

1 **Keep it Cool! Results of a two-year Cool-intervention: a** 2 **descriptive case series study**

3 Nicole Philippens ^{1*}, Ester Janssen ^{1†}, Stef Kremers ¹ and Rik Crutzen ²

4 ¹ Department of Health Promotion, NUTRIM, School of Nutrition and Translational Research
5 in Metabolism, Maastricht University, Maastricht, the Netherlands.

6 ² Department of Health Promotion, CAPHRI, Care & Public Health Research Institute,
7 Maastricht University, Maastricht, the Netherlands.

8
9 † Nicole Philippens and Ester Janssen contributed equally to this work.

10

11 * Corresponding author: n.philippens@maastrichtuniversity.nl

12

13 Keywords: lifestyle, health promotion, intervention, long-term effects, overweight, obesity.

14 **Abstract**

15 *Background:* Coaching on Lifestyle (Cool) is a two-year healthcare intervention for people
16 with overweight or obesity, stimulating weight reduction by promoting sustained healthier
17 behavior. The objective of this study is to investigate the effects of Cool on participants'
18 anthropometrics, personal factors and behavioral factors over the two-year timeframe of
19 Cool.

20 *Methods:* A descriptive case series study, using a broad set of routinely collected data on
21 anthropometrics, personal factors and behavioral factors of adults living across the

22 Netherlands. The data were collected between November 2018 and December 2021 among

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

23 participants of Cool (N=746) at three moments during the intervention: at baseline (T0), at 8
24 months (T1) after completion of phase 1 and at 24 months (T2) after ending Cool. Changes
25 over time were analyzed using paired t-tests comparing baseline to T1 and baseline to T2. In
26 addition, potential differences on outcomes in subgroups based on education level, weight
27 status and group size were examined using paired t-tests and ANOVA-tests.

28 *Results:* The results showed changes in the desired direction on all outcomes at 24 months
29 compared to baseline. The largest effects were seen on perceived health, attentiveness towards
30 meal size and meal composition (large effect size). Mean weight loss was 4.13 kilograms (SD
31 7.54), and mean waist circumference decreased with 4.37 centimeters (SD 8.59), indicating a
32 medium to large effect size. Changes in outcomes were similar across all subgroups
33 consisting of participants with different educational level, BMI at baseline and in different
34 group sizes.

35 *Conclusion:* The study showed sustained effects in weight-related outcomes of Cool over the
36 timeframe of 24 months supporting the two-year set-up of Cool. The outcomes indicate that
37 Cool is appropriate and effective for different group sizes and for a wide variety of
38 participants irrespective of gender, age, BMI at baseline or level of education.

39 **Introduction**

40 Obesity is considered a chronic disease according to the World Health Organisation [1] and
41 the Dutch Health council [2] and it is linked to many other diseases – both physical and
42 mental [3] – and a diminished quality of life [4].

43 Overweight is more common among men (53%) than women (47%), obesity on the other
44 hand is more prevalent in women (17%) than in men (12%). Approximately 41% of people
45 with a higher level of education are overweight whereas this percentages rises to 60% for
46 people with a lower level of education. The proportion of people with severe obesity is twice

47 as high among people with a lower level of education (18%) compared to those with a higher
48 level of education (10%) [5].

49 Consensus has been reached internationally on the importance of an integrated approach to
50 target overweight and obesity, including limited energy intake, healthy food choices and
51 regular physical activity [6]. The Dutch national guidelines have added stress management
52 and sleep as additional essential elements to tackle overweight and obesity [7].

53 As of January 2019, Combined Lifestyle Interventions (CLIs) are part of basic health
54 insurance. Having a basic health insurance is a legal obligation for every person living or
55 working in the Netherlands, and as such a CLI is available for all adults meeting the inclusion
56 criteria (i.e. being obese ($BMI > 30$) or being overweight ($25 < BMI < 30$) combined with
57 comorbidity; and being sufficiently motivated).

58 Coaching on Lifestyle (Cool) is one of these CLIs; a two-year healthcare intervention for
59 people with overweight or obesity, stimulating weight reduction by promoting sustained
60 healthier behavior. The set-up of Cool is in line with the recommendations of the WHO. In
61 addition, the outcomes of the first eight months of Cool, even during COVID-19 and its
62 accompanying restrictions, showed substantial and promising results. Both aspects make
63 Cool an appropriate intervention for people that are overweight or obese [6, 8].

64 Research on long term effects of lifestyle and/or behavior change interventions has been done
65 mainly on specific patient groups and disease related outcome measures [9]. The long-term
66 effects of CLIs (including Cool) in the Netherlands are still unknown, mainly due to the short
67 timespan that the CLI is currently running. So far, effect outcomes over the full intervention
68 period are limitedly available and when outcomes are available, they are showing a
69 stabilization or small relapse in the second year of the intervention [10]. This is the first
70 research on the changes over time on participants over the full Cool-intervention course of 24
71 months.

72

73 We suspect the changes in outcomes to vary with differences in participants and with
74 differences in the context in which the intervention is executed. As Cool aims at higher levels
75 of self-management and self-steering we suspect that the intervention is less effective for
76 people with a lower level of education in line with previous research indicating that level of
77 education shows a strong and positive correlation with health and health related behaviors
78 [11, 12]. A lower BMI at baseline has previously been associated with better program
79 adherence [13] and a higher BMI is associated with unhealthier food choices, i.e. less fruit
80 and vegetable, less fiber, and more fried food consumption [14] whereas on the other hand
81 health interventions seem equally appropriate for different BMI-categories at baseline [15].
82 These contradictory findings sparked our interest to investigate the relationship between BMI
83 at baseline and differences in effect sizes on the outcomes of Cool as well. Furthermore, we
84 are interested in the differences in effect of a large versus a small group size on the Cool
85 outcomes. No consensus has yet been reached on the optimal group size for group
86 interventions, while in Cool these group sizes vary per context. Research in education has
87 shown that a group size of five, compared to fifteen members, enhances participation and
88 satisfaction of the group members [16] whereas groups of nine or more participants bring
89 diversity of thought, experiences, and viewpoints, thereby stimulating active participation of
90 group members [17]. Group lifestyle interventions are usually offered in groups of 10 up to
91 15 participants [18, 19]. These mixed findings do not provide a clear picture on the optimal
92 group size for health interventions. In our definition, aligned with the practice of Cool, large
93 groups consist of 10 or more participants whereas small groups have less than 10 participants.
94 We hypothesize that participants in small Cool groups show larger effect sizes on the
95 outcomes of Cool because a smaller group provides the coach with more time and focus per
96 participant, thereby stimulating active participation and behavior change.

