough to allow the](#)
69 [proposed biomarkers to change in response to the alteration in FV intake, whilst the inclusion and](#)
70 [exclusion criteria ensured a broad range of ages and healthy volunteers were included, to maximise](#)
71 [generalisability of study findings. A sample size of n=10 per group was chosen for two reasons, firstly](#)
72 [because variability data from previous studies suggested that statistically significant increases in](#)
73 [blood-based FV biomarkers would be achieved with such a sample size, whilst this was also](#)
74 [achievable from a manpower perspective, given the intensity of the dietary intervention.](#)

75

76 Participants were recruited using university intranet emails and posters between June 2011 and May
77 2012. Inclusion criteria were: aged between 18 and 65 y; current consumption of FV \leq two portions/d,

78 while exclusion criteria were: body mass index >35 kg/m²; use of high dose vitamins, minerals or
79 dietary supplements likely to affect biomarkers of FV intake; excessive alcohol consumption (defined
80 as >28 units/week for males and >21 units/week for females); food sensitivities or allergies that would
81 interfere with the tolerance of a high FV-rich diet; current smoking; medical conditions or dietary
82 restrictions that would considerably limit the participant's ability to complete the study requirements;
83 history of diabetes; pregnant or lactating; following a weight loss diet. Participants' usual diet was
84 assessed using the 7-day diet history method.

85

86 Participants were randomly assigned, using a block design, to one of three intervention groups, to
87 consume 2, 5 or 8 portions of FV daily for 4 weeks.

88

89 All food, including the FV consumed during the intervention, was provided and there was supervised
90 consumption of two meals per day on weekdays. The menu plan was based around the participant's
91 portions allocation, their likes and dislikes and also their usual energy intakes. [Participants were](#)
92 [therefore free to choose the FV they wished to consume, although a balance of fruit versus vegetables](#)
93 [was encouraged, and variety also promoted.](#) A portion of FV was as recommended by Department of
94 Health (UK) guidelines, e.g. one apple, orange or banana, 3 heaped tablespoons of vegetables, or 150
95 ml fruit juice) [27].

96

97 A fasting blood sample was collected from all participants at baseline and week 4. All bloods were
98 processed within two hours of being drawn and stored at -80°C. Weight and height were also assessed
99 at baseline and week 4, and weight was re-measured at week 2 to ensure it remained constant over the
100 course of the study, [and diets were altered if weight loss or gain was observed.](#) Demographic
101 information was collected on alcohol consumption, smoking status, levels of physical activity (MRC
102 Recent Physical Activity Questionnaire), medication use at baseline and week 4 to ensure there were
103 no changes to these behaviours over the study duration.

104

105 Other fruit and vegetable intervention studies

106 The other FV intervention studies included in the current analysis were randomised interventions,
107 conducted in a similar way to BIOFA V, [and by the same research team](#), except the intervention was
108 less strictly controlled, i.e. FV were provided to the participants rather than whole diet, and there was
109 no supervised consumption. The FA VRIT study randomised participants with hypertension to 1, 3 or
110 6 portions of FV/d for 8 weeks, the ADIT study randomised participants >65 y to 2 or 5 portions of
111 FV/d for 16 weeks, while the FIRST study recruited participants at increased risk of CVD and
112 randomised them to 2, 4 or 7 portions FV/d for 12 weeks. Detailed methodology and analysis of the
113 primary outcomes of these studies have been published [5,7-8], and the studies are summarised in
114 **Table 1.**

115

116 Laboratory analysis

117 All biomarker analysis was conducted blinded to allocated FV group. For all studies, plasma ascorbic
118 acid concentrations were determined by fluorometric assay [28]. Serum concentrations of lutein,
119 zeaxanthin, β -cryptoxanthin, α -carotene, β -carotene and lycopene were measured by reverse phase
120 high performance liquid chromatography (HPLC) [29]. Assays were standardised against appropriate
121 National Institute of Standards and Technology reference materials. These assays are also externally
122 quality assured by participation in the French Society for Vitamins and Biofactors quality assurance
123 scheme.

124

125 Statistical Analysis

126 For baseline characteristics continuous variables are presented as mean (SD) and categorical variables
127 are presented as n (%). Between groups comparisons of baseline characteristics were made using one
128 way ANOVA tests for continuous variables and Chi-square tests for categorical variables.

129

130 Biomarker status variables were summarised as mean (standard deviation) and changes are expressed
131 as mean (95% CI). Changes in micronutrient status were assessed using one-way ANOVA, with a test
132 for linear trend across groups if there were > two intervention groups.

133

134 In BIOFA V, ordinal (proportion odds) logistic regression analyses were utilised to predict allocated
135 FV group (2, 5 or 8 portions of FV per day). Initially, models were fitted with single biomarkers
136 (vitamin C only, lutein only) and then a combined biomarker model was fitted (containing all
137 carotenoids and vitamin C). Biomarker variables entered into models were based upon change
138 between week 4 and baseline and were standardised to calculate (adjusted) odds ratios (OR) per
139 standard deviation increase in each variable and 95% confidence intervals (95% CI). To avoid
140 computational problems when calculating, for the full model, an optimism corrected C-statistic
141 (described later), a separate model was created in which factor analysis (based upon the principal-
142 factor method, including 2 factors and no rotation) was first used to reduce the carotenoid variables to
143 produce two factor score variables. These two factor score variables were entered into a model along
144 with vitamin C. The combined biomarker model was formally compared to the models with vitamin C
145 and lutein model using likelihood ratio tests. The proportional odds assumption was informally
146 checked by comparing the estimates from logistic regression models comparing 2 to 5 items of FV
147 intake per day and 5 to 8 items of FV intake per day.

148
149 The ability of the models to correctly classify FV intake (i.e. 2, 5 or 8 portions of FV per day) were
150 measured by creating 3 by 3 tables of predicted intake based upon the model scores against the
151 observed intake (i.e. the actual category of intake). The cut offs for the predicted intake categories
152 were chosen post-hoc to obtain the correct total number of predicted cases for each outcome category
153 overall. Based upon these 3 by 3 tables, the percentage agreement of predicted and observed intake
154 and a weighted Kappa value was calculated.

