Homework 1: Categorical Data Analysis

Asci Onur Dewilde Brecht Vandenbulcke Stijn

1 Question 1

An **exact binomial test** at a 5% significance level is conducted to examine the following hypothesises:

$$H_0 : \pi = 0.87$$
$$H_a : \pi > 0.87$$

During 20 trials, 20 successes were observed. Table 1 contains, for each value x, the probability that there are at least as much successes as x in case 20 trials are performed. It can be seen that none of these probabilities fall below the significance level $\alpha = 0.05$. Consequently, the 'acceptance region' contains the complete set of possible successes (0 - 20) and thus the rejection region is empty. Furthermore it can be concluded that the obtained value of 20 successes lies within the acceptance region and thus the null hypothesis will not be rejected.

Table 1:

x	P(X = x)	P(X $\geq$ x)
0	1.9 e^{-18}	1.0000
1	2.54 e^{-16}	1.0000
...	...	...
17	0.2347	0.7426
18	0.2618	0.5079
19	0.1844	0.2461
20	0.0617	0.0617

The 95% confidence interval for the estimate of the success probability for each trial of a binomial distribution if 20 success where observed during 20 trials is [0.8316, 1]. This interval is obtained by the use of the definition of Clopper and Pearson. The **exact coverage probability** for a 95% confidence interval for the probability of successes for a sample of size 20 when the population probability $\pi = 0.87$ is 0.9897 when using the Clopper-Pearson method. This coverage probability is 0.9279 when the Wald method is used. These coverage probabilities illustrate that the Clopper-Pearson CI is indeed too conservative and that the coverage probabilities for Wald CI are indeed too low.

2 Question 2

To test whether there is a (positive) effect due to the tutor-supervised learning program, a two-proportion z-test is used. This test is used to compare two observed proportions. The null and alternative hypothesis are formulated by the use of p1 and p2. P1 represents the proportion of students that where successful in their 1st year and participated in the tutor-supervised learning program. P2 represents the proportion of students that where successful but did not participate in the tutor-supervised learning program.

$$H_0 : p_1 - p_2 = 0$$

$$H_a : p_1 - p_2 \geq 0$$

The estimated probability of success for the students that followed the tutor supervised learning program is 0.50 and for the students that didn't 0.25. The obtained p-value of the z-test is 0.0614. This p-value is higher than the significance level $\alpha = 0.05$,therefore we conclude that the proportion of people that succeeds is not significantly different between the two groups. If we use this sample to create a 95% confidence interval of the difference between both proportions, then we obtain [-0.0334, 1.0000]. Note that the Yates continuity correction is not used as we do not specifically want the results to be conservative.

Given the obtained results, we can conclude that the programme has a positive effect on the success rate of the first year students. Moreover, there is a difference of 25% between the estimated probabilities of both groups. Nevertheless, the two-proportion z-test does not consider this difference as significant, however it is very close. This can likely be explained by the fact that the number of students participating the programme (n1) is much lower than the number of students that didn't ask for support (n2). Therefore we suggest to extend the program by one more year, such that more data will be available to obtain a more profound conclusion.

3 Question 3

To test the hypothesis that the drug 'STOP-HEADACHE' will increase headache relief after 2 hours, compared to a placebo, a logistic regression model (model 1) was fitted. This model uses treatment (trt), baseline severity (bassev), gender and age as predictors. Whereby the variables bassev and trt are coded by the use of dummy variables.

Bassev has four levels, 0 (no pain), 1 (mild pain), 2 (moderate pain) and 3 (severe pain). In this dataset only observations with bassev 2 and 3 are included as only patients with a bassev of 2 or 3 were treated for their headache. *Relief* has two levels, 0, which states that there was relief after

treatment and 1 when there was no relief. There was relief if the bassev went from 2 or 3 to 0 or 1. *Treatment* has three levels, placebo (trt = 0), 50 mg (trt = 50) and 100 mg (trt = 100). The variable *age* has two levels 1 (male) and 2 (female). *Age* is a continuous variable ranging from 18 to 80.