97

98 Our objective is to study the effects of Cool after 24 months on anthropometrics, personal
99 factors and behavioral factors of the participants. In addition, we want to analyze potential
100 differences on the outcomes for people with lower education compared to medium or higher
101 education, people with a lower BMI at baseline compared to a higher BMI and for people
102 participating in Cool in groups under or over 10 participants.

103

104 **Materials and Methods**

105 **Cool-intervention**

106 Cool is a Combined Lifestyle Intervention (CLI) including a one-hour intake and phase 1 (8
107 months) in which behavioral change is initiated followed by phase 2 (16 months) in which
108 both behavioral change and behavioral maintenance are targeted. The intervention consists of
109 individual sessions (6 hours in total) and 8 group sessions (1,5 hours each) both in phase 1
110 and phase 2, resulting in a higher density of sessions in phase 1 compared to phase 2. Cool
111 aims at changes in anthropometrics (i.e. weight, BMI and waist circumference) and at an
112 increase in perceived quality of life by stimulating healthier eating habits, less sitting time,
113 more physical activity and attention for sufficient relaxation and high quality sleep.

114

115 Cool is an open CLI, i.e. an intervention without a strict protocol. Coaches may adapt the
116 intervention to the target group and context as long as the main effective elements of Cool
117 (e.g., goal setting, mobilizing social support, positive psychology, self-management and self-
118 monitoring) are respected in implementation. The Cool-coach is a trained and licensed
119 professional who coaches participants towards a predefined set of final objectives on health-
120 related skills and knowledge. Participants are stimulated to take responsibility for their

121 personal lifestyle changes by addressing motivation, personal objectives and behavioral
122 changes. The main objective is to coach and activate participants to a sustained healthier
123 lifestyle in line with their individual needs and personal goals.

124

125 **Study design and population**

126 As Cool is part of regular health care in the Netherlands, a control group receiving no
127 treatment would be both unethical and impractical, making a descriptive case series study the
128 appropriate study design in the Dutch context. The participants are adults living throughout
129 the Netherlands. All participants met the inclusion criteria for participating in a CLI and were
130 referred to Cool by their general practitioner, practice nurse or internist. The decision on a
131 proper fit for inclusion was up to the participant, the referrer and the coach. All participants
132 signed an informed consent regarding data collection for this study.

133

134 **Data collection**

135 We used a lifestyle questionnaire and anthropometric measurements to collect a broad set of
136 data. The lifestyle questionnaire was based on existing validated questionnaires. The outcome
137 measures can be divided into the categories: anthropometrics (i.e. weight/BMI and waist
138 circumference), control and support (i.e. self-mastery and social support), physical activity
139 (i.e. sedentary time and active minutes), diet attentiveness, alcohol use and smoking,
140 perceived fitness (i.e. perceived health, fitness and impact of stress on daily functioning),
141 sleep and stress.

142 During the course of the study, the questionnaire was extended with additional questions
143 covering changes in context (e.g. COVID-19) and adjusted with textual simplifications in
144 both questions and answers preserving the original essence as much as possible.

145 **Datasets**

146 Data were collected from November 2018 until December 2021 at three moments in time: at
147 baseline, during the intake (T0), after completion of phase 1 (T1) and after ending Cool (T2).
148 No information is available on the exact number of participants starting with or dropping out
149 of the intervention as data is submitted only by participants who agreed to share data. In
150 addition, data collection is restricted to two moments during the intervention: after 8 months
151 (at T1) and when ending Cool (at T2). The data at T2 contains participants that completed
152 the intervention (sent in after approximately 24 months) and participants that dropped out
153 earlier in time (sent in at the moment of dropout, which could be at any moment during the
154 two-year intervention). Participants with a T2 measurement were included in the dataset if
155 their (estimated) completion date of Cool was or would have been before the end of the data
156 collection period, i.e. December 31st 2021.

157

158 All analyses were performed between May 2022 and May 2023 and were done on the full
159 dataset (A) consisting of program finishers and dropouts, to provide a realistic reflection of
160 the potential intervention effects in practice. In addition, we analyzed a cleaned dataset (B)
161 including all participants that completely finished the two-year intervention, to portray
162 efficacy of the intervention. Changes over time were measured from baseline to T1 and from
163 baseline to T2.

164

165 **Demographics**

166 At baseline, participants reported their personal characteristics such as gender, date and
167 country of birth, highest completed education, marital status, living situation and occupational
168 status. Educational level was categorized in line with the Dutch Central Bureau of Statistics
169 (CBS) into low (i.e., no education or primary education), intermediate (e.g., secondary

170 education) and high (e.g., tertiary education). The living situation was divided into living
171 together with someone (married or cohabiting) with or without kids and living alone
172 (divorced, unmarried or widowed) with or without kids. The occupational status was
173 categorized into working (e.g. paid work, voluntary work or self-employed) and not working
174 (e.g. stay-at-home, unemployed, retired or student). Country of birth was categorized into
175 Dutch or non-Dutch.

176

177 **Anthropometrics**

178 Under normal conditions anthropometric data (weight, length and waist circumference) were
179 measured by the Cool-coaches with professional equipment according to the guidelines
180 provided by the Dutch Association of General Practitioners (Dutch: Nederlands Huisartsen
181 Genootschap, NHG) [20]. Body weight (kg) was measured in kilogram, rounded off the
182 nearest decimal. Height (m) was measured to the nearest centimeter without shoes. Waist
183 circumference measurements were obtained to the nearest centimeter with a tape measure. As
184 COVID-19 restrictions could have changed the measurement method, additional information,
185 gathered from the Cool-coaches that were the main data suppliers (representing data of 227
186 participants), confirmed that in general, physical measurements took place either by the coach
187 or on a distance of 1.5 meters under direct supervision of the coach.

188

189 **Control and support**

190 The self-mastery questions in the questionnaire were based on the short version of the Pearlin
191 Mastery Scale using four questions (for example “I have little control over the things that
192 happen to me”) and a 5-point Likert scale ranging from strongly agree (1) to strongly disagree
193 (5) [21]. To identify social support, we questioned the perceived support of close ones using a
194 5-point Likert scale ranging from no support at all (1) to a lot of support (5).

195

196 **Physical activity**

197 The outcome measurements on physical activity, diet and personal factors were defined in
198 cooperation with the Dutch Association of Lifestyle Coaches (BLCN) with the objective to
199 capture the essence and map the desired outcomes of lifestyle coaching in a minimum set of
200 questions. Physical activity used questions on sedentary behavior, both on most and least
201 active days (“What is the average number of hours you spent sitting on the day of the week
202 you sit the most?”) and the number of physical activity minutes per day (“What is the average
203 minutes per day that you are physically active (in minimum bouts of 10 minutes)?”).