155
156 The discriminative ability of each model was formally quantified using a C-statistic. C-statistics and
157 confidence intervals were calculated using STATA add-on somersd [30]. This statistic can be
158 interpreted as the probability that, for a pair of individuals with different FV category intake, the one
159 in the higher FV intake category has higher predicted probability for that category from the model.
160 Bootstrap methods were applied to attempt to correct for over-fitting. Specifically, each model was
161 estimated in a bootstrap sample and the modified C-statistic was calculated in the bootstrap sample

162 and in the original sample. This process was repeated 200 times and the average difference in
163 performance in the bootstrap sample and in the original sample was calculated (the optimism) and
164 subtracted from the apparent performance to estimate the internally validated performance [31].

165

166 These analyses were replicated in FIRST (using ordinal logistic regression models with outcome 2, 4
167 and 7 FV per day and explanatory variables based upon the difference in biomarker levels from
168 baseline to week 12) and FA VRIT (using ordinal logistic regression models with outcome 1, 3 and 6
169 FV per day and using explanatory variables based upon the difference in biomarker levels from
170 baseline to week 8). In ADIT similar analyses were conducted using logistic regression, as only two
171 FV intake categories were used (2 and 5 FV per day), with explanatory variables based upon
172 difference in biomarker from baseline to week 16.

173

174 These analyses were conducted to predict allocated FV group, but, because of the reduced dietary
175 control in these studies (supply of FV and close contact with study team to encourage compliance), it
176 is possible that compliance with [the dietary intervention in](#) these [other](#) studies was less than [in](#)
177 BIOFA V. Linear regression analysis with self-reported FV intake as the outcome was therefore also
178 used to calculate an adjusted R^2 value for models containing change in vitamin C alone, lutein alone
179 and a combination of all carotenoids and vitamin C in FA VRIT, FIRST and ADIT.

180

181 Statistical analyses were performed using SPSS version 19.0 for Mac (SPSS Inc, Chicago, IL) and
182 STATA version 12 (StataCorp, College Station, Texas).

183

184 *Results*

185 A total of 30 participants completed the BIOFA V study. **Figure 1** shows the flow of participants
186 through the study. One participant had a missing blood sample at week 4. **Table 2** shows the baseline
187 study population characteristics. Also shown in table 2 are the baseline characteristics according to FV
188 intervention group (2, 5 or 8 portions FV/d). The baseline characteristics were similar in the three FV
189 intervention groups (2, 5 or 8 portions FV/d).

190
191 **Table 3** shows baseline and change in biomarker status in the three FV intervention groups. The
192 groups were similar at baseline, but there were statistically significant between group changes in all
193 biomarkers measured at week 4, with the exception of zeaxanthin and lycopene, although a similar
194 trend was observed for zeaxanthin.

195
196 **Table 4** illustrates the single and combined biomarker models used to predict allocated FV group in
197 BIOFAV participants. In the vitamin C only model, vitamin C was associated with FV group, whilst
198 in the lutein only model, lutein was also associated with FV group. Formal tests indicated that the
199 combined model containing all carotenoids and vitamin C was a better fit than either the vitamin C
200 only ($P<0.001$) model or the lutein only ($P=0.006$) model. The combined model correctly allocated
201 86% of individuals to the correct group compared with 52% in the vitamin C only model and 66% in
202 the lutein only model. These corresponded to Kappa values of 0.85 in the combined model compared
203 with 0.31 in the vitamin C only and 0.54 in the lutein only models. The C-statistic was slightly lower
204 in the lutein only model (0.84) and in the model based upon factor analysis (0.88), and much lower in
205 the vitamin C model (0.68) compared with the full model (0.95). An optimism corrected C-statistic
206 was not calculable for the full model, however correction for optimism reduced the C-statistic for the
207 factor analysis model (0.85) but the C-statistics for the vitamin C model and lutein model were little
208 altered (0.68 and 0.84).

209
210 Results for the other studies are shown in **Table 5**, but the differences between the models were less
211 marked. For instance, the optimism corrected C-statistics was slightly higher in the combined model
212 compared with the vitamin C only model and lutein only model in FA VRIT (0.76, 0.64 and 0.65,
213 respectively) and in ADIT (0.75, 0.72 and 0.67, respectively), but not in FIRST (0.64, 0.58 and 0.66,
214 respectively). Similarly, adjusted r^2 values were higher in the combined model compared with the
215 vitamin C only model and lutein only in FA VRIT (0.22, 0.03 and 0.16, respectively) and in ADIT
216 (0.26, 0.12 and <0.01 , respectively), but not in FIRST (0.11, <0.01 and 0.15, respectively).

217

218 When analyses were replicated to predict allocated FV group in each of the four studies based upon
219 the final biomarker values as opposed to the change variables, similar levels of agreement were
220 demonstrated (data not shown).

221

222 *Discussion*

223 [This study suggests that a combined biomarker panel may better predict FV intake within intervention](#)
224 [studies than consideration of single biomarkers.](#) Plasma vitamin C, and certain carotenoids and
225 flavonoids have been suggested as biomarkers of intake of overall FV [16-20, 23-24, 26, 32-343]. A
226 systematic review of FV biomarkers within FV intervention studies showed that vitamin C and
227 carotenoids are commonly used and do tend to increase [15]. However, there are two problems with
228 the assessment of these biomarkers of overall FV intake: firstly FV is a complex food group, and the
229 content of phytochemicals proposed as biomarkers will vary markedly between different classes of
230 FV, and even within different varieties of the same FV [3435-3637]. Storage and processing of FV
231 can also affect phytochemical content [3738]. Secondly, there are a number of factors that will affect
232 biomarker response to a given FV intake, including inter-individual variation in digestion, absorption
233 and metabolism. Some of this variation will be genetic in origin, but environmental factors such as
234 BMI [3839], smoking [3940], baseline biomarker status [4041,4142] and other aspects of diet (e.g. fat
235 content of meal in which FV are consumed [24, 4243-4546]) will affect bioavailability of the
236 phytochemicals and/or biomarker response. Nutritional biomarkers will therefore never be perfect at
237 reflecting dietary FV consumption on their own, because physiological processes in the body will also
238 impact upon response. Furthermore, there are other dietary sources of some of these compounds and
239 therefore single compounds may not reflect total overall FV consumption. In addition, the plasma
240 concentration of vitamin C has a linear relationship with vitamin C intake up until a specific point,
241 above which plasma concentrations of vitamin C plateau (>5 servings of FV) [19,4647], and therefore
242 vitamin C may be a less useful biomarker at higher levels of intake.