The mathematical representation of **model 1**:

$$Logit(Y_i) = \beta_0 + \beta_1 Age + \beta_2 trt_{50} + \beta_3 trt_{100} + \beta_4 gender2 + \beta_5 bassev3 + \varepsilon_i$$

where:

$$trt_{50} = \begin{cases} 1 \text{ treatment 50 mg is applied} \\ 0 \text{ otherwise} \end{cases} \qquad gender2 = \begin{cases} 1 \text{ gender is female} \\ 0 \text{ gender is male} \end{cases}$$

$$trt_{100} = \begin{cases} 1 \text{ treatment 100 mg is applied} \\ 0 \text{ otherwise} \end{cases} \qquad bassev3 = \begin{cases} 1 \text{ baseline severity is 3} \\ 0 \text{ baseline severity is 2} \end{cases}$$

Table 2 (appendix 5.2) contains the results of fitting model 1. It is visible that the intercept, age, trt50 and trt100 are significant. Moreover, the negative value for the age predictor suggests that the odds for relief are lower with older subjects. Taking the natural exponent of a parameter gives us more information about how the odds of relief increase. For example by taking the 100 mg treatment, the estimated odds of relief are 5.908 times higher than the relief of people that where treated with a placebo, while if you would take the 50 mg treatment, the odds are 4.0825 times higher. The parameters for gender2 and bassev3 are not significant, suggesting there is no significant influence of gender and severity on pain. However, there is still a possibility of interactions between predictors, this will be investigated below.

An ANOVA test to compare model 1 with a model that excludes the treatment predictor. The results of this test are presented in table 3 (appendix 5.2). They indicate that model 1 fits the data a lot better (Table 3).

With the R effects package we can compare the effect of both treatments in model 1 and calculate a 95% confidence intervals (CI) for this effect. Changing the treatment and keeping the other predictors constant, we obtain that the placebo group has a 0.1772 chance of relief, the 50 mg has a success rate of 0.4678 and the 100mg has a success rate of 0.5599. This is visible in figure 1 (appendix 5.3). The CI's for these success rates are respectively [0.1413, 0.2198], [0.4188, 0.5175] and [0.5089, 0.6097]. The confidence intervals of 50mg and 100mg overlap slightly, meaning hat we cannot be certain about the fact that there is a difference in effect of treatment with this model.

Three additional models where fitted to check the consistency of the treatment effect across baseline severity, gender and age by adding interaction terms to model 1. These models can be found in appendix 5.1.

Using the effects package we are able to see the effects of the interaction terms. The Trt50:Bassev3 interaction does not have a significant parameter, while the Trt100:Bassev3 does. We can see in Figure 2 (appendix 5.3) that when a subject has a bassev of 2 or 3, it doesn't matter how much mg they take. However when the pain is severe (bassev = 3), then the odds of relief are higher if 100mg is taken.

The Trt50:Gender2 and Trt100:Gender2 parameters are both insignificant, figure 3 (appendix 5.3) shows that the odds are similar for both genders and treatments. However, the confidence intervals are smaller for females.

The Age:Trt50 and Age:Trt100 parameters are insignificant as well. As can be seen on figure 4 (appendix 5.3), age has a significant effect on the odds, however there is no interaction between treatment. An older person has lowers odds of relief regardless of the treatment. The confidence intervals for the effects of the interaction terms can be found in the appendix.

4 Question 4

Rothwell et al. (2010) performed 8 trials investigating the effects of daily aspirin use on the on risk of death due to cancer. We are given the number of cancer related deaths and total number of participants in the control group and the treatment group.
Using these values, we can calculate the odds ratio and 95% confidence interval for each trial. In order to do this, we can treat the results of each trial as a 2x2 table and calculate $\hat{\theta}$ is computed as:

$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

Furthermore, we can calculate the standard error of the odds ratio of each trial as

$$SE(log\hat{\theta}) = (\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}})^{1/2}$$

And using this standard error we can calculate the 95% CI for each trial as

$$log\hat{\theta} \pm z_{\alpha/2}SE(log\hat{\theta})$$

Accordingly, the odds ratio and 95% CI for each trial can be seen in Table 10 (appendix 5.4).
We can get a pooled estimate for these 8 odds ratios by calculating the Mantel-Haenszel estimator as

$$\hat{\theta}_{MH} = \frac{\sum_{k=1}^{8}(\frac{n_{11k}n_{22k}}{n_{++k}})}{\sum_{k=1}^{8}(\frac{n_{12k}n_{21k}}{n_{++k}})}$$

Which gives us the pooled estimate $\hat{\theta}_{MH}$ 0.79 with 95% CI [0.68, 0.92] and p-value of 0.003.