204

205 **Diet attentiveness, alcohol and smoking**

206 We defined questions on dietary attentiveness, in line with the input of the BLCN, based on
207 the idea that deliberate behavior changes start with being aware of one’s own behavior. We
208 used questions on the attentiveness of participants towards meal composition and meal
209 quantities and attentiveness during the actual consumption of food using a 5-point Likert scale
210 from very little attention (1) to a lot of attention (5). At T1 and T2 an additional question was
211 added regarding changes in eating pattern: a reflection of the perception on healthy diet
212 improvements compared to baseline (“How much healthier have you been eating since the
213 intake of this program?”) with the answers ranging from much healthier (1) to much
214 unhealthier (5). The amount of alcohol and smoking was questioned by numerical values.

215

216 **Perceived fitness**

217 Perceived fitness consisted of questions, in line with the input of the BLCN, on perceived
218 fitness when waking up and during the day, the impact of stress on daily functioning and on
219 perceived health (i.e. feeling good about oneself, the extent of self-care invested and the

220 perception of one’s general health). Questions were answered using a 5-point Likert scale,
221 ranging from not good at all (1) to very good (5).

222

223 **Sleep**

224 We defined a specific set of questions around the sub-constructs: subjective sleep quality,
225 sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep
226 medication and daytime dysfunction, analogous to the validated and widely used PSQI-
227 questionnaire [22]. Each subconstruct was covered by one or two question(s) using a
228 numerical value or a 4-point Likert scale, ranging from ‘never’ (1) to ‘three times per week or
229 more frequently’ (4).

230

231 **Stress**

232 For stress, the validated Perceived Stress Scale questionnaire was used, which exists of ten
233 questions using a 5-point Likert scale from never (1) to always (5) [23].

234

235 **Statistical analyses**

236 **Data preparation**

237 We recoded some of the variables to facilitate interpretation in the sense that a higher/positive
238 score refers to a desirable trend and a lower/negative score to an undesirable trend in the
239 variable. For constructs based on validated questionnaires (i.e. sleep and stress) we adopted
240 the accompanying approach without recoding. Secondly, we performed an exploratory factor
241 analysis and calculated McDonald’s omega to assess the internal structure of items regarding
242 several constructs such as perceived health, self-mastery, sleep and stress in line with Crutzen
243 et al. [24]. These analyses justified summarizing all lifestyle constructs by item score means.
244 Missing data were excluded from the statistical analyses.

245

246 **Effect sizes**

247 For all items and constructs, we ran descriptive statistics. Changes over time were analyzed
248 using paired t-tests comparing baseline to T1 and baseline to T2. Effect sizes were calculated
249 and interpreted in accordance with Lipsey's guidelines for each pair of items or constructs,
250 i.e. an effect size smaller than 0.32 is considered small, an effect size between 0.33 and 0.55
251 is considered medium and an effect size above 0.56 is considered large [25]. To improve
252 comprehensibility effect sizes were represented such that positive values represented change
253 in the desired direction whereas negative values represented change in an undesired direction.
254 All T-tests were performed using SPSS-software (version 27). Missing data were excluded
255 from the statistical analyses.

256 To be considered successful the target for the CLI (including CoolL) is an average 5% weight
257 loss for all participants, as set by the Dutch Partnership Overweight (Dutch: PON), an
258 advisory body for the Dutch government on obesity related health issues. We categorized the
259 outcomes on weight: 5% weight loss or more, between 0 and 5% weight loss, weight
260 stabilization and weight gain, to map the percentage of participants that comply with this
261 target.

262 **Subgroup analyses**

263 We compared different subgroups in sequence to explore potential differences in outcomes,
264 i.e. subgroups based on educational level of the participants, on BMI at baseline and on group
265 size at the start of CoolL. Subgroup analyses were done on the full dataset (A) including
266 program finishers and dropouts. To enable subgroup comparison, we calculated the difference
267 (delta) between T0 and T2 for each construct or variable. As a higher starting weight usually
268 requires less effort to lose a certain amount of weight, we looked at relative (%) weight loss

269 compared to baseline for the BMI-subgroup comparison. For the construct ‘eating pattern’ we
270 used the construct itself as it already includes changes in eating pattern compared to baseline
271 in the formulation of the questions. When comparing two subgroups we performed
272 independent t-tests comparing all delta-variables. In case of multiple subgroups, we ran an
273 ANOVA test on the delta-variables followed by post-hoc Tukey tests to analyze potential
274 differences in effect.

275

276 **Drop-outs**

277 We used logistic regression analysis to determine the main factors related to drop-out. The
278 predictor variables in the logistic regression were based on the pre-defined subgroups of
279 interest (i.e. based on group size, BMI at baseline and educational level) and two additional
280 demographic variables (i.e. age category and gender). For the dropout analysis we used the
281 full dataset (A) excluding the participants (n=22) that could not be assigned as program
282 finisher or dropout due to missing information.

283 The dropout analysis showed no distinct pattern in dropout profiles. However, specific
284 subcategories of some of the constructs were less likely to drop out in comparison to the
285 reference category, i.e. a BMI of 35-40 compared to BMI<30, participating in a group of over
286 10 compared to less than 10 and a higher level of education compared to a lower level of
287 education all were less likely to drop out. The constructs gender and age showed no
288 differences in dropout. See Appendix 1 for the details on dropout percentages and related
289 analyses.

290 As CoolL is part of basic health insurance and data is gathered from all participants, provided
291 that they gave written consent for the use of their anonymized data, selection bias is limited.

292 In addition, we tried to minimize bias by ensuring a check on all analyses by a second

293 researcher, by including both program finishers and dropouts in our analyses and by
294 presenting a complete set of outcomes on all variables and analyses.

295

296 **Ethics**

297 This study was submitted to and approved by the Research Ethics Committee of the Faculty
298 of Health, Medicine and Life Sciences of Maastricht University (FHML-REC/2019/073). All
299 participants gave their written informed consent for their anonymised personal data to be used
300 for research purposes.

301 **Results**

302 **Datasets**

303 We collected data from in total 3780 participants that started Cool between November 2018
304 and December 2021. See Figure 1 for a graphical representation of the dataset selection steps.

305

306 **Figure 1.** Flowchart dataset selection process.

