

## Supplementary material that accompanies the paper “Subpopulation treatment effect pattern plot (STEPP) methods with R and Stata”

SERGIO VENTURINI<sup>1,\*</sup>, MARCO BONETTI<sup>2,3,4</sup>, ANN A. LAZAR<sup>5,6</sup>, BERNARD F. COLE<sup>7</sup>, XIN VICTORIA WANG<sup>8,9</sup>, RICHARD D. GELBER<sup>8,9</sup>, AND WAI-KI YIP<sup>10</sup>

<sup>1</sup>*Department of Economic and Social Sciences, Università Cattolica del Sacro Cuore, Via Bissolati 74, 26100 Cremona, Italy*

<sup>2</sup>*Carlo F. Dondena Research Centre, Università Commerciale L. Bocconi, Via Röntgen 1, 20136 Milano, Italy*

<sup>3</sup>*Bocconi Institute for Data Science and Analytics, Università Commerciale L. Bocconi, Via Röntgen 1, 20136 Milano, Italy*

<sup>4</sup>*Department of Social and Political Sciences, Università Commerciale L. Bocconi, Via Röntgen 1, 20136 Milano, Italy*

<sup>5</sup>*Division of Oral Epidemiology, Department of Preventive and Restorative Dental Sciences, University of California, San Francisco, 3333 California Street, San Francisco, CA 94118, USA*

<sup>6</sup>*Division of Biostatistics, Department of Epidemiology and Biostatistics, University of California, San Francisco, 3333 California Street, San Francisco, CA 94118, USA*

<sup>7</sup>*Department of Mathematics and Statistics, University of Vermont, 33 Colchester Avenue, Burlington, VT 05405, USA*

<sup>8</sup>*Department of Data Science, Dana-Farber Cancer Institute, 450 Brookline Avenue, Mailstop CLS 11007, Boston, MA 02215, USA*

<sup>9</sup>*Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02215, USA*

<sup>10</sup> *Agenus, Inc., 3 Forbes Road, Lexington, MA 02421, USA*

### Efficacy of aspirin in preventing colorectal adenomas

Here, we present a further STEPP analysis using the data originated from the Aspirin/Folate Polyp Prevention Study, a randomized, double-blind, placebo-controlled trial of the efficacy of oral aspirin, folic acid, or both as a chemoprevention agent against colorectal adenomas ([Baron et al., 2003](#)). Here, we limit the analysis to the aspirin component of the study. As usual, the aim of the analysis is the investigation of potential treatment-covariate interaction. More specifically, in this example the outcome corresponds to the occurrence of any pathologically confirmed colorectal adenomas and we are interested in assessing the interaction between the treatment effect and patients' age. A total of 1121 participants were randomized to two aspirin groups, that is 81 mg/day and 325 mg/day, or to the placebo arm. For brevity, we consider only the 81 mg/day arm (a full analysis is available in [Yip et al., 2016](#)). Participants were then followed for 3 years. Findings in the original study report concluded that low-dose aspirin had a moderate chemopreventive effect against colorectal adenomas.

After removing the missing values, we create the outcome and treatment indicators:

```
data("aspirin", package = "stepp")
aspirin_c <- aspirin[complete.cases(aspirin), ]
aspirin_sub <- subset(aspirin_c, DOSE != 325)
aspirin_sub$Y <- as.numeric(aspirin_sub$AD == 1 | aspirin_sub$AL == 1)
aspirin_sub$trt <- ifelse(aspirin_sub$DOSE == 0, 0, 1)
```

---

\*Corresponding author Email: [sergio.venturini@unicatt.it](mailto:sergio.venturini@unicatt.it).

