

1 **Using the Bayesian credible subgroups method to identify populations benefiting from**
2 **treatment: An application to the Look AHEAD trial**

3

4 Anna Coonan^{1*}, John Forbes¹, Patrick Schnell², Joel Smith³

5

6 ¹Graduate Entry Medical School, The University of Limerick, Limerick Ireland.

7 ²Division of Biostatistics, College of Public Health, The Ohio State University, Columbus,
8 OH, USA

9 ³Nuffield Department of Population Health, The University of Oxford, Oxford, UK

10

11 *Corresponding author:

12 Anna.Coonan@ul.ie

13 **Abstract**

14 Traditionally, subgroup analyses are used to assess whether patient characteristics
15 moderate treatment effectiveness with general disregard for issues of multiplicity. Using data
16 from The Action for Health in Diabetes (Look AHEAD) trial in the United States, we aim to
17 identify a subgroup where all of its types of members experience a treatment benefit defined
18 as reducing the likelihood of a major cardiovascular event under an intensive lifestyle and
19 weight-loss intervention. We apply the credible subgroups method to a Bayesian logistic
20 model with a conservative prior that is sceptical of large treatment effect heterogeneity. The
21 covariate profiles for which there is sufficient evidence of treatment benefit are, coarsely,
22 middle-aged women, in poor subjective general health and with moderately to poorly
23 controlled diabetes. There is at least 80% posterior probability that the conditional average
24 treatment effect is positive for all covariate profiles fitting this description, which account for
25 0.5% of trial participants. Conversely, the covariate profiles that are likely to be associated
26 with no benefit are middle aged and older men in excellent subjective general health, with
27 well-controlled diabetes. These profiles apply to less than 2% of trial participants. More
28 information is required to determine treatment benefit or no benefit for the remainder of the
29 trial population.

30

31 **1. Introduction**

32 In general, healthcare providers aim to use information about a person to prevent,
33 diagnose and treat disease, or to promote wellness. This relies on the ability to translate
34 results from trials into clinical practice where the quantity of interest is usually the average
35 treatment effect (ATE). The experimental approach to the estimation of average treatment
36 effects involves random assignment of treatment followed by a comparison of the average

37 outcomes of treated to nontreated individuals that serve as controls [1]. However, ATEs only
38 provide a partial explanation of treatment effects particularly when they vary across
39 experimental units.

40 In randomised controlled trials, average treatment effects tend to be generalised to a
41 wider population and as such the results will be taken to apply equally to all patients in the
42 trial. Participants, however, will vary significantly in baseline predictive and prognostic
43 characteristics, and therefore each patient can report a unique effect with respect to a given
44 treatment. This treatment effect heterogeneity can be masked by the calculation of average
45 effects, leaving potential benefiter and non-benefiters of a given treatment unidentified.
46 Producing results on population averages without recognizing the underlying heterogeneity
47 could lead to welfare losses and increases in healthcare inefficiencies and costs [2, 3].

48 The standard approach for investigating treatment effect heterogeneity is to estimate
49 conditional average treatment effects, or treatment effects among homogeneous groupings of
50 experimental units. These homogenous subgroups are defined by baseline covariates. In order
51 to assess how causal effects vary across subgroups, researchers predict conditional average
52 treatment effects (CATEs) which are average effects conditioned on certain values of
53 observed covariates for which treatment effectiveness varies [4]. Conditional average
54 treatment effects are particularly useful when the researcher wants to explore the
55 heterogeneity of treatment effects in various subgroups defined by a single covariate.

56 Estimating heterogeneous treatment effects, or systematically different outcomes
57 associated with treatment due to moderating patient characteristics, is typically relegated to
58 subgroup analyses that rely on simple treatment-covariate interaction terms [5]. In clinical
59 trials, however, researchers typically aren't concerned with just one subgroup but rather
60 subgroups across a series of standard covariates such as age, ethnicity, and sex. Multiple
61 interaction tests can be misleading as they are often underpowered, suffer from multiplicity

62 errors and do not allow for simultaneous inference [6, 7]. Moreover, subgroups tend to be
63 prespecified to avoid perceptions of data-mining or overfitting the data and therefore restrict
64 the ability of the researcher to test new unforeseen hypotheses [8].

65 There is growing recognition in the field of health economics that more nuanced
66 estimates of patient-specific heterogeneous treatment effects can be estimated and
67 conditioned on a variety of observable and unobservable risk factors. Some methods are
68 derived from the computer science discipline such as machine learning methods where the
69 focus is on predicting patient-specific treatment effects [9-11]. Other methods build on a
70 Bayesian framework where the parameter of interest is unknown but the model allows the
71 researcher to vary the degree of expected treatment effect heterogeneity through *a priori*
72 selection [6, 12, 13]. All of these methods provide insight into the reliability of incorporating
73 heterogeneity in treatment response.

74 Our work builds on a Bayesian framework to classify personalised treatment effects
75 (PTEs), or average outcomes under treatment and control for every covariate point. The PTEs
76 are used to classify covariate points into credible subgroups that allow for inferences about
77 the types of patients expected to benefit or not benefit from an intervention, while accounting
78 for the multiplicity arising from testing a hypothesis at every covariate point. This study is the
79 first major application of the Bayesian credible subgroups method to such a large clinical
80 dataset. Previous research and our own are discussed in the context of a large multicentre
81 randomised controlled trial called the Action for Health in Diabetes (Look AHEAD). In the
82 Look AHEAD trial, insignificant treatment effects for the whole sample and all prespecified
83 subgroups were reported, most likely resulting from latent patient heterogeneity [14].

84 In 2001 the Look AHEAD trial set out to model the long-term effects of an intensive
85 lifestyle intervention (ILI), one specifically aimed at weight loss on cardiovascular morbidity

86 and mortality among patients with type 2 diabetes who were also overweight or obese [14,
87 15]. The ILI included the following: individual and group counselling sessions on diabetes
88 control, aids for calorie management with the option of meal replacement supplements, and
89 supports to increase physical activity. Patients in the control group received diabetes support
90 and education (DSE) counselling only. Prior to the Look AHEAD trial, the effects of lifestyle
91 interventions on cardiovascular morbidity and mortality were mixed at best. Some studies
92 reported no cardiovascular effect and others reported minimal benefits after two decades of
93 follow-up [16-19]. The effects of weight loss on reduced cardiovascular risk factors in the
94 short-run were well-documented, but the Look AHEAD trial set out to identify if significant
95 effects translated to reductions in a primary composite CVD endpoint over a long-term
96 follow-up. In September 2012 the Look AHEAD trial prematurely ended after a futility
97 analysis concluded the weight loss intervention did not reduce the rate of primary
98 cardiovascular events [14]. The median follow-up was 9.6 years. Data are available upon
99 request from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).
100 All participants in Look AHEAD provided written consent to allow their information to be
101 used in future research.

102 A possible explanation for a lack of statistical significance in Look AHEAD was the
103 presence of differing effects for different individuals or patient heterogeneity [14].
104 Secondary analyses were conducted to see if general health benefits resulted from the
105 lifestyle intervention. The documented health benefits range from weight loss, increased
106 physical fitness, better diabetes control, and reduced urinary incontinence [20-22]. A small
107 number of exploratory studies have examined the original CVD endpoint finding positive
108 treatment effects for subgroups defined by subjective general health and diabetes control [23]
109 and percentage weight loss in the first twelve months [20]. Using different methods, these
110 were all attempts to model treatment heterogeneity in Look AHEAD.

111 Behavioural responses and treatment uptake are notoriously heterogeneous with
112 varying degrees of weight loss being reported in the first year [24]. In the instance of the
113 Look AHEAD trial it is likely that patient heterogeneity was able to moderate treatment
114 effectiveness at two entry points: the effect of the intervention on behaviour and uptake, and
115 the effect of behaviour and weight loss on the primary CVD outcome [25]. In Baum et.al.
116 [23] a machine learning algorithm based on decision-trees was used to identify two baseline
117 patient characteristics believed to be instigators behind the trial-wide outcome neutral effects.
118 One covariate was HbA_{1c}, or glycated haemoglobin, which is a measure of diabetes control.
119 The second instigator of treatment effect heterogeneity was a subjective general health
120 indicator derived from the SF-36 questionnaire. These covariates were determined to be the
121 most important variables for defining the range of patient heterogeneity. Defined by HbA_{1c}
122 and subjective general health, two subgroups with positive average treatment effects were
123 identified. However, the HbA_{1c} and subjective general health (SF-36) compositions within
124 these two subgroups were contradictory. One subgroup reported a significant absolute risk
125 reduction defined by the baseline covariates excellent subjective general health (SF-36 \geq 48)
126 and well-controlled diabetes (HbA_{1c} < 6.8%). After participants with an expected adverse
127 treatment response were removed from the analysis, a second subgroup with significant risk
128 reductions was defined by the baseline covariates HbA_{1c} \geq 6.8% or both HbA_{1c} < 6.8% and
129 SF-36 \geq 48. These results led to conclusions stating 85% of Look AHEAD participants
130 averted a cardiovascular event which would be a large overestimation of treatment benefit
131 compared to previous Look AHEAD studies [14, 20, 25, 26]. The mixed composition of these
132 subgroups support previous criticisms stating too much emphasis is placed on the coefficients
133 produced by machine learning algorithms and that they are not designed to collate patients
134 into subgroups. Rather, machine learning methods are better suited for examining unit-level
135 heterogeneity [27].