We test the homogeneity of the odds ratios with the null hypothesis $H_0 : \theta : \theta_{XY(1)} = \theta_{XY(2)} = ... = \theta_{XY(8)}$ using the Breslow-Day test as

$$\sum_{k=1}^{8} (\frac{(n_{11k} - \hat{\mu}_{11k})^2}{Var(\hat{\mu}_{11k})})$$

According to the test result χ^2 (7) = 5.9117, p = .5501 we reject the null hypothesis.
We also conducted the same analyses by using logistic regression. First, we created a full model that consists of the main effects of treatment and trial and also their interaction. The exponentiated regression coefficient for UK-TIA trial gives us the odds ratio of 0.45, showing that the estimated odds of dying from cancer is 0.45 times higher in the control group compared to the treatment group. The odds ratio for ETDRS trial is 1.14, indicating 1.14 times higher estimated odds for dying from cancer in the treatment group.

Next, we calculated the standard error for the BDAT trial as 0.16 by multiplying the coefficient of this trial with the covariance matrix of the logistic regression model. Using this standard error, we calculated the exponentiated 95% CI as [0.85, 1.59].

Then, we created a logistic regression model with both main effects but without the interaction term and computed the pooled odds ratio estimate as 0.79 and 95% CI [0.68, 0.92] by using the exponentiated regression coefficients and confidence intervals. To calculate the associated p-value, we created another model with only the main effect of trial and compared it to the previous model which had both main effects by a Chi-square test. The result of this test yielded a p-value of 0.0029. The findings of the odds ratio, 95% CI and p-value for this model were equivalent to the results of the Mantel-Haenszel estimation we reported in the previous section.

Finally, we tested the homogeneity of the results of these trials by comparing the full model that included the interaction term and the model with only the main effects with a Chi-square test. This comparison yielded a p-value of 0.56, which is equivalent to the result of the Breslow-Day test we conducted in the previous section.

5 Appendix

5.1 Models with interaction terms for question 3

$$Logit(Y_i) = \beta_0 + \beta_1 Age + \beta_2 trt_{50} + \beta_3 trt_{100} + \beta_4 gender2 + \beta_5 bassev3 + \beta_6 bassev3 * Trt_{50} + beta_7 bassev3 * Trt_{100} + \varepsilon_i$$

$$Logit(Y_i) = \beta_0 + \beta_1 Age + \beta_2 trt_{50} + \beta_3 trt_{100} + \beta_4 gender2 + \beta_5 bassev3 + \beta_6 Gender2 * Trt_{50} + \beta_7 Gender2 * Trt_{100} + \varepsilon_i$$

$$Logit(Y_i) = \beta_0 + \beta_1 Age + \beta_2 trt_{50} + \beta_3 trt_{100} + \beta_4 gender2 + \beta_5 bassev3 + \beta_6 Age * Trt_{50} + beta_7 Age * Trt_{100} + \varepsilon_i$$

5.2 Tables for question 3

Table 2: Parameter estimates logistic regression model 1 STOP-HEADACHE

Coefficients	Estimate	Standard deviation	Z value	Pr(>—Z—)
Intercept	-0.959816	0.258645	-3.711	0.000206
Age	-0.013309	0.003475	-3.830	0.000128
trt50	1.406733	0.170487	8.251	<2*10^-16
trt100	1.776334	0.172794	10.280	<2*10^-16
gender2	-0.073582	0.159052	-0.463	0.643629
bassev3	0.214201	0.132032	1.622	0.104730

Table 3: Analysis of Deviance Table: Comparing model 1 with a model excluding the treatment term

Model 0: relief ∼age + gender + bassev

Model 1: relief ∼age + trt + gender + bassev

Resid. Df	Resid. Dev	Df	Deviance	Pr(>chi)
1136	1524.5			
1134	1395.0	2	129.53	<2.2*10^-16

Table 4: Parameter estimates logistic regression model STOP-HEADACHE with Treatment Bassev interaction term