Then, we generate the subpopulations using a unit-based sliding window approach with  $r_1 = 30$  and  $r_2 = 100$ , which produce 8 subpopulations:

```
swin <- new("stwin", type = "sliding", r1 = 30, r2 = 100)
subp <- new("stsubpop")
subp <- generate(subp, win = swin, cov = aspirin_sub$AGE)
summary(subp)
```

Window type: sliding

Number of patients per subpopulation (patspop r2): 100

Largest number of patients in common among consecutive subpopulations (minpatspop r1): 30

Number of subpopulations created: 8

Subpopulation summary information (including all treatments)

Subpopulation	Covariate Summary			Sample size
	Median	Minimum	Maximum	
1	43.00	29.0000	47.0000	107
2	50.00	46.0000	51.0000	128
3	53.00	52.0000	55.0000	118
4	57.00	55.0000	59.0000	127
5	61.00	59.0000	62.0000	115
6	64.00	62.0000	66.0000	103
7	68.00	66.0000	71.0000	101
8	73.00	70.0000	78.0000	89

As usual, one may further inspect the impact of these choices through a sensitivity analysis that involves both  $r_1$  and  $r_2$ .

Since the outcome in this application is binary, we run a STEPP analysis for the following `stmodelGLM` model with logit link,

$$\text{logit}(p) = \beta_0 + \beta_1 \times \text{treatment},$$

where  $p$  is the risk of developing a colorectal adenomas, and produce the corresponding plots (see Figure 1):

```
res <- new("steppes")
modelGLM <- new("stmodelGLM", coltrt = aspirin_sub$trt,
  colY = aspirin_sub$Y, glm = "binomial",
  trts = sort(unique(aspirin_sub$trt)))

res <- estimate(res, subp, modelGLM)
set.seed(101)
nperm <- 2500
res <- test(res, nperm)
print(res, estimate = TRUE, cov = FALSE, test = TRUE)

plot(res, ylabel = "Risk of Colorectal Adenomas",
  xlabel = "Median Age in Subpopulations",
  tlegend = c("Placebo", "81 mg/day aspirin"), legendy = 0.5,
  pointwise = FALSE, ylimit = c(0, 0.7, -0.3, 0.6, 0, 3),
  lsty = c(2, 1), marker = c(15, 17), ncex = 0.8, at = 8,
  legend_diff = c(1, 2))
```

Sample size in treatment 0: 363

Sample size in treatment 1: 366

Total sample size (excluding missing data): 729

Risk estimates for treatment group 0

Subpopulation	Risk	Std. Err.
1	0.2979	0.0667
2	0.3288	0.0550
3	0.5517	0.0653
4	0.6379	0.0631
5	0.5161	0.0635
6	0.4906	0.0687
7	0.5102	0.0714
8	0.5833	0.0712
Overall	0.4711	0.0262

Risk estimates for treatment group 1

Subpopulation	Risk	Std. Err.
1	0.3833	0.0628
2	0.3818	0.0655
3	0.3333	0.0609
4	0.2754	0.0538
5	0.3585	0.0659
6	0.4600	0.0705
7	0.4423	0.0689
8	0.5366	0.0779
Overall	0.3825	0.0254

Effect differences and ratio estimates

trt 0 vs. trt 1

Risk differences

Subpopulation	Risk Difference	Std. Err.
1	-0.0855	0.0916
2	-0.0531	0.0855
3	0.2184	0.0893
4	0.3626	0.0829
5	0.1576	0.0915
6	0.0306	0.0984
7	0.0679	0.0992
8	0.0467	0.1055
Overall	0.0886	0.0365

Odds ratios

Subpopulation	log Odds Ratio	Std. Err.	Odds Ratio
1	-0.3820	0.4150	0.6825
2	-0.2319	0.3730	0.7930
3	0.9008	0.3804	2.4615
4	1.5340	0.3838	4.6366
5	0.6465	0.3829	1.9088
6	0.1226	0.3950	1.1304
7	0.2726	0.3995	1.3134
8	0.1899	0.4287	1.2091
Overall	0.3631	0.1504	1.4377