243

244 Despite these limitations, efforts to improve the ability of biomarkers to predict FV intake would be
245 valuable in nutritional epidemiology. Campbell et al, [26] initially suggested combining biomarkers,

246 | and summing carotenoid status [4748] or flavonoid urinary excretion [4849-5051] has previously been
247 | attempted. Such an approach, however, will take greater account of the more predominant compounds
248 | (e.g. lycopene when summing total carotenoids), and therefore a more sophisticated approach may be
249 | required. A recent publication has conducted a similar analysis to that performed here in a single, less
250 | well-controlled FV intervention [5452].

251 |
252 | Because of the limitations of the FV biomarkers when examining overall FV intake, we proposed that
253 | a combined biomarker approach ([vitamin C and six carotenoids](#)) may be able to take into account the
254 | diversity and variety of bioactive compounds found within FV and would be more likely to capture the
255 | total amount of FV consumed. ~~Hence the current analysis aimed to determine whether a model~~
256 | ~~including a panel of biomarkers (vitamin C and six carotenoids) was better at predicting FV intake~~
257 | ~~within a FV intervention study than models examining individual biomarkers. These~~ The analyses
258 | utilised a tightly controlled FV intervention study (~~all food was provided and two meals/d on~~
259 | ~~weekdays consumed under supervision;~~ (BIOFA V) and three previously-conducted, less tightly
260 | controlled FV interventions (FA VRIT, FIRST and ADIT). The combined biomarker model performed
261 | better in the more tightly controlled intervention, BIOFA V. Similar patterns of results were observed
262 | for the other studies, but differences between the combined biomarker and individual biomarker
263 | models were less marked, and this was particularly true for the FIRST study, where the combined
264 | model was not significantly better than the lutein only model. This suggests that an integrated panel of
265 | biomarkers in intervention studies may obtain a more accurate and precise measure of total FV
266 | consumption, but the observed differences detected between studies do need to be explored.

267 |
268 | It is likely that the difference detected within the current studies arose due to the less intensive and less
269 | dietary control in the other three dietary interventions compared to BIOFA V, although efforts were
270 | made in each of these [other](#) interventions to maximise compliance with the allocated intervention and
271 | minimize inaccuracies in self-reporting. For example, FV was delivered to the participants in all three
272 | studies, participants were contacted weekly by telephone to monitor compliance with the intervention
273 | study, and participants were encouraged to report any lack of compliance with the intervention. It is

274 possible that there was a lack of adherence to the dietary intervention, and therefore the analysis
275 according to allocated FV group would not have been appropriate. However, the association between
276 biomarker response and self-reported FV intake (in portions/d) was also calculated and revealed
277 similar associations. It is still possible that the self-reported measures utilised to assess numbers of
278 portions of FV consumed were inaccurately reported. There may also have been between-study
279 differences in the FV self-selected by participants or in the composition of the FV consumed, which
280 could have impacted on biomarker responses. For example, the range of FV consumed within
281 BIOFAV may have been less broad than in the other studies, although this is unlikely, [and similar](#)
282 [guidance regarding selection of FV \(a balance of FV, encouragement to maximise variety\) was given](#)
283 [in all studies](#). Similarly, the FV supplier was the same in all studies, and all fieldwork duration
284 periods were long enough to account for between-study seasonal differences in FV composition.
285 Another major difference between BIOFAV and other FV intervention studies was that BIOFAV
286 recruited young healthy individuals, whilst the other studies recruited either older participants or those
287 at high CVD risk. BIOFAV participants were non-smokers, and had BMIs in the normal range,
288 whereas those in the other studies included smokers, and participants were overweight. It is possible
289 that there is an age-related reduction in FV biomarker response, or that biomarker responses are
290 reduced with increasing weight [3839] and in smokers [3940] and that this contributed to the
291 differences observed.

292

293 The analyses presented here have several strengths. They are the first to examine combining FV
294 biomarkers to improve prediction of FV intake in a tightly controlled intervention, or other than simple
295 summation [4748-5051]. The BIOFAV study was strictly controlled, and we are confident that
296 participants consumed the number of FV portions they were allocated to, therefore our observation of
297 improved prediction of FV intake using a combined panel of biomarkers is novel and robust. The
298 ability to test this hypothesis in further FV interventions is also a strength. There are, however, some
299 accompanying limitations – the overall absolute performance of the model to predict intake,
300 particularly in the less controlled studies, was relatively low (e.g. an r^2 of 0.26 in ADIT), indicating
301 there is substantial variability in FV intake not explained by the included biomarkers, and that what we

302 are commenting on is the relative improvement with a combined biomarker approach. A further
303 limitation is that internal validation was not possible for the combined model in BIOFA V, due to the
304 small number of observations relative to the parameters estimated, and therefore it was not possible to
305 correct for over fitting/optimism in the models. Although internal validation was performed for the
306 other models, a better estimate of the prognostic power of these models would be determined using
307 external validation. The combined biomarker approach could be extended and incorporate further
308 biomarkers (e.g. flavonoids) and account for environmental factors that could affect biomarker
309 response (e.g. BMI, smoking), while the ability of such an approach, compared with single
310 biomarkers, to predict individual intakes on a population level would require further model
311 development and testing. The utility of a combined biomarker approach to predict different patterns of
312 FV intake (e.g. predominantly fruit or vegetable) [or diets containing different classes of FV \(e.g. green](#)
313 [vegetables, root vegetables, pulses, fruits\), or different varieties of the same FV classes](#) could also be
314 explored. A combination of multiple measures of dietary assessment may provide more accurate
315 estimates of true dietary intake, for example combining biomarkers and self-reported measures of
316 dietary intake using regression calibration equations [[5253,5354](#)], and this could represent a natural
317 extension of combining biomarkers when measuring FV intake. The biomarkers measured here are
318 commonly used in FV intervention studies and therefore the suggested approach, if confirmed as
319 useful, would be at no extra cost.