136 Machine learning methods are useful for detecting and analysing heterogeneity, but
137 they do not directly address the question of primary interest to the current study: specifically
138 whom does treatment benefit? Previously-applied machine learning methods do not adjust for
139 multiplicity or provide simultaneous inferences, defined as a situation where all types of
140 individuals within a subpopulation satisfy a common statement [6]. The most relevant
141 statement here is that all covariate points in the subgroup correspond to a positive PTE. We
142 use a Bayesian credible subgroups method for simultaneous inference that through careful
143 choice of priors on treatment-covariate interactions, allows us to tune our *a priori* belief in
144 the strength of treatment effect heterogeneity. Our model includes a conservative prior that is
145 sceptical of large treatment effect heterogeneity but still allows for strong evidence from the
146 posterior distribution to overwhelm the prior information.

147 The inferential goal of this analysis is different from previous secondary Look
148 AHEAD studies that confirm the overall treatment effect is neutral and any subsequently
149 identified subgroups are defined by estimated average treatment effects. These estimates
150 continue to apply assumptions of treatment effect homogeneity within all levels of selected
151 covariates. Instead, we aim to capitalise on inherent patient heterogeneity in order to identify
152 the subgroup consisting of only types of patients expected to benefit, not a subgroup of any
153 types patients that together yield a positive average effect. We explore the mechanisms
154 behind the statistical insignificance of the lifestyle intervention. This is done by examining
155 the possibility of a treatment-benefiting subgroup defined by the baseline covariates sex, age,
156 blood glucose percentage (HbA_{1c}) and a measure of subjective general health (SF-36).

157 This article is structured as follows: In Section 2 we present the methodology and
158 selection of priors. Section 3 describes the data from the Look AHEAD and defines the
159 descriptive statistics. Section 4 presents the results from the credible subgroups analysis with

160 varying credible levels. Section 5 concludes with a discussion and suggestions for future
161 research.

162

163 **2. Methodology**

164 Data provided by the NIDDK repository was fully anonymised before we accessed
165 them. Our research received ethical approval from University of Limerick application number
166 2017_12_19_EHS. All participants in Look AHEAD provided written consent to allow their
167 information to be used in future research.

168 This analysis uses the Bayesian credible subgroups method to construct two bounding
169 subgroups for a benefiting subgroup B that is unknown. The benefiting subgroup is defined to
170 consist of all covariate profiles with a conditional average treatment effect above a threshold
171 set to zero ($\delta = 0$), intended to capture all positive treatment effects regardless of magnitude.
172 Inferences about the subgroup B are made by constructing bounding subgroups, D and S,
173 similar to the procedure for making inferences about a scalar quantity by constructing interval
174 estimates. The bounding subgroups are constructed with the goal that one, the exclusive
175 credible subgroup (D), contains only covariate profiles with a positive treatment effect. The
176 other, the inclusive credible subgroup (S), is intended to contain all covariate profiles with a
177 positive treatment effect. In the construction of these bounds, treatment effects are estimated
178 and tested at all points throughout the covariate space, thereby introducing multiple testing
179 issues. However, by estimating a subgroup where all of its members satisfy some criteria,
180 such as having a treatment effect above some threshold of significance, this statement of
181 simultaneity accounts for multiplicity due to performing a hypothesis test at each covariate
182 point. The result is a decreased risk of conducting a family wise type I error.

183 The method of credible subgroups is a two-step regression classification method.
184 Parameters of a linear or non-linear regression model are estimated, and a simultaneous
185 credible band of some level around the regression surface is constructed. Covariate points
186 with a lower bound above some threshold form the exclusive credible subgroup (D), and
187 those where the upper bound is greater than the threshold become the inclusive credible
188 subgroup (S). As an example of the first step, estimation of a standard linear model includes
189 predictive and prognostic covariates generally described as follows:

$$190 \qquad \qquad \qquad E[Y_i|x_i, z_i, t_i] = x_i'\beta + t_i z_i'\gamma \qquad (1)$$

191 Y is the response, x and z are vectors of prognostic and predictive covariates where predictive
192 covariates are those interacted with treatment allocation. β and γ are the effect parameters
193 and $t \in (0,1)$ the treatment allocation. Additionally, the model includes a conservative prior
194 on γ that represents the belief that treatment-covariate interactions are unlikely by shrinking
195 coefficients towards zero, but that still allows for strong evidence of treatment effect
196 heterogeneity from the posterior distribution to overrule this. The second step, classification,
197 comes from the posterior distribution of patients' personalised treatment effects (PTEs).
198 These are used to classify all combinations of covariates in the covariate space according to a
199 credible subgroup pair: the exclusive credible subgroup (D) and the inclusive credible
200 subgroup (S). PTEs provide a clinically useful interpretation of the expected difference in
201 outcomes for patients that share the same baseline characteristics.

202 In our analysis the covariate space is comprised of age, sex, subjective general health
203 measured by the SF-36 general health questionnaire, and diabetes control measured a
204 glycated haemoglobin percentage (HbA_{1c}). All of the covariates included in this paper are
205 standard medical and demographic variables included in studies examining behavioural
206 treatments for type 2 diabetes and the confounding effects of obesity [14, 17, 18, 20, 23, 28].

207 The model of the personalised treatment effect is the difference in the expectation of the
208 outcome (y) conditional on the predictive covariates (z), as follows:

$$209 \quad \Delta(z) \equiv E[Y|x, z; t = 1] - E[Y|x, z; t = 0] \quad (2)$$

210 From the posterior distribution of personalised treatment effects the covariate space is divided
211 into three parts based on a decision rule used to define the credible subgroup pair (D,S). This
212 is comprised of the covariate profiles with lower bounds for treatment effects above some
213 threshold ($\delta > 0$) that define subgroup D, those in a complement subgroup S^c with upper
214 bounds for treatment effects at most equal to the threshold ($\delta = 0$) and an uncertainty region
215 $S \setminus D$ (S remove D) where more information is required for classification. The trichotomy of
216 the covariate space is illustrated in Fig 1. The defining property of the credible subgroups is
217 as follows:

$$218 \quad P(D \subseteq B \subseteq S | \text{data}) \geq 1 - \alpha, \quad (3)$$

219 where D and S are subsets in the covariate space with a bound on the posterior probability
220 that these two sets bracket the benefiting subset B. Formally, D is the exclusive credible
221 subgroup because the posterior probability that D contains only z for which $\Delta(z) > \delta$, and
222 simultaneously, S is the inclusive subgroup that contains all z such that $\Delta(z) > \delta$, is at least
223 $1 - \alpha$. The credible level of the subgroup pair ($1 - \alpha$) is also an approximate lower bound
224 on the frequentist probability that $D \subseteq B \subseteq S$. In this study, D is the reported estimate with
225 posterior probability $1 - \alpha$ that the reported subset characterises exclusively covariate
226 combinations with an expected treatment effect above the threshold ($\delta = 0$).

227 The objective of this analysis is to identify a credible subgroup pair such that, at a
228 credible level of at least 80%, D represents a subgroup that contains *only* types of patients
229 with a positive treatment effect and S represents a subgroup that contains *all* types of patients
230 with a positive treatment effect. Defining an exclusive credible subgroup (D) where all

231 members have a treatment effect above some threshold satisfies properties of simultaneous
232 inference and adjusts for multiplicity [6]. The credible subgroups method allows for
233 inferences that state it is with high probability that the personalised treatment effect ($\Delta(z)$) is
234 positive at all covariate points (z) in the exclusive credible subgroup (D).