Supremum test results

trt 0 vs. trt 1

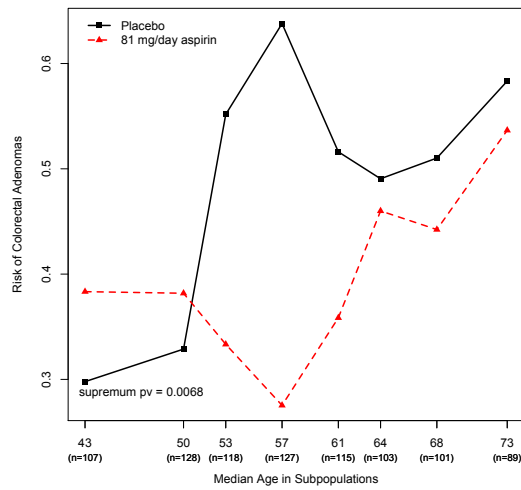
Interaction p-value based on effect estimate differences: 0.0068

Chi-square interaction p-value based on effect estimate differences: 0.0432

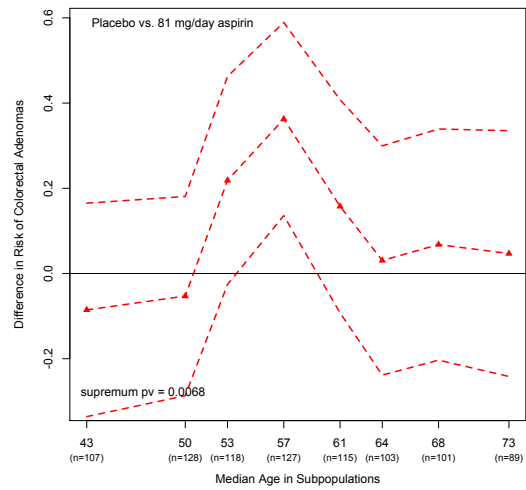
Both the output and the figures show that the risk for the placebo group is higher than the risk for the treatment group for all subpopulations apart from the first two. In particular, the risk for the placebo group rises quickly with age, while the risk for the treatment group first decreases but then starts increasing too, and the two converge at around age 60. In addition, the largest differences are observed in correspondence of the middle-aged subgroups. Similarly, the odds ratios are in favor of the treatment group in the same subpopulations. We also note that the p-value for the supremum test is very small signaling that the data provide evidence for heterogeneity in the treatment effects among the subpopulations. On the one hand, these results confirm the original study findings; that is, the 81 mg/day aspirin dose reduces the risk of adenomas compared with placebo. On the other hand, they also suggest that patients benefiting the most from the low-dose aspirin are those aged around 57 years.

## References

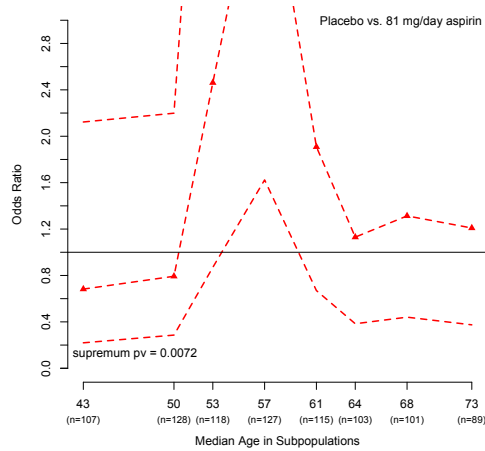
- Baron JA, Cole BF, Sandler RS, Haile RW, Ahnen D, Bresalier R, et al. (2003). A Randomized Trial of Aspirin to Prevent Colorectal Adenomas. *The New England Journal of Medicine*, 348(10): 891–899.
- Yip WK, Bonetti M, Cole BF, Barcella W, Wang XV, Lazar A, et al. (2016). Subpopulation Treatment Effect Pattern Plot (STEPP) Analysis for Continuous, Binary, and Count Outcomes. *Clinical Trials*, 13(4): 382–390.



(a)



(b)



(c)

Figure 1: aspirin data. Plot of: (a) risk estimates of developing colorectal adenomas, (b) absolute treatment effect measured as the difference in the risk of developing colorectal adenomas estimates (aspirin minus placebo; a value below zero suggests that aspirin is better), (c) relative treatment effect measured as the odds ratio estimates (aspirin vs. placebo; a value less than one suggests that aspirin is better). Panels (b) and (c) also report the corresponding 95% confidence regions.