320
321 In conclusion, there was some evidence that a combined model including a range of FV biomarkers (a
322 carotenoid panel and vitamin C) performed better at predicting allocated FV within a strictly
323 controlled dietary intervention study than the use of a model with vitamin C only or lutein only. A
324 similar pattern was observed in three less intensive FV interventions, although differences between
325 models were less clear cut. Therefore, a combined biomarker approach to assess overall FV
326 consumption may more accurately predict FV intake within intervention studies than the use of single
327 biomarkers. The utility of such an approach to predict population level intake of FV remains to be
328 tested.

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345 *Acknowledgments*

346 Grateful thanks to the technical assistance of Shrobona Bhattacharya in terms of food preparation and
347 delivery and Sarah Gilchrist in terms of biomarker analysis.

348 *Statement of authors' contributions to manuscript*

349 KMA, JM, MCM, ISY and JVW designed research; AJM, LLH, CRD and CEN conducted research;
350 CRC and AJM analyzed data; AJM, JVW and CRC wrote the paper; JVW had primary responsibility
351 for final content. All authors read and approved the final manuscript. There are no conflicts of
352 interest.

353 *Conflict of interest*

354 The authors declare that they have no conflict of interest.

355

356 **References**

- 357 (1) Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, Leschik-Bonnet E, Müller MJ,
358 Oberritter H, Schulze M, et al. (2012) Critical review: vegetables and fruit in the prevention of chronic
359 diseases. *Eur J Nutr* 51:637–663.
- 360 (2) He FJ, Nowson CA, Lucas M, MacGregor GA (2007) Increased consumption of fruit and vegetables is
361 related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. *J Hum Hypertens*
362 21:717-728.
- 363 (3) He FJ, Nowson CA, MacGregor GA (2006) Fruit and vegetable consumption and stroke: meta-analysis
364 of cohort studies. *Lancet* 367:320-326.
- 365 (4) Macready AL, George TW, Chong MF, Alimbetov DS, Jin Y, Vidal A, Spencer JP, Kennedy OB,
366 Tuohy KM, Minihane AM, et al; FLAVURS Study Group (2014) Flavonoid-rich fruit and vegetables
367 improve microvascular reactivity and inflammatory status in men at risk of cardiovascular disease--
368 FLAVURS: a randomized controlled trial. *Am J Clin Nutr* 99:479-489.
- 369 (5) Wallace IR, McEvoy CT, Hunter SJ, Hamill LL, Ennis CN, Bell PM, Patterson CC, Woodside JV,
370 Young IS, McKinley MC (2013) Dose-response effect of fruit and vegetables on insulin resistance in
371 people at high risk of cardiovascular disease: a randomized controlled trial. *Diabetes Care* 36:3888-3896.
- 372 (6) Neville CE, Young IS, Gilchrist SE, McKinley MC, Gibson A, Edgar JD, Woodside JV (2014) Effect
373 of increased fruit and vegetable consumption on bone turnover in older adults: a randomised controlled
374 trial. *Osteoporos Int* 25:223-233.
- 375 (7) Gibson A, Edgar JD, Neville CE, Gilchrist SE, McKinley MC, Patterson CC, Young IS, Woodside JV
376 (2012) Effect of fruit and vegetable consumption on immune function in older people: a randomized
377 controlled trial. *Am J Clin Nutr* 96:1429-1436.
- 378 (8) McCall DO, McGartland C, McKinley MC, Patterson CC, Sharpe P, McCance D, Young IS, Woodside
379 JV (2009) Dietary intake of fruits and vegetables improves microvascular function in hypertensive subjects
380 in a dose dependent manner. *Circulation* 119:2153-2160.
- 381 (9) Bingham SA (2001) Biomarkers in nutritional epidemiology. *Public Health Nutr* 5:821-827.
- 382 (10) Lee RD, Nieman DC (2003) *Nutritional Assessments – Third Edition*. McGraw-Hill: New York.
- 383 (11) Livingstone MBE, Black AE (2003) Markers of the Validity of Reported Energy Intake. *J Nutr*
384 133:895S-920S.

- 385 (12) Woodside JV, Young IS, McKinley MC (2013) Fruits and vegetables: measuring intake and
386 encouraging increased consumption. *Proc Nutr Soc* 72:236-245.
- 387 (13) Hunter D (1990) Biochemical indicators of dietary intake. In *Nutritional Epidemiology*, pp. 143–216
388 [W Willett, editor]. New York: Oxford University Press.
- 389 (14) Crews H, Alink G, Anderson R, Braesco V, Holst B, Maiani G, Ovesen L, Scotter M, Solfrizzo M, van
390 den Berg R et al. (2001) A critical assessment of some biomarker approaches linked with dietary intake. *Br*
391 *J Nutr* 86:S5–S35.
- 392 (15) Baldrick FR, Woodside JV, Elborn JS, Young IS, McKinley MC (2011) Biomarkers of fruit and
393 vegetable intake in human intervention studies: a systematic review. *Crit Rev Food Sci Nutr* 51:795–815.
- 394 (16) Block G, Norkus E, Hudes M, Mandel S, Helzlsouer K (2001) Which plasma antioxidants are most
395 related to fruit and vegetable consumption? *Am J Epidemiol* 154:1113-1118.
- 396 (17) Al-Delaimy WK, Ferrari P, Slimani N, Pala V, Johansson I, Nilsson S, Mattisson I, Wirfalt E, Galasso
397 R, Palli D et al. (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables: individual-level
398 correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Eur J Clin Nutr*
399 59:1387–1396.
- 400 (18) Al-Delaimy WK, Slimani N, Ferrari P, Key T, Spencer E, Johansson I, Johansson G, Mattisson I,
401 Wirfalt E, Sieri S et al. (2005) Plasma carotenoids as biomarkers of intake of fruits and vegetables:
402 ecological-level correlations in the European Prospective Investigation into Cancer and Nutrition (EPIC).
403 *Eur J Clin Nutr* 59:1397–1408.
- 404 (19) Padayatty SJ, Levine M (2008) Fruit and vegetables: think variety, go ahead, eat! *Am J Clin Nutr*
405 87:5-7.
- 406 (20) Zino S, Skeaff M, Williams S, Mann J (1997) Randomised controlled trial of effect of fruit and
407 vegetable consumption on plasma concentrations of lipids and antioxidants. *BMJ* 314:1787.
- 408 (21) Riso P, Visioli F, Erba D, Testolin G, Porrini M (2004) Lycopene and vitamin C concentrations
409 increase in plasma and lymphocytes after tomato intake. Effects on cellular antioxidant protection. *Eur J*
410 *Clin Nutr* 58:1350-1358.
- 411 (22) Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S, Lubin R, Thurnham DI, Key TJ,
412 Roe L et al. (1997) Validation of dietary assessment methods in the UK arm of EPIC using weighed
413 records, and 24-hour urinary nitrogen and potassium and vitamin C and carotenoids as biomarkers. *Int J*