235

236 **Fig_1. Trichotomy of Predictive Covariate Space.** Visualisation of the trichotomy of the
237 covariate space based on the credible subgroup pair (D,S) relative to the true benefiting
238 subgroup (B), adapted from Schnell et.al. [6]

239

240 The primary endpoint in Look AHEAD is a binary composite variable measuring the
241 presence or absence of at least one CVD outcome. We fit a Bayesian logistic regression
242 model with the endpoint likelihood $Y_i|\eta_i \sim \text{Bernoulli}(\text{logit}^{-1}, \eta_i)$ where $\eta_i = x_i'\beta + t_i z_i'\gamma$.
243 We use the vague prior $\beta \sim \text{Normal}(0, 10^3)$ for prognostic effects and a conservative prior
244 $\gamma \sim \text{Normal}(0, 1)$ for predictive effects. In other words, the prognostic effect variances are set
245 to 1000 and those corresponding to predictive effect variances set to 1. The benefiting
246 subgroup is one for which the ILI treatment is superior to the diabetes support and education
247 (DSE) control. This is in line with current clinical recommendations and the belief that
248 physical activity and controlled caloric intake is beneficial for patients with type 2 diabetes
249 who may be overweight or obese. We set the threshold of a log odds ratio of primary CVD
250 occurrence to $-\delta_1 > 0$. The sign is switched for δ because we want reductions in risk of the
251 primary CVD endpoint. Due to the insignificant results reported in the original trial, we wish
252 to capture treatment benefits of any magnitude in order to avoid automatically eliminating
253 covariate profiles that may not achieve a pre-defined level of interest. The model is fit using
254 the NIMBLE package [29] in the R software [30] with 10,000 Gibbs sampler iterations after

255 1000 burn-in iterations. Convergence appears to be mixing well and near-immediate from
256 autocorrelation plots for Markov Chains. Our analysis is conducted using only complete
257 cases eliminating 8 observations which renders a sample size of 4893.

258

259 **3. Data**

260 Look AHEAD was a multicentre randomised controlled trial that tested the effect of
261 an intensive lifestyle intervention focused on diet and exercise on long-term cardiovascular
262 disease outcomes among patients with type 2 diabetes who were overweight or obese [14,
263 15]. Between August 2001 and April 2004 the trial recruited 5145 participants to undergo
264 randomisation. 2570 were assigned to the intensive lifestyle intervention (ILI) group and
265 2575 were assigned to the diabetes support and education control (DSE) group. The
266 intervention was modelled on group behaviour programs tailored for obese and overweight
267 patients through individual and group counselling sessions. These sessions continued for an
268 initial 6 months and decreased in frequency throughout a 12-month follow-up. Intervention
269 design was tailored for a weight-loss goal of losing 10% initial body weight in the first year
270 and a total weight-loss of 7% (or more) initial body weight by trial's end. This would be
271 achieved by limiting caloric intake with the option of meal replacement supplements and
272 increasing physical fitness through weekly exercise regimes. The diabetes support and
273 education (DSE) control group provided nutrition and exercise training for a period of four
274 years after randomisation.

275 The primary outcome was a composite indicator of CVD morbidity and mortality
276 consisting of CVD death, non-fatal myocardial infarction (MI), non-fatal stroke, and
277 hospitalization for angina [14]. The number of CVD events was a concern throughout the
278 trial resulting in the fourth CVD indicator, hospitalisation from angina, being added to the

279 primary outcome two years into the trial [31]. The event rate of the CVD endpoint ended up
280 being less than half the projected rate in the DSE control group, resulting in a very low
281 overall rate of events [14]. Insignificant treatment effects were reported for secondary CVD
282 composite measures and for all prespecified subgroups including age, sex and ethnicity.

283 Table 1 presents the descriptive statistics for patients in the treated and control groups
284 of the remaining 4901 trial participants after 244 were omitted early in the trial for consent
285 reasons. The characteristics of the patients in the two groups were similar at baseline. The
286 ILI treatment group was on average 58.8 years old, had a BMI of 35.9, weighed 100.9kg and
287 was 41% men. The DSE control group was on average 59.1 years old, had a BMI of 36.0,
288 weighed 101.2kg and was 41% men. For the entire sample 75% of patients had moderately or
289 poorly controlled diabetes defined as having an HbA_{1c} percentage greater than or equal to
290 6.5.

291

292 **Table 1. Characteristics of Patients at Baseline**

	Intensive Lifestyle Intervention (ILI) (n = 2448)	Diabetes Support and Education (DSE) (n = 2453)	Entire Trial Sample (n = 4901)
Age (years)	58.8 (6.7)	59.1 (6.8)	58.9 (6.8)
Male	1014 (41%)	1016 (41%)	2030 (41%)
History of cardiovascular disease	355 (15%)	335 (14%)	690 (14%)
Diabetes duration (years)	6.7 (6.5)	6.7 (6.3)	6.7 (6.4)
Insulin use	362 (15%)	385 (16%)	747 (15%)
Weight (kg)	100.9 (19.6)	101.2 (18.8)	101.1 (19.2)
BMI (kg/m ²)	35.9 (6.0)	36.0 (5.7)	36.0 (5.9)

HbA _{1c} (%)	7.2 (1.1)	7.3 (1.2)	7.3 (1.2)
HbA _{1c} (>=6.5%)	1798 (73%)	1867 (76%)	3665 (75%)
Systolic blood pressure (mm Hg)	128.5 (17.2)	129.8 (17.1)	129.1 (17.2)
SF-36 General Health (GH) Questionnaire	47.2 (9.1)	47.3 (8.7)	47.2 (8.9)

293 Data are in mean (SD) or n (%)

294

295 4. Results

296 The credible subgroups method is used to divide the covariate space into a trichotomy
297 which will identify the profiles of patients expected to benefit from treatment, those expected
298 not to benefit from treatment and those for whom more information is needed. The Look
299 AHEAD trial is used in our analysis and includes an experimental intensive lifestyle
300 intervention (ILI) versus a diabetes support and education (DSE) control. The covariates are
301 believed to be both prognostic and predictive and are defined by baseline measurements for
302 subjective general health measured by the SF-36 questionnaire, age, sex and diabetes control
303 (HbA_{1c}). The endpoint of interest is a binary composite indicator measuring the presence of
304 four possible macro cardiovascular events.

305 The first stage of the credible subgroups method is the estimation of a Bayesian
306 logistic regression with vague priors on prognostic effects and conservative priors on
307 predictive effects. From this stage our objective is to discuss treatment-covariate interactions
308 at the nominal credible level of 95% for which standard subgroup analyses were not powered
309 to detect. The model is fit with 10,000 Gibbs sampler iterations after 1000 burn-in iterations.
310 The posterior mean, standard deviations and credible bounds of effect parameters are
311 reported in Table 2. Significant effects are reported for the treatment-control interaction and
312 individual covariates general health and HbA_{1c}. These results are not adjusted for multiplicity

313 and are considered incomplete because a lack of significance on the age and sex interaction
314 terms does not imply a homogeneous treatment effect across all levels of these covariates.
315 The estimates in Table 2 do not identify the covariate profiles of patients explicitly expected
316 to benefit and therefore do not directly address our research question. The interpretation that
317 we wish to avoid is that the treatment only interacts with general health and HbA_{1c} and that
318 any potential treatment effect exists across all levels of the other covariates.

319

320

321 **Table 2. Posterior Summaries of Bayesian Logistic Parameters**

	Posterior Mean	Posterior Standard Deviation	95% Credible Bounds
Intercept	-2.41*	0.14	-2.18 to -2.63
General Health	-0.07	0.06	0.03 to -0.17
Age	0.43*	0.06	0.34 to 0.53
Sex	0.82*	0.11	0.64 to 0.99
HbA _{1c}	0.39*	0.13	0.17 to 0.61
Treatment	-0.22	0.18	-0.52 to 0.07
General Health × Treatment	-0.19*	0.08	-0.05 to -0.32
Age × Treatment	-0.01	0.08	-0.15 to 0.13
Sex × Treatment	-0.07	0.16	-0.33 to 0.18
HbA _{1c} × Treatment	0.32*	0.18	0.02 to 0.61

322

323 Continuous covariates are standardized. Estimates identified as significant according to a
324 95% credible level are denoted by *.