307

308 **Demographics**

309 Of all participants in the full dataset (A) a total of 28% were male and 72% female. This ratio
310 is in line with the data from the national CLI-monitor [26]. Most participants (93%) had a
311 Dutch background. In total, 69% of the participants had a lower or intermediate level of
312 education; 30% did not have a steady job (anymore) and approximately two third of the
313 participants were living together with a partner (see Table 1).

314 The cleaned dataset (B), containing only respondents that finished the intervention, showed in
315 general a similar demographic picture except for the educational level of the participants: this

316 dataset contained relatively more participants with a higher-level education and less
 317 participants with a lower-level education.

318

319 **Table 1.** Demographics of the participants

Category	Demographic	Number of participants (%) full dataset (A)	Number of participants (%) cleaned dataset (B)
Gender	Male	203 (28%)	105 (28%)
	Female	519 (72%)	272 (72%)
	Other	1 (0.1%)	1 (0.3%)
Age	Until 35 years	75 (11%)	41 (12%)
	35 – 44 years	90 (13%)	42 (12%)
	45- 54 years	166 (25%)	89 (26%)
	55 – 64 years	198 (29%)	103 (30%)
	65+	147 (22%)	70 (20%)
Living situation	Single	157 (22%)	85 (22%)
	Single parent	48 (7%)	30 (8%)
	Living together with kids	284 (39%)	150 (40%)
	Living together without kids	200 (28%)	98 (26%)
	Other	33 (4%)	15 (4%)
Country of birth	Dutch	668 (93%)	348 (93%)
	Non-Dutch	49 (7%)	27 (7%)
Working situation	Employed	515 (70%)	276 (71%)
	Unemployed	217 (30%)	111 (29%)
Education	Lower level	197 (28%)	88 (23%)
	Intermediate level	293 (41%)	158 (42%)
	Higher level	223 (31%)	129 (35%)

Participants	Total number	746	396
---------------------	---------------------	------------	------------

320

321 All results on the anthropometric and personal factors in the full dataset (A) are summarized
322 in Table 2 whereas more detailed information is provided in appendix II.

323

324 **Anthropometrics**

325 The anthropometric measurements, i.e. weight, BMI and waist circumference, all showed a
326 medium effect size in the desired direction at T1 increasing slightly at T2. Participants
327 showed on average a decrease of 4.1 kg weight, 1.4 BMI point and 4.4 cm waist
328 circumference after two years of CoolL.

329 Three quarters of all participants showed weight loss during 24 months of CoolL and 32% of
330 all participants showed more than 5% weight loss. On average participants lost 3.8% weight
331 during these 24 months.

332 The CoolL finishers (dataset B) showed slightly better outcomes at T2, i.e. an average weight
333 loss of 4.7 kg, a decrease of 1.6 BMI point and a decline of 5.5 cm in waist circumference at
334 T2 (all large effect sizes).

335

336 **Personal factors and feeling fit**

337 Participants experienced an increased feeling of self-mastery at T2 (small effect size) and an
338 improvement in perceived health both at T1 and T2 (large effect size). Feeling fit when
339 waking up, showed an improvement with a small effect size both at T1 and T2 whereas
340 feeling fit during the day showed no effect at T1 and a small effect size at T2. No effect was
341 found on perceived support and influence of stress on daily functioning both at T1 and T2
342 compared to baseline.

343 The CoolL finishers (dataset B) showed similar effects and effect sizes.

Table 2. Overview of changes over time in anthropometrics and personal factors in complete population and in subgroups

Changes over time on full dataset (A) at T1 and T2								ΔT0T2 comparing subgroups		
Construct/ factor	T0 M (SD)	T1 M (SD)	T2 M (SD)	Δ T0T1 [95% CI]	Cohen's d** T0T1	Δ T0T2 [95% CI]	Cohen's d** T0T2	P-value (T-test) on LLE vs IHLE ¹	P-value (ANOVA) on BMI ²	P-value (T-test) on group size ³
Anthropometrics										
Weight	105.63 (18.61)	101.65 (17.59)	101.39 (18.75)	-3.26 [-3.82; - 2.70]*	0.53	-4.13 [-4.74; -3.51]*	0.55	0.07	0.02*	0.97
BMI	35.97 (5.29)	34.90 (5.40)	34.60 (5.41)	-1.12 [-1.31; - 0.93]*	0.53	-1.40 [-1.61; -1.19]*	0.55	0.15	0.00*	0.84
Waist circumference	116.38 (13.1)	112.25 (14.20)	111.72 (14.49)	-3.42 [-4.27; - 2.56]*	0.42	-4.37 [-5.17; -3.57]*	0.51	0.21	0.12	0.95
Personal factors and feeling fit										
Self-mastery	2.54 (0.81)	2.44 (0.79)	2.42 (0.73)	-0.06 [-0.13; 0.01]	0.10	-0.11 [-0.18; -0.03]*	0.15	0.40	0.99	0.21
Perceived health	8.93 (2.27)	10.40 (2.08)	10.47 (2.29)	1.39 [1.16; 1.62]*	0.58	1.56 [1.35; 1.77]*	0.64	0.35	0.48	0.22
Fitness (waking)	2.45 (1.01)	2.68 (0.85)	2.70 (0.89)	0.20 [0.11; 0.30]*	0.20	0.25 [0.17; 0.34]*	0.25	0.21	0.76	0.10

Changes over time on full dataset (A) at T1 and T2								ΔT0T2 comparing subgroups		
Construct/ factor	T0 M (SD)	T1 M (SD)	T2 M (SD)	Δ T0T1 [95% CI]	Cohen's d** T0T1	Δ T0T2 [95% CI]	Cohen's d** T0T2	P-value (T-test) on LLE vs IHLE ¹	P-value (ANOVA) on BMI ²	P-value (T-test) on group size ³
Fitness (daytime)	2.63 (0.92)	2.70 (0.86)	2.84 (0.85)	0.06 [-0.04; 0.16]	0.06	0.21 [0.12; 0.29]*	0.22	0.46	0.88	0.07
Support	3.72 (1.07)	3.72 (0.98)	3.80 (0.94)	-0.003 [- 0.12;0.11]	-0.003	0.02 [-0.09; 1.13]	0.02	0.41	0.81	1.00
Influence of stress on daily functioning	2.21 (0.970)	2.27 (0.91)	2.19 (0.89)	0.05 [-0.04; 0.14]	0.05	-0.03 [-0.11; 0.06]	-0.03	0.84	0.39	0.01*

¹ Comparison of two subgroups: participants with a lower level of education (LLE) to participants with an intermediate to higher level of education (IHLE).

² Comparison of different subgroups: participants with a BMI 25-30, BMI 30-35, BMI 35-40 and BMI 40+.