- 414 Epidemiol 26:S137–S151.
- 415 (23) Jansen MC, van Kappel AL, Ocké MC, Van 't Veer P, Boshuizen HC, Riboli R, Bueno-de-Mesquita
416 HB (2004) Plasma carotenoid levels in Dutch men and women, and the relation with vegetable and fruit
417 consumption. *Eur J Clin Nutr* 58:1386–1395.
- 418 (24) Drewnowski A, Rock CL, Henderson SA, Shore AB, Fischler C, Galan P, Preziosi P, Hercberg S
419 (1997) Serum beta-carotene and vitamin C as biomarkers of vegetable and fruit intakes in a community-
420 based sample of French adults. *Am J Clin Nutr* 65:1796–1802.
- 421 (25) Ferrari P, Al-Delaimy WK, Slimani N, Boshuizen HC, Roddam A, Orfanos P, Skeie G, Rodríguez-
422 Barranco M, Thiebaut A, Johansson G et al. (2005) An approach to estimate between- and within-group
423 correlation coefficients in multicenter studies: plasma carotenoids as biomarkers of intake of fruits and
424 vegetables. *Am J Epidemiol* 162:591–598.
- 425 (26) Campbell DR, Gross MD, Martini MC, Grandits GA, Slavin JL, Potter JD (1994) Plasma carotenoids
426 as biomarkers of vegetable and fruit intake. *Cancer Epidemiol Biomarkers Prev* 3:493–500.
- 427 (27) National Health Service. Rough guide- Fruit and vegetable portion sizes. Internet:
428 http://www.nhs.uk/Livewell/5ADAY/Documents/Downloads/5ADAY_portion_guide.pdf (accessed 25th
429 July 2014).
- 430 (28) Vuilleumier JP, Keck E (1989) Fluorometric assay of vitamin C in biological materials using a
431 centrifugal analyser with fluorescence attachment. *J Micronut Analysis* 5:25-34.
- 432 (29) Craft NE, Wise SA, Soares JH (1992) Optimisation of an isocratic high performance liquid
433 chromatography separation of carotenoids. *J Chromatogr* 589:171-176.
- 434 (30) Newson R (2006) Confidence intervals for rank statistics: Somers' D and extensions. *Stata Journal*
435 6:309-334.
- 436 (31) Steyerberg EW (2010) *Clinical Prediction Models: A Practical Approach to Development, Validation,*
437 *and Updating.* Springer.
- 438 (32) Le Marchand L, Hankin JH, Carter FS, Essling C, Luffey D, Franke AA, Wilkens LR, Cooney RV,
439 Kolonel LN (1994) A pilot study on the use of plasma carotenoids and ascorbic acid as markers of
440 compliance to a high fruit and vegetable dietary intervention. *Cancer Epidemiol Biomarkers Prev* 3:245-
441 251.

- 442 (33) Tucker KL, Chen H, Vogel S, Wilson PW, Schaefer EJ, Lammi-Keefe C (1999). Carotenoid intakes,
443 assessed by dietary questionnaire, are associated with plasma carotenoid concentrations in an elderly
444 population. *J Nutr* 129:438–445.
- 445 [\(34\) Chong MF, George TW, Alimbetov D, Jin Y, Weech M, Macready AL, Spencer JP, Kennedy OB,](#)
446 [Minihane AM, Gordon MH, Lovegrove JA \(2013\). Impact of the quantity and flavonoid content of fruits](#)
447 [and vegetables on markers of intake in adults with an increased risk of cardiovascular disease: the](#)
448 [FLAVURS trial. *Eur J Nutr* 52:361-78.](#)
- 449 (3435) Szeto YT, Tomlinson B, Benzie IF (2002) Total antioxidant and ascorbic acid content of fresh fruits
450 and vegetables: implications for dietary planning and food preservation. *Br J Nutr* 87:55-59.
- 451 (3536) Heinonen, MI (1990) Carotenoids and provitamin A activity of carrot (*Daucus carota* L.) cultivars. *J*
452 *Agric Food Chem* 38:609–612.
- 453 (3637) Kurilich, AC, Juvik JA (1999) Quantification of carotenoids and tocopherol antioxidants in *Zea*
454 *mays*. *J Agric. Food Chem* 47:1948–1955.
- 455 (3738) Kalt W (2005) Effects of production and processing factors on major fruit and vegetable
456 antioxidants. *J Food Science* 70:R11-19.
- 457 (3839) Vioque J, Weinbrenner T, Asensio L, Castelló A, Young IS, Fletcher A (2007) Plasma
458 concentrations of carotenoids and vitamin C are better correlated with dietary intake in normal weight than
459 overweight and obese elderly subjects. *Br J Nutr* 97:977-986.
- 460 (3940) Alberg A (2002) The influence of cigarette smoking on circulating concentrations of antioxidant
461 micronutrients. *Toxicology* 180:121-137.
- 462 (4041) Jacques, PF, Hartz SC, McGandy RB, Jacob RA, Russell RM (1987) Ascorbic acid, HDL, and total
463 plasma cholesterol in the elderly. *J Am Coll Nutr* 6:169–174.
- 464 (4142) Sanchez-Moreno C, Cano MP, de Ancos B, Plaza L, Olmedilla B, Granado F, Martin A (2003)
465 Effect of orange juice intake on vitamin C concentrations and biomarkers of antioxidant status in humans.
466 *Am J Clin Nutr* 78:454–460.
- 467 (4243) Brown MJ, Ferruzzi MG, Nguyen ML, Cooper DA, Eldridge AL, Schwartz SJ, White WS (2004)
468 Carotenoid bioavailability is higher from salads ingested with full-fat than with fat-reduced salad dressings
469 as measured with electrochemical detection. *Am J Clin Nutr* 80:396–403.
- 470 (4344) Unlu NZ, Bohn T, Clinton SK, Schwartz SJ (2005) Carotenoid absorption from salad and salsa by