325

326 Fig 2 presents the covariate space which constitutes the exclusive credible subgroup D
327 at the 80% credible level. Credible subgroup membership was determined using the
328 credsubs package [32]. This analysis uses a Bayesian logistic model to estimate posterior
329 mean treatment effects from endpoint likelihoods that follow a Bernoulli distribution.
330 Ellipsoidal contours of the standard errors are centred on a majority of the data points which

331 induce the curved boundaries of the credible subgroup [6]. The covariate profiles classified to
332 the exclusive credible subgroup D are middle-aged women in poor subjective general health
333 with moderately to poorly controlled diabetes. There is at least 80% posterior probability that
334 the conditional average treatment effect is beneficial ($\delta < 0$) for all types of patients with
335 covariate points in D, accounting for multiplicity due to testing a hypothesis for each
336 covariate point, and thus promoting the lifestyle intervention as an appropriate treatment for
337 this subgroup. Conversely, the S-complement subgroup identifies the covariate profiles of
338 those who may not benefit, represented by either a null or adverse treatment effect ($\delta \geq 0$).
339 The covariate profiles of patients in this region are middle-aged and older men, in excellent
340 subjective general health with well-controlled diabetes. Therefore, the lifestyle intervention is
341 not to be considered an appropriate treatment for members of this subgroup. A figure of
342 similar threshold and credible level with a vague prior on predictive and prognostic effects is
343 included in the Supplement S_Fig1. This figure supports the conclusions presented in Table
344 2.

345

346 **Fig 2. Trichotomy of Baseline Covariates at 80% Credible Level.**

347

348 The funnel plot in Fig 3 shows the predicted treatment effects across the entire
349 covariate space and their coloured classification according to the credible subgroups pair
350 (D,S). A majority of covariate space is classified as the $S \setminus D$ region where more information
351 is required to classify points as being members of the exclusive credible subgroup. The
352 funnel plot displays the proportion of covariates that are members of the S-complement
353 subgroup and have a treatment effect at most equal to zero. This graph identifies the minority
354 of covariate profiles with sufficient evidence of an expected treatment benefit represented by

355 the green area. The minority of points classified to the exclusive credible subgroup supports
356 the idea that the heterogeneity of patients with these covariate profiles could have been lost in
357 the original trial-wide average estimates. When applying the trichotomy of the covariate
358 space to 4893 observations in the Look AHEAD trial, there is sufficient evidence of benefit
359 for less than 0.5% of the Look AHEAD population. Likewise, there is sufficient evidence of
360 an expected null or adverse effect for less than 2% of the trial population.

361

362 **Fig 3. Exclusive (Green) and Inclusive (Yellow) Credible Subsets at 80% Credible**
363 **Level**

364

365 Fig 4 shows the contours of the maximum credible level that allows assignment of a
366 covariate point to the exclusive credible subgroup. Contours with negative values indicate
367 assignment to the complement of the inclusive credible subgroup at the credible level of the
368 modulus. The 80% contour aligns with the previously identified boundary of the exclusive
369 credible subgroup shown in Fig 2. This figure displays the trade-off that exists in statistical
370 reliability versus restrictiveness of subgroup composition. As the maximum credible level is
371 relaxed, a greater portion of the covariate space is classified in either the exclusive credible
372 subgroup or the complement of the inclusive credible subgroup.

373

374 **Fig 4. Maximum Credible Level Contours for Exclusive Credible Subgroup Inclusion**

375

376 In Fig 5 we set the treatment effect threshold to zero but relax the credible levels from
377 80% to 50%. A figure illustrating the trichotomy of the covariate space at a 65% credible

378 level is included in the Supplement S_Fig2. As expected, when the credible level is relaxed,
379 the proportion of covariate points allocated to the exclusive credible subgroup (D) and S-
380 complement subgroup (S^c) increase. As the credible level is relaxed, the covariate
381 composition of the exclusive credible subgroup widens to include both men and women for a
382 wider range of ages. The two most important covariates for differentiating patient benefitters
383 from non-benefitters are having a low HbA_{1c} percentage and having good subjective general
384 health. The similarity between these two covariates is to be expected as good control over
385 one's diabetes, which is a physical health indicator, could translate to patients' positive
386 subjective evaluations of their own health. This conclusion supports previous research that
387 has found HbA_{1c} and subjective general health (SF-36) are two of the most important
388 covariates for determining treatment effect heterogeneity in the Look AHEAD trial [23]. The
389 HbA_{1c} composition of the exclusive and inclusive credible subgroups is largely consistent
390 across sex, age and general health covariates.

391 At the 80% credible level, the application of the covariate profiles of the exclusive
392 credible subgroup, suggests evidence of treatment benefit for less than 0.5% of the Look
393 AHEAD population. Applying the covariates that constitute the S-complement subgroup
394 suggests evidence of a null or adverse treatment effect for 2% of the Look AHEAD
395 population. At the more lenient 50% credible level, the application of the covariate profiles of
396 the exclusive credible subgroup would render a subpopulation of less than 10% of the Look
397 AHEAD trial. While the true benefiting population may be larger, the expansive inconclusive
398 region across two credible levels does not provide supporting evidence of this. Our results
399 across all credible levels differ from previous controversial findings that suggested as much
400 as 85% of the Look AHEAD population benefitted from the ILI based on positive average
401 estimates [23].

402

403 **Fig 5. Trichotomy of Baseline Covariates at 50% Credible Level.**

404

405 **5. Discussion**

406 In this analysis of the Look AHEAD trial, we identified a subgroup that benefited
407 from an intense behavioural lifestyle intervention in terms of reduced cardiovascular risk.
408 The aim of the trial was to model the effects of a lifestyle intervention that focused on
409 weight-loss and physical activity on the long-term cardiovascular outcomes of patients with
410 type 2 diabetes who were overweight or obese. After the original trial reported no between
411 group differences in the occurrence of the primary CVD endpoint for patients randomised to
412 the treatment or control, commentators suggested this was the result of latent patient
413 heterogeneity. Recent developments in treatment effect estimation call for a move away from
414 *ad hoc* subgroup analyses and encourage the development of new methodologies designed to
415 capture the richness that patient heterogeneity provides. By fitting a Bayesian logistic
416 regression model, we were able to include prior information of the degree of treatment effect
417 heterogeneity and still allow for strong evidence from the posterior distribution to overrule
418 any assumptions. Using the credible subgroups method, we classified covariate profiles
419 according to a credible subgroup pair that allowed for the identification of a subgroup in
420 which all types of patients exhibit a positive treatment effect. This subgroup was coarsely
421 comprised of the following covariate profiles: middle-aged women in poor subjective general
422 health with moderately to poorly controlled diabetes ($HbA_{1c} \geq 6.5\%$). There exists at least
423 80% posterior probability that the treatment effect is positive for patients with these covariate
424 profiles located within the exclusive credible subgroup (D), accounting for multiplicity due to
425 testing a hypothesis for each covariate point. A second subgroup was identified as the S-
426 complement where its members exhibited a null or adverse treatment effect. The covariate

427 profiles of this group were middle-aged and older men, in excellent subjective general health
428 with well-controlled diabetes ($HbA_{1c} < 6.5\%$). Applying the covariate profiles of the
429 exclusive credible subgroup to the Look AHEAD participants rendered an eligible
430 subpopulation of less than 0.5%. Members of this subpopulation would be expected to
431 experience a positive treatment effect.

432 The covariate composition of the exclusive credible subgroup is consistent with
433 studies from similar trials that reported women and those in poor health experienced the
434 greatest reduction in CVD mortality after participating in a weight loss intervention [17, 33,
435 34]. Additionally, the identification of a S-complement group supports previous findings that
436 a selection of patients can have null or adverse cardiovascular responses to weight loss
437 interventions [20, 34]. In our analysis, as the credible levels were relaxed, the covariates poor
438 subjective general health and poorly controlled diabetes were consistently the ones to have an
439 expected treatment benefit while the covariates good subjective general health with well-
440 controlled diabetes had an expected null or adverse treatment effect.

441 The innovation of this paper firstly comes from the inferences made possible by the
442 credible subgroups method: that there is a high posterior probability that all types of patients
443 within the exclusive credible subgroup D experience a positive treatment effect. This
444 addresses a different research question to previous publications that have explored patient
445 heterogeneity but relied on sample average calculations. Our findings compared to those
446 previous demonstrate that it is possible to calculate average estimates that do not capture the
447 responses of minority subgroups with treatment effects that deviate from the mean. Secondly,
448 the credible subgroups method allows us to make statements of simultaneity where all
449 subgroup members experience a treatment effect above some threshold that adjusts for issues
450 of multiplicity due to hypothesis tests in multiple subgroups. This helps to avoid the

451 perception of data dredging which underpins prevalent scepticism of standard subgroup
452 analyses.