³ Comparison of two subgroups: participants in group sizes of less than 10 participants to group sizes of 10 or more participants.

* p<0.05

** Effect size: positive values represent change in desired direction, negative values represent change in undesired direction

353 All results on the behavioral factors in the full dataset (A) are summarized in Table 3 whereas
354 more detailed data is provided in appendix II.

355

356 **Behavioral factors**

357 No effect was found at T1 for sedentary time (both least and most active days) and a small
358 effect was found at T2: participants spent around half an hour less time on sitting both on
359 least and most active days. Physical activity showed a small effect size both at T1 and T2 with
360 an average increase of 18 minutes at T2. The outcomes on sleep showed that participants
361 experienced a higher quality of sleep at T1 and T2, both with a small effect size. In addition,
362 participants experienced less stress at T1 and T2 (both medium effect size) and participants
363 smoked less at T1 and T2 (small effect size).

364 The dietary outcomes showed that participants paid more attention to meal composition and
365 to the amount of food they consume compared to baseline, both constructs showed a large
366 effect size at T1 and T2. In addition, participants were more attentive during actual
367 consumption of food both at T1 and T2 (both medium effect size). When drinking alcohol,
368 participants consumed on average one unit less alcohol at T1 (small effect size) and this effect
369 was sustained until T2 (small effect size).

370 Regarding change in eating pattern compared to baseline, participants indicated an
371 improvement at T2 compared to T1 with a small effect size.

372

373 The Cool finishers (dataset B) showed deviating outcomes on physical activity (no effect at
374 T1) and smoking (no effect at T1 or T2).

375

376

377

Table 3. Overview of changes over time in behavioral factors in complete population and in subgroups

Changes over time on full dataset (A) at T1 and T2								ΔT0T2 comparing subgroups		
Construct/ factor	T0 M (SD)	T1 M (SD)	T2 M (SD)	Δ T0T1 [95% CI]	Cohen's d** T0T1	Δ T0T2 [95% CI]	Cohen's d** T0T2	P-value (T-test) on LLE vs IHLE ¹	P-value (ANOVA) on BMI ²	P-value (T-test) on group size θ^3
Sedentary time (least active)	9.32 (3.87)	8.92 (3.64)	8.78 (3.46)	-0.28 [-0.63; 0.07]	0.08	-0.65 [-0.99; -0.31]*	0.18	0.14	0.57	0.55
Sedentary time (most active)	6.23 (3.50)	6.22 (3.38)	5.92 (3.22)	-0.06 [-0.40; 0.27]	0.02	-0.45 [-0.76; -0.15]*	0.14	0.18	0.33	0.49
Active minutes	90.01 (113.32)	100.99 (96.50)	108.84 (110.96)	11.04 [0.76; 21(.32)]*	0.11	17.91 [6.73; 29.09]*	0.16	0.00*	0.82	0.02*
Sleep (summary)	6.89 (4.23)	5.85 (4.06)	5.76 (3.95)	-1.04 [-1.46; -0.62]*	0.28	-1.07 [-1.47; -0.67]*	0.29	0.45	0.31	0.81
Stress (summary)	14.50 (6.79)	12.51 (6.38)	12.24 (6.25)	-2.03 [-2.71; -1.36]*	0.35	-2.28 [-2.97; -1.58]*	0.37	0.35	0.37	0.93
Smoking	1.06 (4.21)	0.57 (3.15)	0.67 (3.14)	-0.44 [-0.69; -0.20]*	0.14	-0.46 [-0.75; -0.17]*	0.12	0.25	0.25	0.91
Meal composition	2.84 (0.99)	3.47 (0.89)	3.55 (0.87)	0.63 [0.52; 0.74]*	0.59	0.67 [0.57; 0.77]*	0.63	0.42	0.03*	0.70
Amounts of food	2.64 (0.93)	3.42 (0.89)	3.45 (0.94)	0.77 [0.67; 0.88]*	0.76	0.73 [0.63; 0.84]*	0.67	0.65	0.15	0.76
Attentive to consuming	2.80 (1.12)	3.34 (0.94)	3.33 (0.96)	0.51 [0.40; 0.61]*	0.48	0.56 [0.46; 0.66]*	0.51	0.83	0.24	0.78
Alcohol	1.74 (2.87)	0.66 (1.39)	0.59 (1.46)	-1.06 [-1.26; -0.87]*	0.41	-1.09 [-1.30; -0.89]*	0.43	0.12	0.18	0.97
Eating pattern***	N/A	3.98 [0.64]	4.04 [0.69]	N/A	N/A	0.13 [0.06; 0.20]*	0.18	0.36	0.60	0.45

379
380
381
382
383
384
385
386

¹ Comparison of two subgroups: participants with a lower level of education (LLE) with intermediate to higher level of education (IHLE).

² Comparison of different subgroups: participants with a BMI 25-30, BMI 30-35, BMI 35-40 and BMI 40+.

³ Comparison of two subgroups: participants in group sizes of less than 10 participants versus groups of 10 or more participants.

* $p < 0.05$

** Effect size: positive values represent change in desired direction, negative values represent change in undesired direction

*** Measurement at T1 and T2 only: estimate of improvement in eating pattern compared to baseline, $\Delta T0T2$ represents difference in estimate between T1 and T2.

387

388 **Subgroup analyses**

389 We compared subgroups based on the categorization of different constructs, i.e. educational level, BMI at
390 baseline and group size at the start of the two-year Cool-intervention. All subgroup analyses were done on
391 the full dataset (A).

392

393 **Subgroup: educational level**

394 Comparing participants with a lower level of education (LLE) to an intermediate or higher level of
395 education (IHLE) showed in general no differences in effects. The only difference between both subgroups
396 was found in active minutes. Participants with a lower level of education showed a larger increase from
397 baseline to T2: 49 minutes of increased physical activity compared to 7 minutes for participants with a
398 medium to higher level of education. See Tables 2 and 3 for the p-values on the subgroup comparison on
399 educational level and appendix III for more detailed outcomes of the subgroup analysis.

400

401 **Subgroup: BMI at baseline**

402 For most constructs no differences in effects were observed between participants of different BMI-
403 categories. The comparison showed differences only on change in BMI, percentage weight loss and meal
404 composition between T0 and T2. See tables 2 and 3 for the p-values on the subgroup comparison on BMI-
405 category and appendix IV for more detailed outcomes of the subgroup analysis. Looking into more detail,
406 these differences were found for a limited number of categories (see appendix IV, table b).