- 471 humans is enhanced by the addition of avocado or avocado oil. *J Nutr* 135:431–436.
- 472 | (4445) Hussein L, El-Tohamy M (1990) Vitamin A potency of carrot and spinach carotenes in human
473 | metabolic studies. *Int J Vit Nutr Res* 60:229–235.
- 474 | (4546) Castenmiller JJM, West CE, Linssen JPH, van het Hof KH, Voragen AGJ (1999) The food matrix
475 | of spinach is a limiting factor in determining the bioavailability of β -carotene and to a lesser extent of lutein
476 | in humans. *J Nutr* 129:349–355.
- 477 | (4647) Dehghan M, Akhtar-Danesh N, McMillan CR, Thabane L (2007) Is plasma vitamin C an
478 | appropriate biomarker of vitamin C intake? A systematic review and meta-analysis. *Nutr J* 6:41–53.
- 479 | (4748) Woodside JV, Young IS, Gilchrist SE, Vioque J, Chakravarthy U, de Jong PT, Rahu M, Seland J,
480 | Soubrane G, Tomazzoli L, et al. (2013) Factors associated with serum/plasma concentrations of vitamins
481 | A, C, E and carotenoids in older people throughout Europe: the EUREYE study. *Eur J Nutr* 52:1493-1501.
- 482 | (4849) Krogholm KS, Haraldsdottir J, Knuthsen P, Rasmussen SE (2004) Urinary total flavonoid excretion
483 | but not 4-pyridoxic acid or potassium can be used as a biomarker for the intake of fruits and vegetables. *J*
484 | *Nutr* 134:445-451.
- 485 | (4950) Nielsen SE, Freese R, Kleemola P, Mutanen M (2002) Flavonoids in human urine as biomarkers for
486 | intake of fruits and vegetables. *Cancer Epidemiol Biomarkers Prev* 11:459-466.
- 487 | (5051) Mennen LI, Sapinho D, Ito H, Bertrais S, Galan P, Hercberg S, Scalbert A (2006) Urinary
488 | flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. *Br J Nutr* 96:191-198.
- 489 | (5152) Yannan J, Gordon MH, Alimbetov D, Chong MF, George TW, Spencer JPE, Kennedy OB, Tuohy
490 | K, Minihihane AM, Lovegrove JA, for the FLAVURS Study Group (2014) A Novel Combined Biomarker
491 | including Plasma Carotenoids, Vitamin C, and Ferric Reducing Antioxidant Power Is More Strongly
492 | Associated with Fruit and Vegetable Intake than the Individual Components. *J Nutr* Epub ahead of print
493 | September 17, 2014, doi: 10.3945/jn.114.192856.
- 494 | (5253) Freedman LS, Midthune D, Carroll RJ, Tasevska N, Schatzkin A, Mares J, Tinker L, Potischman N,
495 | Kipnis V (2011) Using regression calibration equations that combine self-reported intake and biomarker
496 | measures to obtain unbiased estimates and more powerful tests of dietary associations. *Am J Epidemiol*
497 | 174:1238-1245.
- 498 | (5354) Freedman LS, Kipnis V, Schatzkin A, Tasevska N, Potischman N (2010) Can we use biomarkers in
499 | combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol*

500 Perspect Innov 7:2.

501

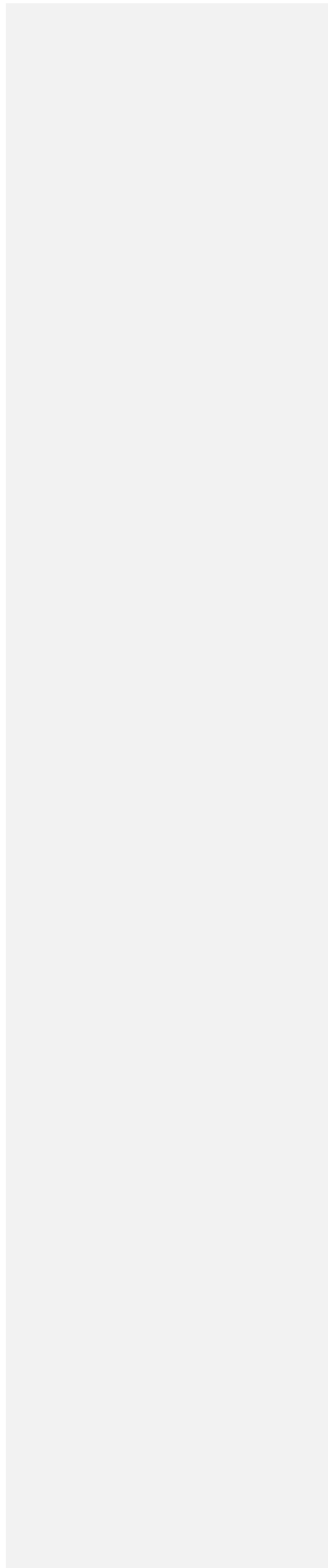


Table 1. Summary details of the intervention studies included in the combined biomarker analysis.