453 Previous research has used exploratory endpoints and *ad hoc* subgroup analysis to
454 identify significant treatment effects reported in secondary studies. Using the original CVD
455 endpoint, our analysis shows that as the credible level is relaxed, the range of the covariate
456 profiles for exclusive subgroup membership expands. There exists a trade-off between
457 statistical confidence and subgroups too narrowly defined to represent any application in
458 medical care. As the credible level is relaxed from 80% to 50%, the proportion of covariate
459 points classified to the exclusive credible subgroup increases. Our research identifies a
460 minority subgroup consisting of covariate profiles with an expected treatment benefit. Based
461 on an 80% credible level, these covariate profiles are used to classify less than 0.5% of the
462 Look AHEAD trial who would simultaneously be expected to experience a positive treatment
463 effect. This contradicts previous research stating 85% of Look AHEAD participants did
464 benefit from treatment as identified by positive average estimates [23].

465 The findings of this secondary analysis are not meant to overshadow the conclusions
466 from the original Look AHEAD trial. Rather, we demonstrate an alternative method of
467 exploring treatment effect heterogeneity that does not suffer from the limitations of standard
468 subgroup analyses. The credible subgroups method defines subgroups, based on a number of
469 baseline covariates, which consist of all the types of patients expected to benefit or not
470 benefit from a given treatment while accounting for multiplicity. This research applies the
471 credible subgroups method to the Look AHEAD trial to illustrate the trichotomy of the
472 covariate space and identifies a large number of covariate points located in an uncertainty
473 region. This combined with the restrictive covariate profiles of those who would be classified
474 to the exclusive credible subgroup, explains how so few patients with an expected treatment
475 benefit could have been lost in trial-wide average estimates.

476 This analysis of the Look AHEAD trial has some limitations. At first glance the
477 output from the credible subgroups method may not be as straightforward to interpret
478 compared to standard subgroup analyses, but the fundamental objective of identifying
479 individual covariate profiles of patients expected to benefit cannot be achieved through
480 simple average calculations. Secondly, the credible subgroups method is not designed for
481 variable selection. Instead, the credible subgroups method requires a list of covariates
482 believed to have predictive value when interacted with treatment which requires *a priori*
483 knowledge of intervening factors. Thirdly, the credible subgroups method computationally
484 struggles with increasing covariates due to a reduction in statistical power and the inclusion
485 of many continuous covariates [6]. The possible biases produced by these limitations are
486 similar to those found in standard subgroup analyses. Regarding the biological plausibility of
487 the chosen covariates in particular, all are standard medical and demographic variables in
488 studies examining behavioural treatments for type 2 diabetes and the confounding effects of
489 obesity [14, 17, 18, 20, 23, 28]. We conclude that covariates previously identified as being
490 important for exploring treatment heterogeneity may only explain a small portion of the
491 entire covariate space that is useful for identifying positive treatment effects. In fact, variable
492 selection itself introduces multiplicity problems, and accounting for this source of
493 multiplicity as well would necessitate even more cautious conclusions. In particular, the
494 inconclusive region ($D \setminus S$) would need to be enlarged for any given credible level. Some
495 study of variable selection via Stochastic Search Variable Selection (SSVS) within the
496 regression step of the credible subgroups approach has been done [35], which shows this
497 effect in the sense that $D \setminus S$ is larger when not conditioning on a selected model. As this
498 relates to the comparison with Baum et.al. [23], we have shown that accounting for only
499 some but not all sources of multiplicity does not allow as optimistic conclusions as the
500 previous analysis. This result would be made even more extreme by accounting for additional

501 sources of multiplicity. Finally, point estimates for treatment effects in the selected subgroups
502 D and S-complement are biased away from zero due to selection bias. Thus conclusions
503 drawn from the credible subgroups method should be limited to hypothesis testing, and plots
504 such as Figure 3 should be interpreted only as diagnostics and not as post-selection effect
505 estimates.

506 There exists much scope for future research in the synergy of the credible subgroups
507 method into adaptive enrichment trial design. The objectives of adaptive trial design and the
508 credible subgroups method are the same: to increase the likelihood of identifying a true
509 subgroup of patients expected to benefit from a given therapy [36]. Adaptive trial design
510 includes prospectively planned analyses that can change the course of an on-going trial based
511 on accumulating the data of subjects within the trial [37]. Adaptive trial design stands to
512 benefit from the credible subgroups method if used throughout interim analyses and at early
513 stage enrollee selection. Sequential trial design permits stopping early for efficacy based on
514 interim analyses [38, 39], which could be defined by the trichotomy of the covariate space
515 using the credible subgroups method. Overwhelming evidence of treatment benefit or harm
516 would allow for early termination while the trial would continue to gain information about
517 the covariate space where classifications are uncertain. Additionally, the credible subgroups
518 method could be used in an adaptive population-enrichment design where biomarkers are
519 identified that would predict response or lack of response to a given therapy [40]. This
520 information could be used in confirmatory trials to limit entry of patients to stage 3
521 randomisation [41]. In this research, biomarkers for diabetes and obesity that are related to
522 the effectiveness of the lifestyle intervention would be of particular interest. Future
523 incorporation of the credible subgroups method into adaptive enrichment trial design can
524 magnify the potential treatment effect, reduce costs and increase trial efficiency.

525

526 **6. Acknowledgements**

527 The work of AC and JF was funded by the Health Research Board (grant number HRB-RL-
528 2013-11). The work of JS was funded by the National Institute for Health Research
529 Biomedical Research Centre at the University of Oxford (grant number NIHR-BRC-1215-
530 20008).

531 Data are available upon request from the National Institute of Diabetes and Digestive and
532 Kidney Diseases (NIDDK) at <https://www.niddk.nih.gov/>.

533

534 **6. Supplement**

535 **S_Fig1. Trichotomy of Baseline Covariates at 80% Credible Level with Vague Prior on**
536 **Predictive and Prognostic Effects.**

537 **S_Fig2. Trichotomy of Baseline Covariates at 65% Credible Level**

538 **References**

- 539 1. Cameron AC, Trivedi PK. *Microeconometrics: methods and applications*: Cambridge
540 university press; 2005.
- 541 2. Basu A. ESTIMATING PERSON-CENTERED TREATMENT (PeT) EFFECTS USING INSTRUMENTAL
542 VARIABLES: AN APPLICATION TO EVALUATING PROSTATE CANCER TREATMENTS. *J Appl Econ*
543 (Chichester Engl). 2014;29(4):671-91.
- 544 3. Basu A, Jena AB, Philipson TJ. The impact of comparative effectiveness research on health
545 and health care spending. *J Health Econ*. 2011;30(4):695-706.
- 546 4. Heckman J. Instrumental variables: A study of implicit behavioral assumptions used in
547 making program evaluations. *J Human Res*. 1997:441-62.
- 548 5. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and
549 baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med*.
550 2002;21(19):2917-30.
- 551 6. Schnell P, Tang Q, Walter OW, Carlin BP. A Bayesian credible subgroups approach to
552 identifying patient subgroups with positive treatment effects. *Biometrics*. 2016;72(4):1026-36.
- 553 7. Ruberg SJ, Lei C, Yanping W. The mean does not mean as much anymore: finding sub-groups
554 for tailored therapeutics. *Clin Trials*. 2010;7(5):574-83.
- 555 8. Olken BA. Promises and perils of pre-analysis plans. *J Econ Perspect*. 2015;29(3):61-80.
- 556 9. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: Springer series in*
557 *statistics* New York, NY, USA.; 2001.
- 558 10. Athey S, Imbens GW. Machine learning methods for estimating heterogeneous causal
559 effects. *Stat*. 2015;1050(5).
- 560 11. Imai K, Ratkovic M. ESTIMATING TREATMENT EFFECT HETEROGENEITY IN RANDOMIZED
561 PROGRAM EVALUATION. *Ann Appl Stat*. 2013;7(1):443-70.
- 562 12. Green DP, Kern HL. Modeling heterogeneous treatment effects in survey experiments with
563 Bayesian additive regression trees. *Public Opin Q*. 2012;76(3):491-511.
- 564 13. Simon R. Bayesian subset analysis: application to studying treatment-by-gender interactions.
565 *Stat Med*. 2002;21(19):2909-16.
- 566 14. Look AHEAD Research Group. Cardiovascular effects of intensive lifestyle intervention in
567 type 2 diabetes. *N Engl J Med*. 2013;369(2):145-54.
- 568 15. Look AHEAD Research Group. Look AHEAD (Action for Health in Diabetes): design and
569 methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2
570 diabetes. *Control Clin Trials*. 2003;24(5):610-28.
- 571 16. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle interventions for
572 patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. *Ann Intern Med*.
573 2013;159(8):543-51.
- 574 17. Li G, Zhang P, Wang J, Gregg EW, Yang W, Gong Q, et al. The long-term effect of lifestyle
575 interventions to prevent diabetes in the China Da Qing Diabetes Prevention Study: a 20-year follow-
576 up study. *Lancet*. 2008;371(9626):1783-9.
- 577 18. Uusitupa M, Peltonen M, Lindström J, Aunola S, Ilanne-Parikka P, Keinänen-Kiukaanniemi S,
578 et al. Ten-year mortality and cardiovascular morbidity in the Finnish Diabetes Prevention Study—
579 secondary analysis of the randomized trial. *PLoS One*. 2009;4(5):e5656.
- 580 19. Li G, Zhang P, Wang J, An Y, Gong Q, Gregg EW, et al. Cardiovascular mortality, all-cause
581 mortality, and diabetes incidence after lifestyle intervention for people with impaired glucose
582 tolerance in the Da Qing Diabetes Prevention Study: a 23-year follow-up study. *Lancet Diabetes*
583 *Endocrinol*. 2014;2(6):474-80.
- 584 20. Look AHEAD Study Group, Gregg EW. Association of the Magnitude of Weight Loss and
585 Physical Fitness Change on Long-term CVD outcomes: The Look AHEAD Study. *Lancet Diabetes*
586 *Endocrinol*. 2016;4(11):913.