407 The outcomes for participants with a larger BMI at baseline showed equal effect sizes on most constructs
408 and better outcomes on weight loss percentage/BMI. The only exception was the attentiveness to meal
409 composition: participants with a BMI<30 at baseline showed more improvement on this construct compared
410 to participants with a BMI 40+.

411

412 **Subgroup: Group size at baseline**

413 A comparison between the outcomes of participants that start in small groups (group size <10) versus large
414 groups (group size 10+) showed similar results. Only for two constructs differences were observed, i.e. the
415 influence of stress on daily functioning (table 2) and the number of daily active minutes (table 3). For
416 participants in smaller groups (<10) the influence of stress on daily functioning showed an increase at T2
417 compared to baseline (i.e. a more positive influence of stress on daily functioning) whereas for participants
418 in larger groups (10+) the influence of stress showed a decrease at T2 (i.e. a more negative influence). For
419 participants in smaller groups the physical active minutes increased on average at T2 with 38 minutes
420 whereas for participants in larger groups the physical active minutes increased with 9 minutes compared to
421 baseline. See appendix V for more detailed outcomes of the subgroup analysis on group sizes.

423 **Discussion**

424 The results of the study showed changes in the desired direction on all outcomes at 24 months compared to
425 baseline. The largest effect sizes were found on perceived health, attentiveness towards meal size and meal
426 composition (large), followed by weight loss/BMI and waist circumference (medium to large). Medium
427 effect sizes were found on attentiveness to consuming, alcohol intake and stress perception. All other
428 behaviors showed small effect sizes whereas very few outcomes showed no effect.

429
430 Looking at changes in the timeframe of baseline to 8 months, the pattern is similar to the outcomes from
431 previous Cool-research [8]. In addition, the present study showed sustained and improved results in Cool-
432 participants, including enlarged weight loss over the full term of 24 months, even though they were exposed
433 to the COVID-19 pandemic and restrictions. Durable weight loss (i.e. weight maintenance) can be defined
434 as intentional weight loss that has been maintained for at least 6 months [27]. In general, initial weight loss
435 is considered relatively easy whereas the opposite is true for durable weight loss. Follow-up measurements
436 of lifestyle programs usually report weight regain compared to baseline after one year [10, 28-30]
437 underlining the importance of the 24 months duration of Cool with a continued focus on behavior change
438 and behavior maintenance in phase 2 of the program.

439

440 The results of the full dataset compared to the cleaned dataset were quite similar, indicating limited selective
441 dropout. Participants that finished CoolL show in broad terms slightly improved outcomes compared to the
442 participants in the full dataset, probably because, on average, dropouts participated 11 months in CoolL,
443 whereas participants that finished CoolL received additional months of guidance and support. In addition, the
444 dropout-group consisted of participants with a range of results on both ends: participants with such positive
445 results early in the program needing no further assistance in behavior maintenance and participants with
446 such negative or less encouraging results wishing no longer to continue the intervention. These results on
447 both ends of the spectrum, seem a reasonable explanation for the average result on the dropout group in
448 total.

449

450 Despite the cut-off date that was applied to balance the dropouts in the dataset, the number of dropouts was
451 still relatively high compared to earlier research on CoolL [31]. As we suspected the COVID-19 pandemic
452 and its restrictions to have a major influence on dropout rates, a quick analysis on the monthly dropouts
453 from June 2019 until June 2021 showed that dropouts more than doubled when restrictions resulting from
454 the COVID-19 pandemic came into effect starting March 2020.

455

456 Although the group of lower educated was somewhat overrepresented in the dropout group, the educational
457 level did not seem to interfere with the achieved effects as participants with a lower level of education
458 showed identical effects in outcomes as participants with higher educational levels, similar to the outcomes
459 of comparable health interventions like SLIMMER [32]. Tentatively, the present study provides indications
460 that CoolL does not enlarge health inequalities and even shows potential to decrease these inequalities, under
461 the condition that participants with a lower level of education can be guided towards sustained participation
462 in CoolL. Physical active minutes provided the only exception as participants with a lower level of education
463 showed a larger increase despite more physical active minutes at baseline, which might be due to a
464 confusing question on (bouts of) active minutes generating a mix of active minutes or bouts of 10 minutes.

465 Participants with a lower level of education were more likely to dropout during Cool which is in line with a
466 higher dropout rate of MetSLIM, a CLI for low SES, compared to the regular SLIMMER CLI [32] and with
467 earlier findings illustrating that higher education serves as a protective factor against dropout[33]. In
468 conclusion, Cool is just as effective for participants with a lower level of education, though extra effort is
469 needed to prevent dropout for this target group.

470
471 The outcomes for participants with a larger BMI at baseline showed equal effect sizes on most constructs
472 and even better outcomes on weight loss/BMI with attentiveness to meal composition as the only exception.
473 These outcomes indicate that the Cool-intervention is appropriate and effective for all BMI-categories.

474
475 When comparing the outcomes of participants in small groups versus large groups, the similarities stand out,
476 as only differences were found on the impact of stress on daily functioning and on physical active minutes,
477 both in favor of a smaller group size. However, a smaller group size is related to more dropouts, potentially
478 due to the fact that a larger group size increases the chance of finding a suitable buddy or role model among
479 the group as group participants and Cool-coach stay together from start to end. These findings leave the
480 ideal group size for Cool undecided leaving room for the Cool-coach to act on personal preferences as an
481 extra group member provides more income but requires extra effort in individual support and group
482 dynamics.

483 484 **Limitations and strengths**

485 A control group for comparison was no option, as the CLI Cool is part of basic health care in the
486 Netherlands. Therefore, the results of Cool should be labelled as changes over time instead of effects as we
487 cannot rule out interference with other factors and variables. We are less hesitant in addressing these
488 changes to Cool given the average effect size of the changes, previous results of Cool and the comparison
489 to similar interventions.

490 During the time of the study the questionnaire was revised with minor changes. We intended to keep the line
491 of questioning and answering the same, but we cannot rule out any effect on the study. The impact of

492 COVID-19 on the intervention, coaches and participants can be considered a second limitation though the
493 results of the CoolL-intervention on participants during COVID-19 are substantial and encouraging [8].
494 We cannot distinguish between dropouts and loss to follow-up as we are dependent on the respondents to
495 hand-over their results. As a consequence, we cannot provide exact numbers on participants and dropouts of
496 CoolL. We compared the cleaned dataset with the full dataset, the latter including dropouts. The comparison
497 is insightful but does not support firm conclusions on the (missing) effects of dropouts.
498 Conversely, there are several strengths to this study. This is the first study with a two-year follow-up
499 measurement in participants of CoolL: the outcomes in the long run, the nationwide inclusion and the broad
500 scope of the research provide valuable insights on the long-term effects of CoolL. In addition, the study is
501 based on data provided by people that participate in CoolL in a real-life setting. As CoolL is part of basic
502 health care insurance, it is accessible to everybody meeting the criteria and the outcomes are generalizable to
503 those participating in real life.