Study	Number of participants completing intervention	Study Duration	Portions/d	Dietary intervention intensity	Age (y)	Health Status	Assessment visits when blood collected
BIOFAV	30	4 weeks (no wash out required as 1-2 portions /day FV consumers recruited)	2, 5 or 8	All food provided, including FV. Two meals/d consumed under supervision.	18-65	Healthy	0, 2 and 4 weeks
FIRST [5]	89	12 weeks (+4 week run-in-period)	2, 4 or 7	FV supplied weekly. Compliance encouraged by weekly telephone call.	40-77	High risk of CVD	0 and 12 weeks of intervention period
FAVRIT [8]	117	8 weeks (+4 week run-in-period)	1, 3 or 6	FV supplied weekly. Compliance encouraged by weekly telephone call.	40-65	Hypertensive	0 and 8 weeks of intervention period
ADIT [6,7]	82	16 weeks (no wash out required as ≤ 2 portions /day FV consumers recruited)	2 or 5	FV supplied weekly. Compliance encouraged by weekly telephone call.	65-85	Healthy	0, 6, 12 and 16 weeks

Formatted Table

Table 2. Baseline characteristics of BIOFAV participants (n=30) according to FV intervention group

Participant Characteristic	Total	2 portion group (n=11)	5 portion group (n=9)	8 portion group (n=10)
Age (years)^a	28.5 (11.8)	25.9 (10.8)	28.2 (9.6)	31.7 (14.9)
Sex n (% female)^b	15 (50)	6 (54.5)	4 (44.4)	5 (50)
Weight (kg)^a	69.7 (13.7)	70.9 (17.8)	70.4 (8.6)	67.8 (13.5)
Height (m)^a	1.70 (0.1)	1.70 (0.1)	1.7 (0.1)	1.69 (0.07)
BMI (kg/m²)^a	24.0 (4.1)	24.6 (5.4)	23.8 (2.7)	23.7 (3.6)
Waist (cm)^a	81.7 (12.2)	82.0 (13.8)	83.7 (11.2)	79.6 (12.0)
Hip (cm)^a	100.7 (6.9)	101.5 (8.9)	100.6 (5.2)	99.8 (6.3)
Systolic blood pressure (mmHg)^a	109.6 (29.2)	109.3 (23.7)	107.1 (32.2)	112.2 (34.5)
Diastolic blood pressure (mmHg)^a	72.7 (21.6)	74.8 (22.1)	70.6 (23.8)	72.3 (21.2)
Past smoker n (%)^b	6 (20)	2 (18.2)	2 (22.2)	2 (20.0)
Alcohol consumers n (%)^b	22 (73.4)	9 (81.8)	7 (77.8)	6 (60.0)
Never or occasionally	8 (26.7)	2 (18.2)	2 (22.2)	4 (40.0)
Once or twice a week	20 (66.7)	9 (81.8)	6 (66.7)	5 (50.0)
Three to five times a week	2 (6.7)	0 (0)	1 (11.1)	1 (10.0)
Full time education (years)^a	16.1 (2.8)	15.9 (1.5)	16.2 (3.1)	16.3 (3.8)
Using medication n (%)^b	13 (43.3)	4 (36.4)	5 (55.6)	4 (40.0)
Employment^b				
Student	18 (60)	8 (72.7)	4 (44.4)	6 (60)
Full time employment	11 (36.7)	3 (27.3)	4 (44.4)	4 (40)
Unemployed	1 (3.3)	0 (0)	1 (11.1)	0 (0)

^aContinuous variables presented as mean (SD). ^bCategorical variables presented as n(%).

Table 3. Serum and plasma FV biomarker status of BIOFAV participants at baseline and during intervention in those consuming 2, 5 or 8 portions of FV/d

		<i>n</i>	Baseline	Change at 4 wk
Vitamin C (μmol/l)	2 portions/d	11	58.4 (16.3)	-7.7 (-20.9,5.5)
	5 portions/d	9	61.0 (13.1)	14.0 (2.9,25.1)
	8 portions/d	9	52.6 (20.5)	11.7 (-6.7,30.0)
	<i>P</i> -value		0.45	0.04
Lutein (μmol/l)	2 portions/d	11	0.16 (0.07)	-0.01 (-0.03,0.01)
	5 portions/d	9	0.15 (0.05)	0.03 (0.02,0.05)
	8 portions/d	9	0.16 (0.03)	0.06 (0.03,0.08)
	<i>P</i> -value		0.89	<0.001
Zeaxanthin (μmol/l)	2 portions/d	11	0.04 (0.02)	0.002 (-0.003,0.006)
	5 portions/d	9	0.05 (0.02)	0.006 (0.003,0.009)
	8 portions/d	9	0.05 (0.01)	0.008 (-0.002,0.018)
	<i>P</i> -value		0.69	0.11
β-cryptoxanthin (μmol/l)	2 portions/d	11	0.10 (0.09)	-0.01 (-0.03,0.02)
	5 portions/d	9	0.08 (0.04)	0.03 (0.004,0.06)
	8 portions/d	9	0.08 (0.02)	0.06 (0.01,0.11)
	<i>P</i> -value		0.41	0.005
α-carotene (μmol/l)	2 portions/d	11	0.14 (0.10)	0.08 (0.02,0.14)
	5 portions/d	9	0.10 (0.06)	0.13 (0.03,0.23)
	8 portions/d	9	0.13 (0.08)	0.34 (0.13,0.54)
	<i>P</i> -value		0.80	0.003
β-carotene (μmol/l)	2 portions/d	11	0.34 (0.25)	0.14 (0.01,0.26)
	5 portions/d	9	0.24 (0.19)	0.36 (0.05,0.67)
	8 portions/d	9	0.48 (0.44)	0.88 (0.31,1.46)
	<i>P</i> -value		0.33	0.002
Lycopene (μmol/l)	2 portions/d	11	0.51 (0.26)	0.064 (-0.032,0.159)
	5 portions/d	9	0.37 (0.18)	0.102 (-0.071,0.276)
	8 portions/d	9	0.48 (0.26)	0.003 (-0.218,0.224)
	<i>P</i> -value		0.80	0.55

Values are mean (SD) with changes expressed as mean (95% CI). Change calculated as wk 4 – baseline. Both baseline variables and changes were compared between 2, 5 and 8 portions/d groups using one way analysis of variance with linear trend fitted.