- 587 21. Look AHEAD Research Group. Reduction in weight and cardiovascular disease risk factors in
588 individuals with type 2 diabetes: one-year results of the look AHEAD trial. *Diabetes Care*. 2007.
589 22. Breyer BN, Phelan S, Hogan PE, Rosen RC, Kitabchi AE, Wing RR, et al. Intensive lifestyle
590 intervention reduces urinary incontinence in overweight/obese men with type 2 diabetes: results
591 from the Look AHEAD trial. *J Urol*. 2014;192(1):144-9.
592 23. Baum A, Scarpa J, Bruzelius E, Tamler R, Basu S, Faghmous J. Targeting weight loss
593 interventions to reduce cardiovascular complications of type 2 diabetes: a machine learning-based
594 post-hoc analysis of heterogeneous treatment effects in the Look AHEAD trial. *Lancet Diabetes
595 Endocrinol*. 2017;5(10):808-15.
596 24. Wadden TA, Butryn ML, Hong PS, Tsai AG. Behavioral treatment of obesity in patients
597 encountered in primary care settings: a systematic review. *JAMA*. 2014;312(17):1779-91.
598 25. Gregg EW, Wing R. Looking again at the Look AHEAD study. *Lancet Diabetes Endocrinol*.
599 2017;5(10):763-4.
600 26. Pi-Sunyer X. The Look AHEAD trial: a review and discussion of its outcomes. *Current nutrition
601 reports*. 2014;3(4):387-91.
602 27. Chernozhukov V, Demirer M, Duflo E, Fernandez-Val I. Generic machine learning inference
603 on heterogenous treatment effects in randomized experiments. *National Bureau of Economic
604 Research*; 2018.
605 28. Li G, Zhang P, Wang J, An Y, Gong Q, Gregg EW, et al. Cardiovascular mortality, all-cause
606 mortality, and diabetes incidence after lifestyle intervention for people with impaired glucose
607 tolerance in the Da Qing Diabetes Prevention Study: a 23-year follow-up study. *Lancet Diabetes
608 Endocrinol*. 2014;2(6):474-80.
609 29. Nimble Development Team. NIMBLE: An R Package for programing with BUGS models. 0.4
610 ed2015.
611 30. R Core Team. R. A Language and Environment for Statistical Computing. Vienna, Austria R
612 Foundation for Statistical Computing; 2019.
613 31. Brancati FL, Evans M, Furberg CD, Geller N, Haffner S, Kahn SE, et al. Midcourse correction to
614 a clinical trial when the event rate is underestimated: the Look AHEAD (Action for Health in
615 Diabetes) Study. *Clin Trials*. 2012;9(1):113-24.
616 32. Schnell P, Carlin B. credsubs: Credible Subsets R package version 1.0. 2017.
617 33. Unick JL, Beavers D, Jakicic JM, Kitabchi AE, Knowler WC, Wadden TA, et al. Effectiveness of
618 lifestyle interventions for individuals with severe obesity and type 2 diabetes: results from the Look
619 AHEAD trial. *Diabetes Care*. 2011:DC_110874.
620 34. Harrington M, Gibson S, Cottrell RC. A review and meta-analysis of the effect of weight loss
621 on all-cause mortality risk. *Nutr Res Rev*. 2009;22(1):93-108.
622 35. Schnell P. Credible subgroups: identifying the population that benefits from treatment.
623 University of Minnesota Digital Conservancy: <http://hdl.handle.net/11299/188890>; 2017.
624 36. Bauer P, Bretz F, Dragalin V, König F, Wassmer G. Twenty-five years of confirmatory adaptive
625 designs: opportunities and pitfalls. *Stat Med*. 2016;35(3):325-47.
626 37. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for
627 Biologics Evaluation and Research (CBER). *Adaptive Designs for Clinical Trials of Drugs and Biologics*.
628 Silver Spring, MD: Food and Drug Administration 2018.
629 38. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika*.
630 1977;64(2):191-9.
631 39. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979:549-
632 56.
633 40. Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N Engl J Med*. 2016;375(1):65-74.
634 41. Simon N, Simon R. Adaptive enrichment designs for clinical trials. *Biostatistics*.
635 2013;14(4):613-25.