Coefficients	Estimate	Standard Error	Z value	Pr(>—Z—)
Intercept	-0.605682	0.289432	-2.093	0.036380
Age	-0.013492	0.003495	-3.861	0.000113
Trt50	1.168917	0.261190	4.475	7.63*10^-6
Trt100	1.051267	0.266932	3.938	8.21*10^-5
Gender2	-0.044410	0.160333	-0.277	0.781788
Bassev3	-0.391774	0.276394	-1.417	0.156353
Trt50:Bassev3	0.385677	0.344658	1.119	0.263135
Trt100:Bassev3	1.201138	0.351762	3.415	0.000639

Table 5: Effects of the Treatment Bassev interaction term
relief ∼Age + trt + gender +bassev + trt*bassev

trt	bassev	Lower limit	Effect	Upper limit
0	2	0.1539612	0.2146440	0.2910175
50	2	0.3918975	0.4679841	0.5455892
100	2	0.3601867	0.4388355	0.5206820
0	3	0.1143203	0.1559166	0.2090752
50	3	0.4031657	0.4664663	0.5308654
100	3	0.5726995	0.6372561	0.6972159

Table 6: Parameter estimates logistic regression model STOP-HEADACHE with Treatment Gender interaction term

Coefficients	Estimate	Standard Error	Z value	Pr(>—Z—)
Intercept	-1.167810	0.360987	-3.235	0.001216
Age	-0.013305	0.003476	-3.828	0.000129
Trt50	1.537811	0.390615	3.937	8.25*10^-5
Trt100	2.193276	0.393120	5.579	2.42*10^-8
Gender2	0.193169	0.350875	0.551	0.581953
Bassev3	0.205912	0.132280	1.557	0.119557
Trt50:Gender2	-0.165110	0.433942	-0.380	0.703583
Trt100:Gender2	-0.524007	0.437418	-1.198	0.230934

Table 7: Effects of the Treatment Gender interaction term
relief ∼Age + trt + gender +bassev + trt*gender

trt	gender	Lower limit	Effect	Upper limit
0	1	0.09028394	0.1557945	0.2554899
50	1	0.35407702	0.4620614	0.5737267
100	1	0.51169607	0.6232622	0.7231304
0	2	0.1424814	0.1829198	0.2317337
50	2	0.4144314	0.4690427	0.5244056
100	2	0.4855650	0.5430381	0.5993878

Table 8: Parameter estimates logistic regression model STOP-HEADACHE with Treatment Age interaction term

Coefficients	Estimate	Standard Error	Z value	Pr(>—Z—)
Intercept	-0.588561	0.399866	-1.472	0.14105
Age	-0.021592	0.007800	-2.768	0.00563
Trt50	0.898984	0.461148	1.949	0.05124
Trt100	1.337805	0.478379	2.797	0.00517
Gender2	-0.071073	0.158981	-0.447	0.65484
Bassev3	0.213767	0.132005	1.619	0.10537
Age:Trt50	0.011111	0.009446	1.176	0.23949
Age:Trt100	0.009601	0.009655	0.994	0.32004

Table 9: Effects of the Treatment Age interaction term

relief ∼Age + trt + gender +bassev + trt*age

Effect

	Treatment		
Age	0	50	100
20	0.27944975	0.5433930	0.6416716
30	0.23810170	0.5172922	0.6136572
50	0.16868716	0.4649500	0.5554906
60	0.14053146	0.4389946	0.5257186
80	0.09597858	0.3882035	0.4658370

Lower 95% Confidence Limits

	Treatment		
Age	0	50	100
20	0.19698298	0.4559845	0.5480180
30	0.18099266	0.4497451	0.5403973
50	0.13247997	0.4157985	0.5045210
60	0.10218446	0.3819105	0.4671435
80	0.05452737	0.3002371	0.3701199

Upper 95% Confidence Limits

	Treatment		
Age	0	50	100
20	0.3800993	0.6282083	0.7256350
30	0.3064872	0.5842132	0.6821089
50	0.2123675	0.5147919	0.6053187
60	0.1902201	0.4977400	0.5835944
80	0.1634918	0.4841127	0.5641376

5.3 Figures for question 3

Figure 1:

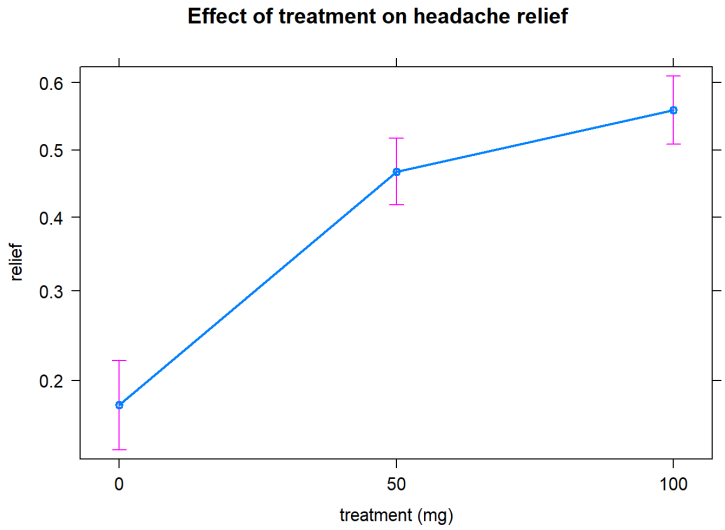


Figure 2:

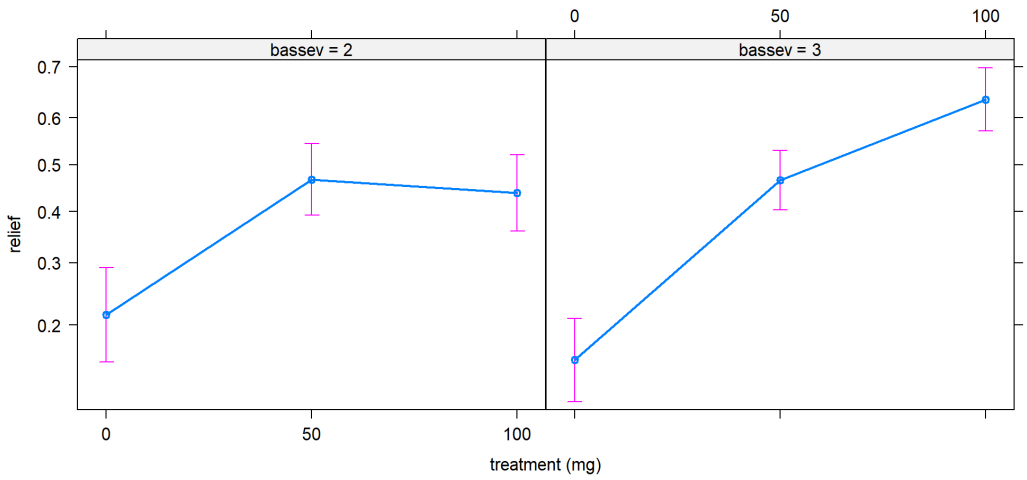


Figure 3:

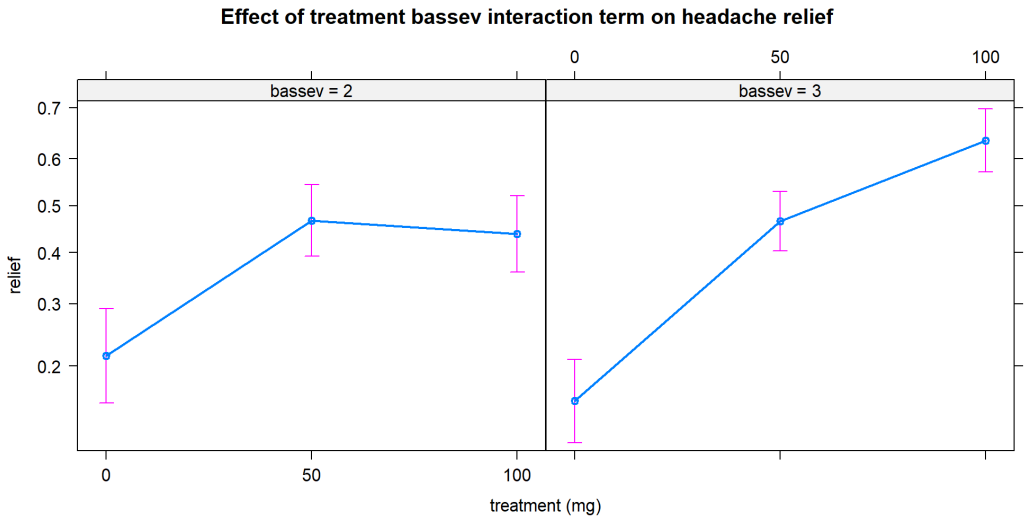


Figure 4:

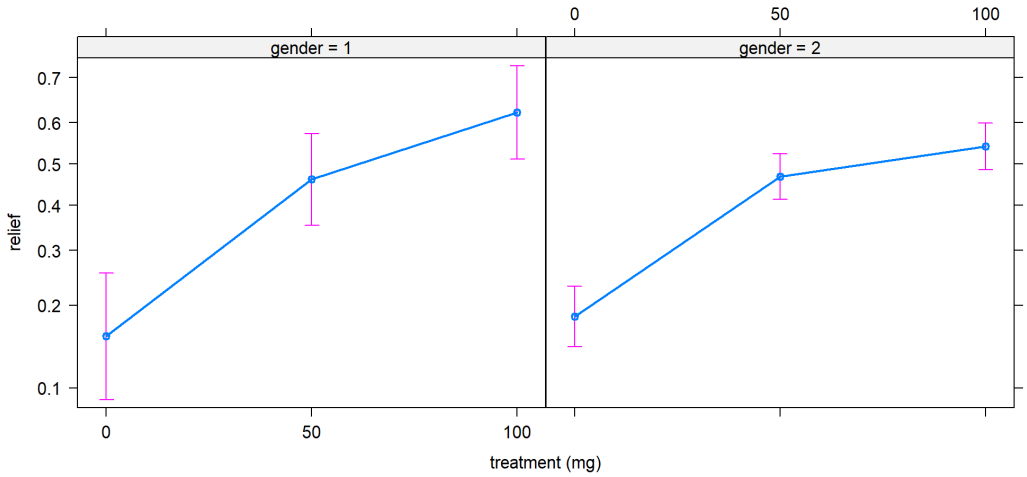
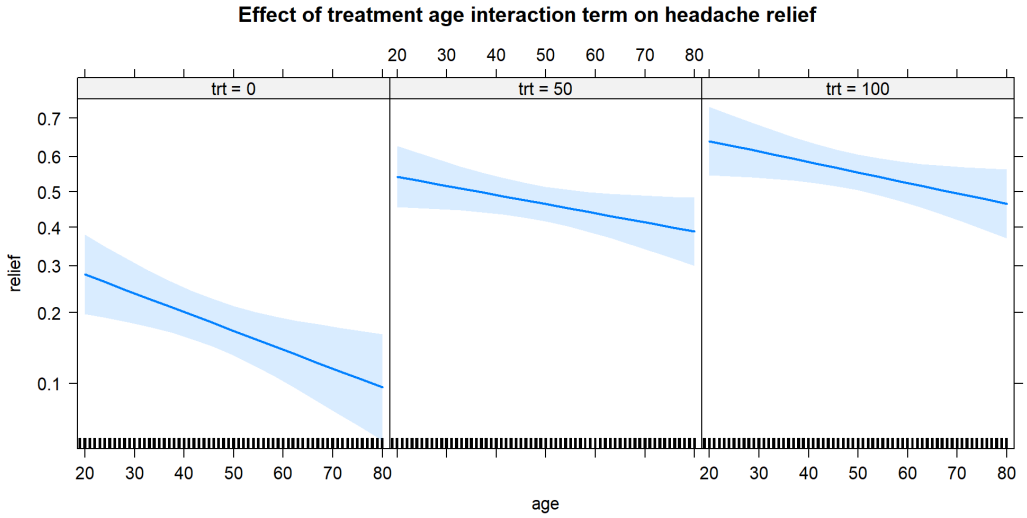


Figure 5:


Effect of treatment age interaction term on headache relief
trt = 0
trt = 50
trt = 100
relief
0.7
0.6
0.5
0.4
0.3
0.2
0.1
20
30
40
50
60
70
80
age

5.4 Table for question 4

Table 10: Odds ratio and 95% CI for each trial

Trial	Odds Ratio	95% CI
BDAT	0.79	0.54 - 1.15
UK-TIA	0.45	0.25 - 0.82
ETDRS	1.14	0.56 - 2.35
SAPAT	0.53	0.24 - 1.15
TPT	0.83	0.62 - 1.11
JPAD	0.80	0.40 - 1.57
POPADAD	0.80	0.47 - 1.37
AAA	0.86	0.63 - 1.17