

Semestrální práce do předmětu SME

Diana Varšíková

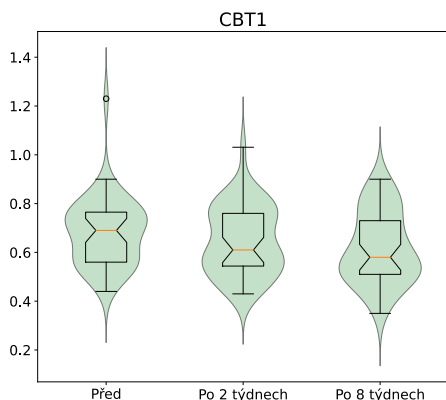
10. července 2023

1 Zadání

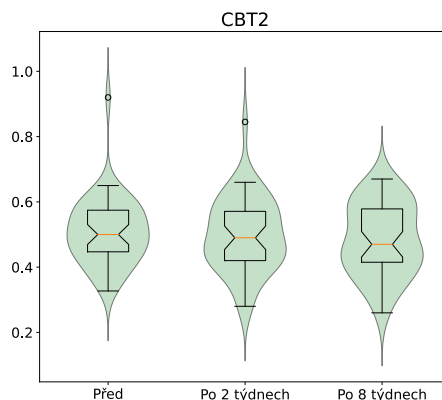
Byla provedena studie vlivu podávání jistého léku na pokles jistých rozměrových parametrů lidského oka (CBT1, CBT2, CBT3) a nitroočního tlaku (IOP). Všechny hodnoty byly změřeny na začátku léčby (baseline), po dvou a osmi týdnech léčby. Pomocí analýzy rozptylu jednoduchého třídění otestujte rozdílnosti hodnot parametrů CBT1, CBT2, CBT3 a IOP naměřených v jednotlivých návštěvách. Ověřte podmínky pro použití analýzy rozptylu a výsledky, včetně základních deskriptivních statistik (průměry, odchylky, odlehlá pozorování atd.), uveďte ve formě tabulek a grafů.

2 Popis dat

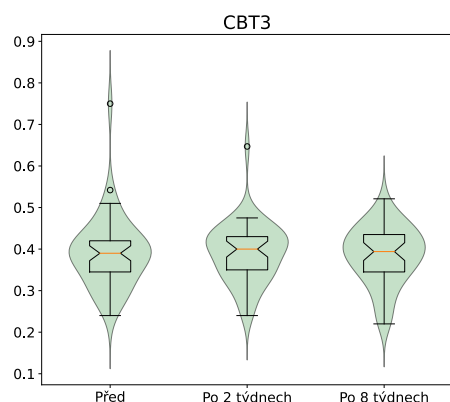
Měření bylo prováděno na celkem 43 pacientech. Bylo opakováno před začátkem podávání léků, 2 týdny po začátku a 8 týdnů po začátku. Cílem analýzy je otestovat rozdílnost hodnot parametrů CBT1, CBT2, CBT3 a IOP naměřených při jednotlivých návštěvách lékaře a tedy ověřit vliv léku na dané parametry. Pro první náhled na dané veličiny lze použít krabicové grafy.



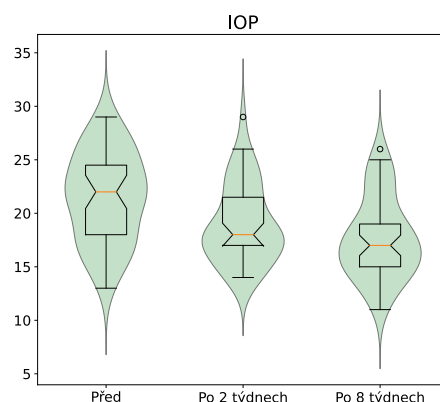
(a) Krabicový graf CBT1



(b) Krabicový graf CBT2



(a) Krabicový graf CBT3



(b) Krabicový graf IOP

Po jistém průzkumu se zdá, že veličiny CBT by mohly popisovat spektrální citlivosti lidského oka pro červenou, modrou a zelenou barvu. Tyto veličiny běžně nabývají hodnot od 0 do 1. V přiloženém datasetu se nachází dvě hodnoty vyšší než 1 (konkrétně 1.23 a 1.031) u dvou různých pacientů. Jelikož se zároveň jedná o odlehlé hodnoty od ostatních, tito dva pacienti budou z následné analýzy vynecháni. Ty odlehlé hodnoty, které jsou ale stále v rozmezí $[0,1]$ budou v datasetu zanechány. Zbývají tedy data o 41 pacientech.