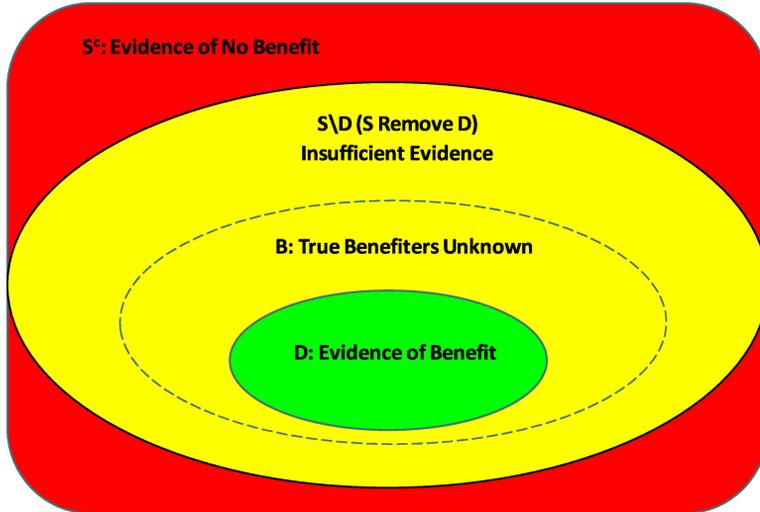
636

S^c: Evidence of No Benefit

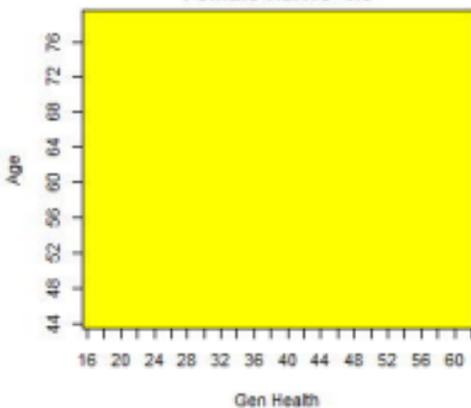
S\D (S Remove D)
Insufficient Evidence

B: True Benefitters Unknown

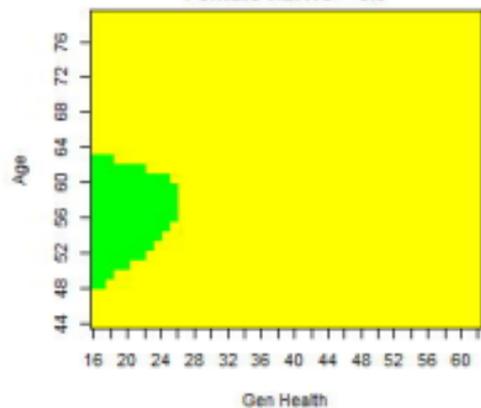
D: Evidence of Benefit



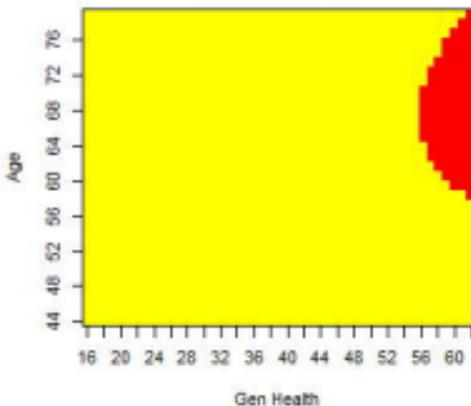
Female HbA1c<6.5



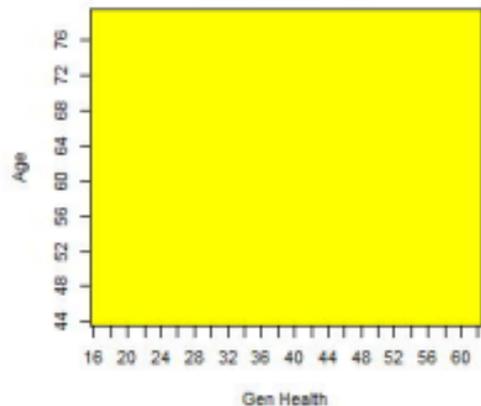
Female HbA1c>=6.5



Male HbA1c<6.5



Male HbA1c>=6.5

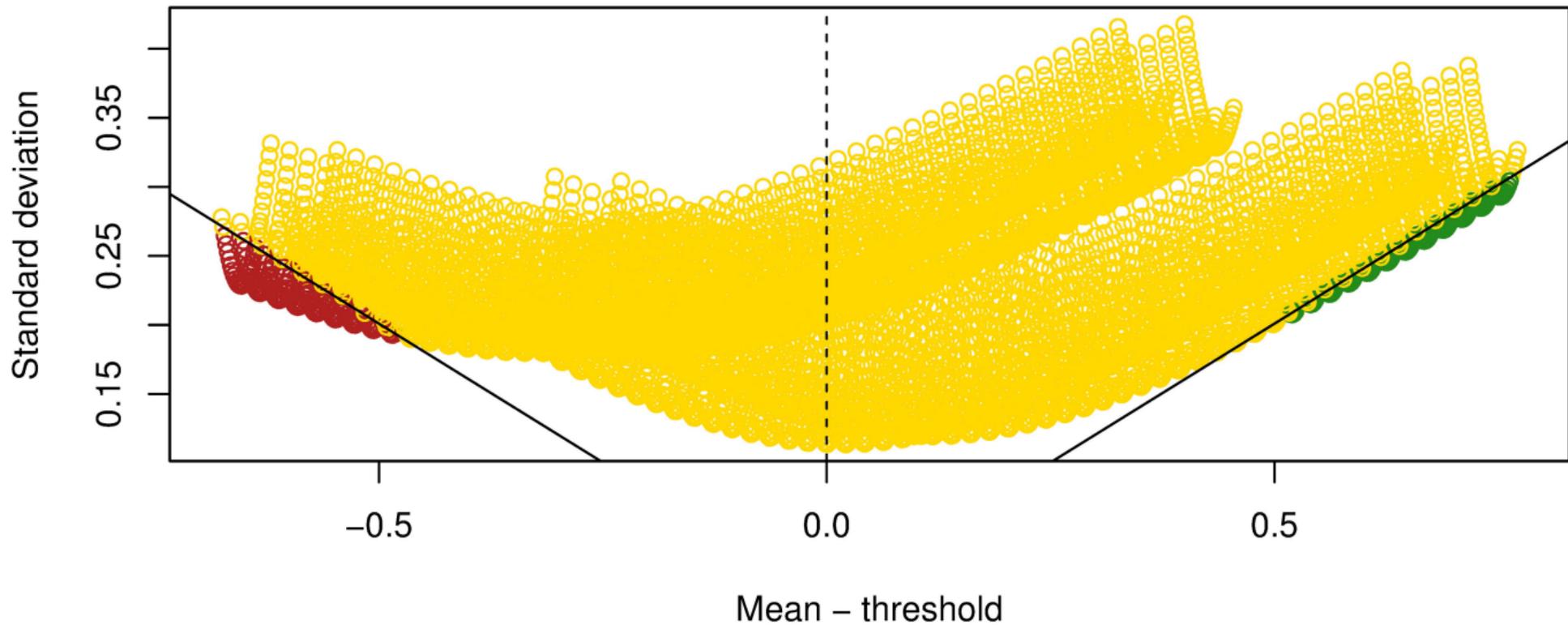


■ D

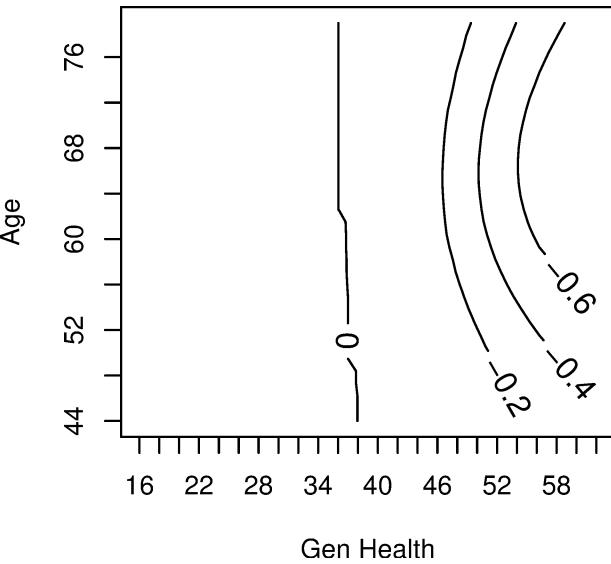
■ S remove D

■ S-complement

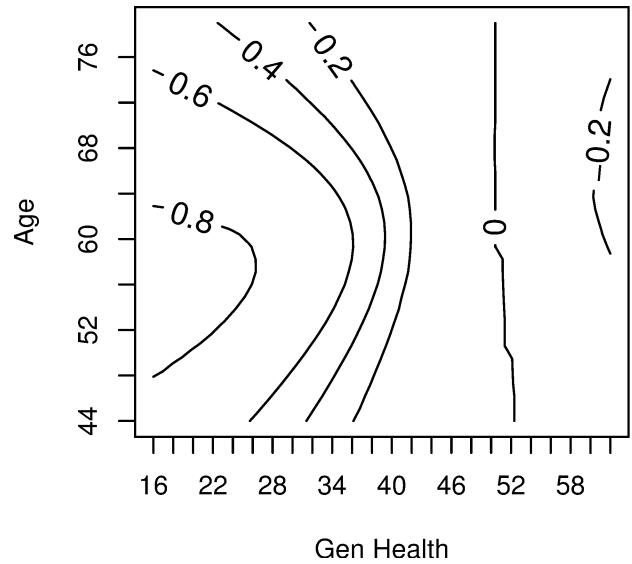
Funnel plot of credible subsets



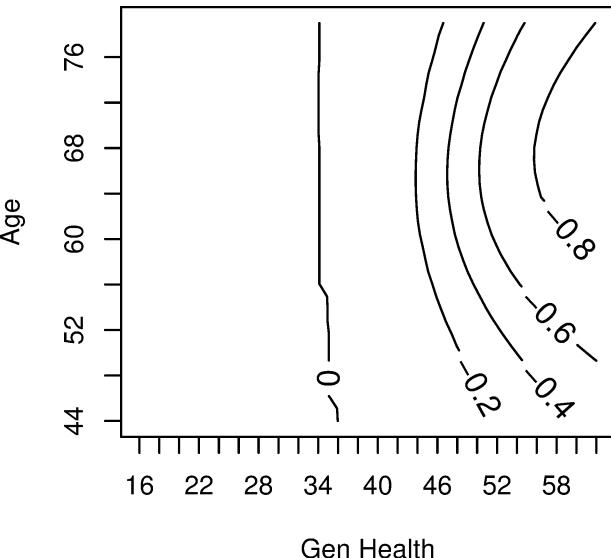
Female HbA1c<6.5



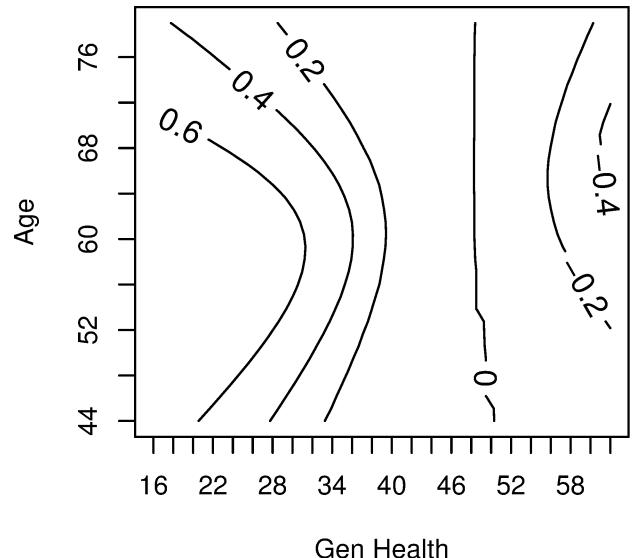
Female HbA1c>=6.5



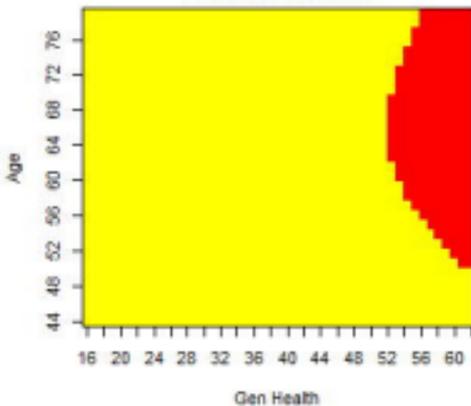
Male HbA1c<6.5