505 **Recommendations for future research**

506 This research provides an overview of the changes over a two-year time frame of the participants of CoolL,
507 showing more and less expected outcomes. Our recommendations for future research seize on these current
508 outcomes:

- 509 • More research on the two-year follow-up of CLIs in the Netherlands.
- 510 • More research into the optimal group size for health interventions, in support of explicit guidelines
511 for the healthcare workers.
- 512 • In-depth research into dropouts of the CLI, providing an overview of risk factors for dropout as well
513 as recommendations to prevent dropout.

514 **Conclusion**

515 The effects of CoolL on its participants show sustained and even enlarged weight loss when comparing phase
516 1 to phase 2 of CoolL, supporting the two-year set-up of CoolL with frequent contact moments and more
517 attention for behavioral maintenance in the second part of the intervention. The outcomes indicate that CoolL

518 is appropriate and effective for different group sizes and for a wide variety of participants irrespective of
519 gender, age, BMI at baseline or level of education.

521 Acknowledgements

522 The authors like to thank the Cool-coaches and the Cool-participants for their efforts to make this research
523 possible.

525 References

- 526 1. James WPT. WHO recognition of the global obesity epidemic. *International Journal of Obesity*.
527 2008;32(7):S120-S6.
- 528 2. Overgewicht en obesitas. Den Haag: Gezondheidsraad 2003
- 529 3. Luppino FS, de Wit LM, Bouvy PF, Stijnen T, Cuijpers P, Penninx BW, et al. Overweight, obesity,
530 and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of general*
531 *psychiatry*. 2010;67(3):220-9.
- 532 4. Taylor VHAPMD, Forhan MAPP, Vigod SNAPMD, McIntyre RSPMD, Morrison KMAPMD. The
533 impact of obesity on quality of life. *Best Practice & Research Clinical Endocrinology & Metabolism*.
534 2013;27(2):139-46.
- 535 5. How many adults are overweight? The Hague: Central Bureau of Statistics; 2022 [Available from:
536 <https://longreads.cbs.nl/nederland-in-cijfers-2022/hoeveel-volwassenen-hebben-overgewicht/>.
- 537 6. Organization WH. Factsheet Obesity and overweight 2021 [updated 09-06-2021. Available from:
538 <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
- 539 7. Bos V, van Dale D, Leenaars K. Werkzame elementen van Gecombineerde Leefstijlinterventies.
540 Rijksinstituut voor Volksgezondheid en Milieu; 2019.
- 541 8. Janssen E, Philippens N, Kremers S, Crutzen R. The combined lifestyle intervention CoolL in times
542 of COVID-19: a descriptive case series study. 2023.
- 543 9. Wing RR. Long-term effects of a lifestyle intervention on weight and cardiovascular risk factors in
544 individuals with type 2 diabetes mellitus: four-year results of the Look AHEAD trial. *Arch Intern Med*.
545 2010;170(17):1566-75.
- 546 10. Duijzer G, Haveman-Nies A, Jansen SC, Beek JT, van Bruggen R, Willink MGJ, et al. Effect and
547 maintenance of the SLIMMER diabetes prevention lifestyle intervention in Dutch primary healthcare: a
548 randomised controlled trial. *Nutrition & diabetes*. 2017;7(5):e268.
- 549 11. Wikström K, Peltonen M, Eriksson JG, Aunola S, Ilanne-Parikka P, Keinänen-Kiukaanniemi S, et al.
550 Educational attainment and effectiveness of lifestyle intervention in the Finnish Diabetes Prevention Study.
551 *Diabetes Res Clin Pract*. 2009;86(1):e1-5.
- 552 12. Wardle J, Steptoe A. Socioeconomic differences in attitudes and beliefs about healthy lifestyles. *J*
553 *Epidemiol Community Health*. 2003;57(6):440-3.
- 554 13. Burgess E, Hassmén P, Pumpa KL. Determinants of adherence to lifestyle intervention in adults with
555 obesity: a systematic review. *Clin Obes*. 2017;7(3):123-35.
- 556 14. Moore CJ, Cunningham SA. Social position, psychological stress, and obesity: a systematic review. *J*
557 *Acad Nutr Diet*. 2012;112(4):518-26.
- 558 15. Barte JC, Veldwijk J, Teixeira PJ, Sacks FM, Bemelmans WJ. Differences in weight loss across
559 different BMI classes: a meta-analysis of the effects of interventions with diet and exercise. *Int J Behav*
560 *Med*. 2014;21(5):784-93.