Table 4. Ordinal logistic regression to predict allocated FV group in BIOFAV participants utilising single or combined biomarker models

		Vitamin C only model		Lutein only model		All carotenoids and vitamin C model		Factor analysis model (all Carotenoids)	
		OR (95% CI)	P	OR (95% CI)	P	OR (95%CI)	P	OR (95% CI)	P
Estimates^a: OR (95% CI) per standard deviation increase	Vitamin C	2.45 (1.09, 5.54)	0.03			1.87 (0.51, 6.93)	0.35	1.59 (0.65, 3.87)	0.312
	Lutein			8.18 (2.19, 30.54)	0.002	30.75 (1.28, 736.55)	0.04		
	Zeaxanthin					0.23 (0.02, 2.46)	0.23		
	Lycopene					0.32 (0.09, 1.11)	0.07		
	β-cryptoxanthin					4.11 (0.49, 34.48)	0.19		
	α-carotene					1.82 (0.04, 92.06)	0.77		
	β-carotene					2.19 (0.04, 107.34)	0.69		
	First factor (footnote) ^b							12.71 (2.88, 56.04)	0.001
Second Factor (footnote) ^c							2.28 (0.73, 7.14)	0.157	
Model Performance	P value compared with full model ^d	<0.001		0.006		Reference		<0.001	
	Percentage agreement ^e	52 % (15/29)		66 % (19/29)		86 % (25/29)		75 % (22/29)	
	Weighted Kappa	0.31		0.54		0.85		0.69	
	C-statistic	0.68 (0.50, 0.86)		0.84 (0.71, 0.96)		0.95 (0.89, 1.00)		0.88 (0.76, 1.00)	
	Optimism corrected C-statistic	0.68		0.84		^f		0.85	

^a Estimates based upon ordinal logistic regression with outcome intake in groups (2, 5 and 8 portions per day) and explanatory variables change in biomarker values at 4 weeks.

^b First factor calculated from factor analysis based upon all carotenoids (change at 4 weeks). Calculate score = 0.22 X Lutein + 0.20 X Zeaxanthin + 0.02 X Lycopene + 0.09 X β-cryptoxanthin + 0.16 X α-carotene + 0.62 X β-carotene.

^c Second factor calculated from factor analysis based upon all carotenoids (change at 4 weeks). Calculate score = 0.41 X Lutein + 0.30 X Zeaxanthin + 0.04 X Lycopene + 0.14 X β-cryptoxanthin + -0.52 X α-carotene + -0.11 X β-carotene.

^d P value for likelihood ratio test comparing each model to model containing all carotenoids and vitamin C

^e Comparing actual category of intake with predicted category of intake based upon model cut offs (chosen to obtain the correct proportion in each category of intake).

^f Optimism could not be calculated due to large number of parameters compared with number of observations.

Table 5. Regression analysis to predict allocated FV group in FAVRIT, FIRST and ADIT studies utilising single or combined biomarker model approach

		Vitamin C model	Lutein model	All carotenoids and vitamin C model
FAVRIT study^a				
Estimates: OR (95% CI) per standard deviation increase	Vitamin C	1.53 (1.05, 2.24)		1.71 (1.14, 2.56)
	Lutein		2.15 (1.40, 3.30)	1.69 (0.87, 3.27)
	Zeaxanthin			0.68 (0.32, 1.41)
	Lycopene			0.52 (0.30, 0.91)
	β -cryptoxanthin			6.23 (2.65, 14.63)
	α -carotene			1.33 (0.60, 2.94)
	β -carotene			0.60 (0.27, 1.32)
P compared with full model ^d		<0.001	<0.001	Reference
Model Performance	Percentage agreement ^e	41% (41/100)	44% (44/100)	61% (61/100)
	Weighted Kappa	0.21	0.24	0.47
	C-statistic	0.64 (0.55, 0.73)	0.65 (0.57, 0.74)	0.80 (0.71, 0.88)
	Optimism corrected C-statistic	0.64	0.65	0.76
	Adjusted r^{2f}	0.03	0.16	0.22
FIRST study^b				
Estimates: OR (95% CI) per standard deviation increase	Vitamin C	1.14 (0.77, 1.69)		0.98 (0.64, 1.49)
	Lutein		1.96 (1.25, 3.08)	2.89 (1.30, 6.46)
	Zeaxanthin			0.25 (0.05, 1.13)
	Lycopene			0.89 (0.53, 1.52)
	β -cryptoxanthin			2.48 (1.09, 5.63)
	α -carotene			2.61 (0.23, 29.30)
	β -carotene			0.77 (0.36, 1.64)
P compared with full model ^d		0.02	0.40	Reference
Model Performance	Percentage agreement ^e	40% (33/83)	41% (34/83)	45% (37/83)
	Weighted Kappa	0.12	0.19	0.30
	C-statistic	0.58 (0.48, 0.68)	0.66 (0.57, 0.75)	0.70 (0.61, 0.79)
	Optimism corrected C-statistic	0.58	0.66	0.64
	Adjusted r^{2f}	<0.01	0.15	0.11
ADIT study^c				
Estimates: OR (95% CI) per standard deviation increase	Vitamin C	2.39 (1.37, 4.17)		2.37 (1.24, 4.54)
	Lutein		1.70 (0.99, 2.93)	1.59 (0.74, 3.43)
	Zeaxanthin			2.76 (0.25, 30.63)
	Lycopene			1.03 (0.50, 2.12)
	β -cryptoxanthin			1.91 (0.94, 3.92)
	α -carotene			5.08 (0.62, 41.71)
	β -carotene			0.28 (0.10, 0.74)
P compared with full model ^d		0.02	<0.001	Reference
Model Performance	Percentage agreement ^e	70% (55/79)	62% (49/79)	75% (59/79)
	Kappa	0.39	0.24	0.49
	C-statistic	0.73 (0.62, 0.84)	0.68 (0.56, 0.80)	0.82 (0.72, 0.91)
	Optimism corrected C-statistic	0.72	0.67	0.75
	Adjusted r^{2f}	0.12	<0.01	0.26

^a Estimates based upon ordinal logistic regression with outcome intake in groups (1, 3 and 6 portions per day) and explanatory variables change in biomarker values at 8 weeks.

^b Estimates based upon ordinal logistic regression with outcome intake in groups (2, 4 and 7 portions per day) and explanatory variables change in biomarker values at 12 weeks.

^c Estimates based upon logistic regression with outcome intake in groups (2 and 5 portions per day) and explanatory variables change in biomarker values at 16 weeks.

^d P value for likelihood ratio test comparing each model to model containing all carotenoids and vitamin C.

^e Comparing actual category of intake with predicted category of intake based upon model cut offs, chosen to obtain the correct proportion in each category of intake.

^f Adjusted r^2 calculated using multiple linear regression with actual self-reported FV intake as the outcome and change in biomarker values as explanatory variables.

Legends for figures

Figure 1: Flow chart illustrating recruitment, randomisation and participants' progression through BIOFAV dietary intervention study.