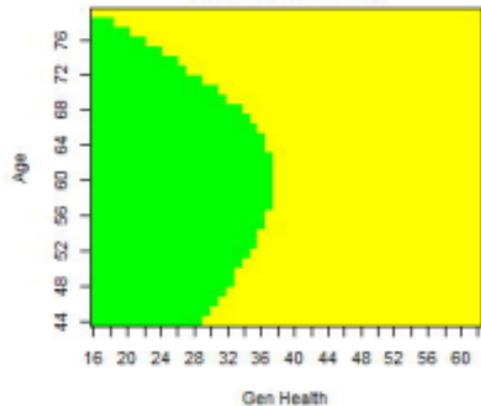
Male HbA1c>=6.5



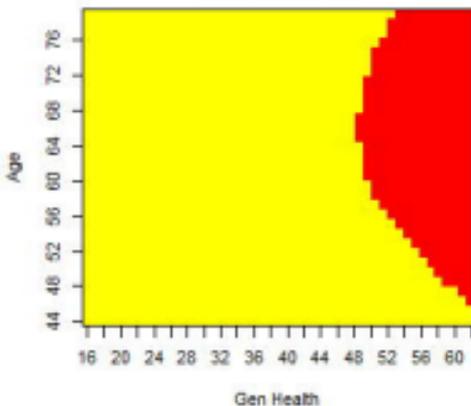
Female HbA1c<6.5



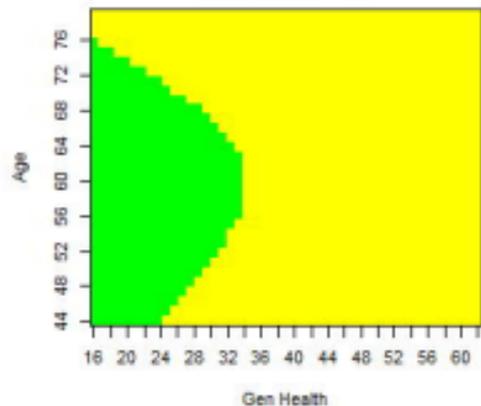
Female HbA1c>=6.5



Male HbA1c<6.5



Male HbA1c>=6.5

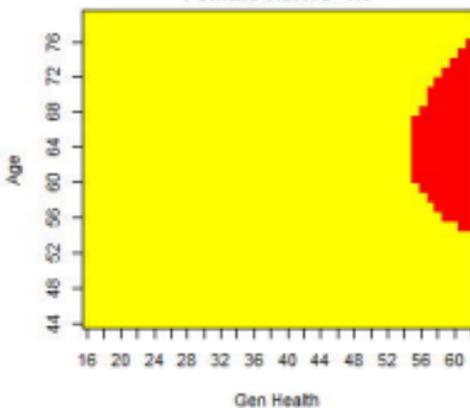


■ D

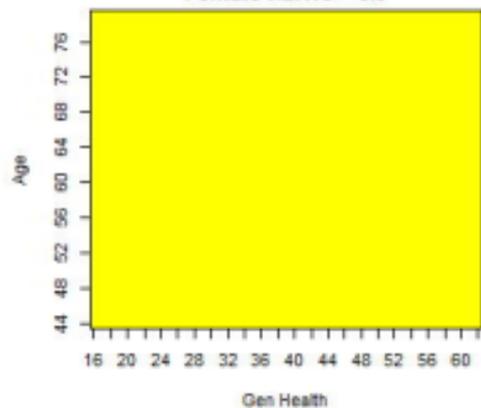
■ S remove D

■ S-complement

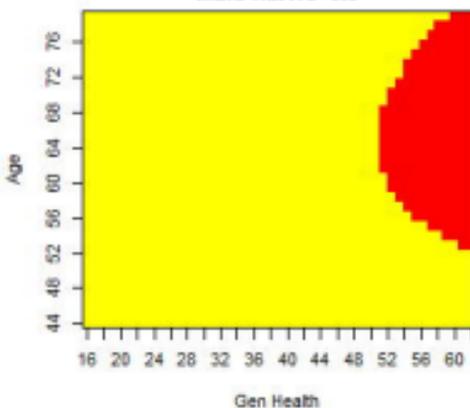
Female HbA1c<6.5



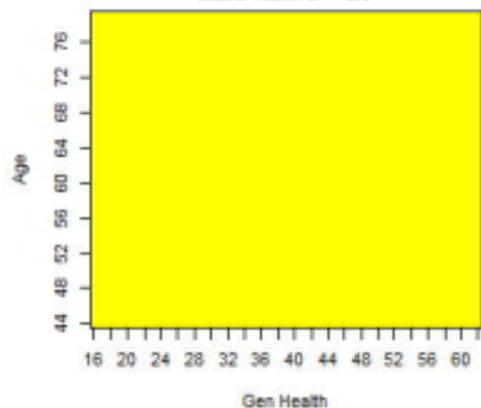
Female HbA1c>=6.5



Male HbA1c<6.5



Male HbA1c>=6.5

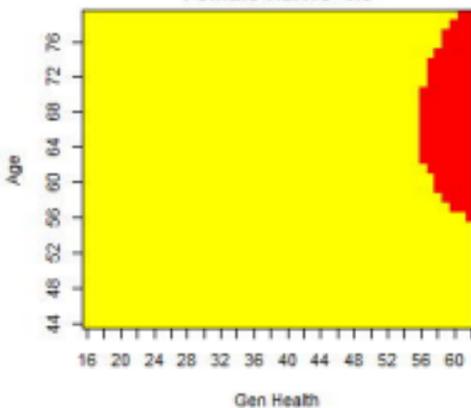


■ D

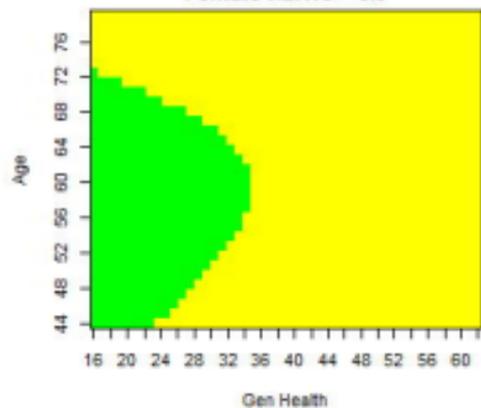
■ S remove D

■ S-complement

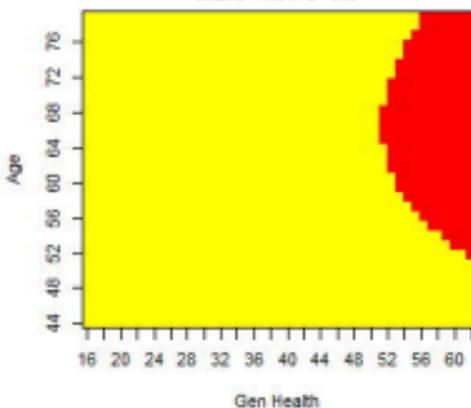
Female HbA1c<6.5



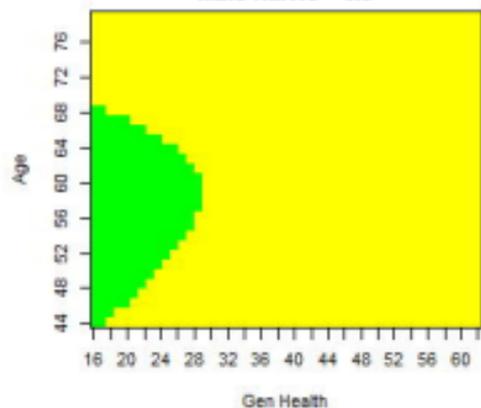
Female HbA1c>=6.5



Male HbA1c<6.5



Male HbA1c>=6.5



■ D

■ S remove D

■ S-complement