- 561 16. Kooloos JGM, Klaassen T, Vereijken M, Van Kuppeveld S, Bolhuis S, Vorstenbosch M.
562 Collaborative group work: Effects of group size and assignment structure on learning gain, student
563 satisfaction and perceived participation. *Medical Teacher*. 2011;33(12):983-8.
- 564 17. Lohman MC, Finkelstein M. Designing groups in problem-based learning to promote problem-
565 solving skill and self-directedness. *Instructional Science*. 2000;28:291-307.
- 566 18. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care*.
567 2002;25(12):2165-71.
- 568 19. Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, Williamson D, et al. The Look AHEAD
569 study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring)*.
570 2006;14(5):737-52.
- 571 20. Protocol BMI en middelomtrek meten: Nederlands Huisartsen Genootschap 2016 [cited 2023
572 January 12]. Available from: <https://www.cvrmindehuisartsenpraktijk.nl/downloads>.
- 573 21. Eklund M, Erlandsson L-K, Hagell P. Psychometric properties of a Swedish version of the Pearlin
574 Mastery Scale in people with mental illness and healthy people. *Nordic Journal of Psychiatry*.
575 2012;66(6):380-8.
- 576 22. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A
577 new instrument for psychiatric practice and research. *Psychiatry Research*. 1989;28(2):193-213.
- 578 23. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. *Journal of Health and
579 Social Behavior*. 1983;24(4):385-96.
- 580 24. Crutzen R, Peters G-JY. Scale quality: alpha is an inadequate estimate and factor-analytic evidence is
581 needed first of all. *Health psychology review*. 2017;11(3):242-7.
- 582 25. Lipsey MW. Design sensitivity : statistical power for experimental research. Newbury Park, CA:
583 Sage Publications; 1990.
- 584 26. Oosterhoff M, de Weerd A, Feenstra T, de Wit A. Jaarrapportage monitor GLI 2022. Stand van
585 zaken gecombineerde leefstijlinterventie. Annual report – Monitor Combined lifestyle intervention 2022
586 Combined lifestyle intervention progress report: Rijksinstituut voor Volksgezondheid en Milieu RIVM;
587 2022.
- 588 27. Elfhag K, Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors
589 associated with weight loss maintenance and weight regain. *Obes Rev*. 2005;6(1):67-85.
- 590 28. Skender ML, Goodrick GK, Del Junco DJ, Reeves RS, Darnell L, Gotto AM, et al. Comparison of 2-
591 year weight loss trends in behavioral treatments of obesity: diet, exercise, and combination interventions. *J
592 Am Diet Assoc*. 1996;96(4):342-6.
- 593 29. Jehn ML, Patt MR, Appel LJ, Miller ER, 3rd. One year follow-up of overweight and obese
594 hypertensive adults following intensive lifestyle therapy. *J Hum Nutr Diet*. 2006;19(5):349-54.
- 595 30. Lindström J, Louheranta A, Mannelin M, Rastas M, Salminen V, Eriksson J, et al. The Finnish
596 Diabetes Prevention Study (DPS): Lifestyle intervention and 3-year results on diet and physical activity.
597 *Diabetes Care*. 2003;26(12):3230-6.
- 598 31. van Rinsum C, Gerards S, Rutten G, Philippens N, Janssen E, Winkens B, et al. The Coaching on
599 Lifestyle (Cool) Intervention for Overweight and Obesity: A Longitudinal Study into Participants' Lifestyle
600 Changes. *International journal of environmental research and public health*. 2018;15(4).
- 601 32. Bukman AJ, Teuscher D, Meershoek A, Renes RJ, van Baak MA, Feskens EJ. Effectiveness of the
602 MetSLIM lifestyle intervention targeting individuals of low socio-economic status and different ethnic
603 origins with elevated waist-to-height ratio. *Public health nutrition*. 2017;20(14):2617-28.
- 604 33. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic
605 review of the literature. *Obes Rev*. 2011;12(11):912-34.
- 606

608 Supporting information

609 **S1 Table. Predictors for dropout.** Logistic regression outcomes on predictors for dropout.

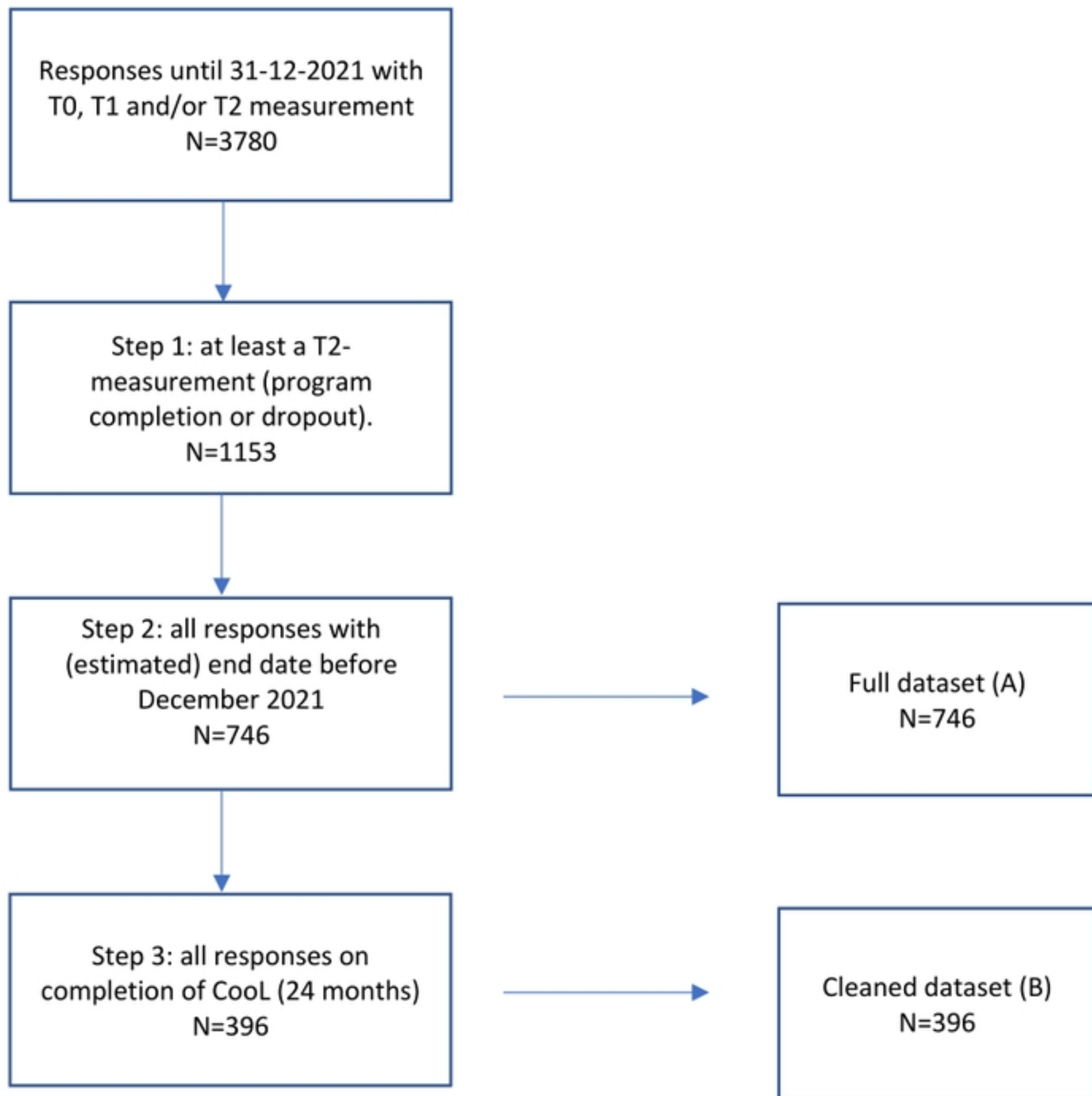
610 **S2 Table. Changes over time Cool finishers.** Detailed overview on changes over time on participants that
611 completely finished Cool.

612 **S3 Table. Subgroup comparison by educational level.** Detailed overview of the subgroup comparison on
613 educational level.

614 **S4 Table. Subgroup comparison by BMI category.** Detailed overview of the subgroup comparison on
615 BMI category.

616 **S5 Table. Subgroup comparison by group size.** Detailed overview of the subgroup comparison of larger
617 group size versus smaller group size.

618



Figure