Normální hodnoty nitroočního tlaku se pohybují v rozmezí 10-20 mm Hg, avšak při onemocnění očí, například zákalech, mohou dosahovat i vyšších hodnot. Při hodnotách nitroočního tlaku nad 30 se jedná již o vážné akutní zákaly. Pro nitrooční tlak se tedy žádné odlehlé pozorování nevynechalo.

Z krabicových grafů jde vidět, že na začátku mají pacienti daleko vyšší hodnoty nitroočního tlaku než běžná populace. Dále je patrné že IOP s časem klesá, což by mohlo naznačovat, že pacienti pozitivně reagují na lék. Pro účely analýzy jsou veličiny označeny předponami W0, W2 a W8. V následující tabulce jsou shrnuty deskriptivní statistiky po vyřazení odlehlých pozorování.

Tabulka 1: Deskriptivní statistiky po odstranění odlehlých pozorování

	CBT1			CBT2			CBT3			IOP		
	W0	W2	W8	W0	W2	W8	W0	W2	W8	W0	W2	W8
průměr	0.67	0.64	0.62	0.50	0.49	0.49	0.39	0.39	0.39	21.46	19.05	17.76
std	0.12	0.13	0.15	0.08	0.10	0.11	0.09	0.06	0.07	4.30	3.67	3.77
min	0.44	0.43	0.35	0.33	0.28	0.26	0.24	0.24	0.22	13.00	14.00	11.00
25%	0.56	0.54	0.50	0.44	0.42	0.41	0.34	0.35	0.35	18.00	17.00	15.00
50%	0.69	0.60	0.58	0.50	0.49	0.48	0.39	0.40	0.40	22.00	18.00	17.00
75%	0.76	0.75	0.75	0.57	0.57	0.58	0.42	0.43	0.44	25.00	22.00	19.00
max	0.90	0.83	0.90	0.65	0.66	0.67	0.75	0.48	0.52	29.00	29.00	26.00

3 ANOVA - předpoklady

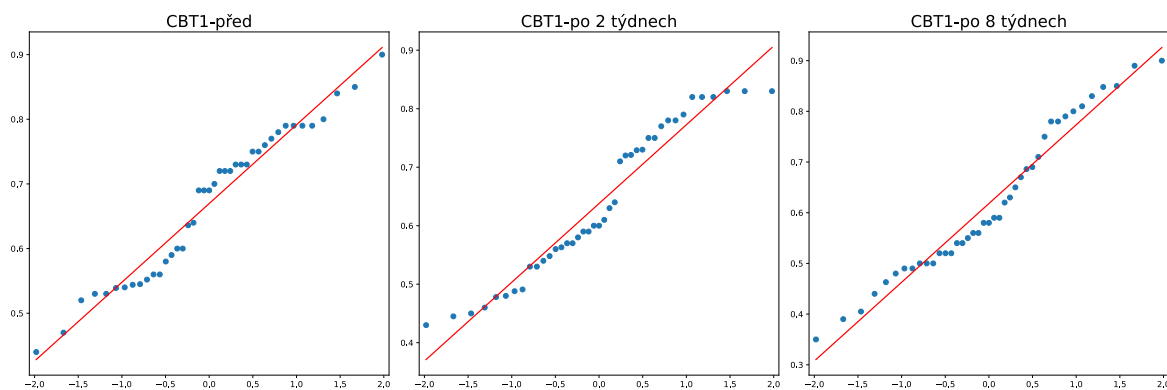
Pro ověření rozdílnosti parametrů před a po podávání léku použijeme metodu ANOVA, jejíž hypotézou je rovnost středních hodnot při jednotlivých návštěvách. Alternativou je nerovnost středních hodnot. V případě zamítnutí hypotézy je tedy nutné ještě vyšetřit, které hodnoty zamítnutí způsobily.

Před použitím samotné metody je nutno ověřit předpoklady pro její funkčnost. Je tedy pro každou ze 4 trojic potřeba ověřit následující :

1. normalitu dat
2. stejné rozptyly dat
3. nezávislost dat

Vzhledem k tomu, že se díváme na vývoj jedné testovací skupiny v čase, tak neplatí předpoklad nezávislosti výběrů a vhodnější by bylo použít Repeated Measures ANOVA metodu. Zde bude ale použita pouze jednofaktorová ANOVA a následně správnost nalezených výsledků ověříme ještě párovým t-testem.

K ověření normality dat použijeme pro každou z 12 veličin QQ plot a Lillieforsův test. Výpočet testu se provede funkcí `lilliefors` z Python knihovny `statsmodels`.

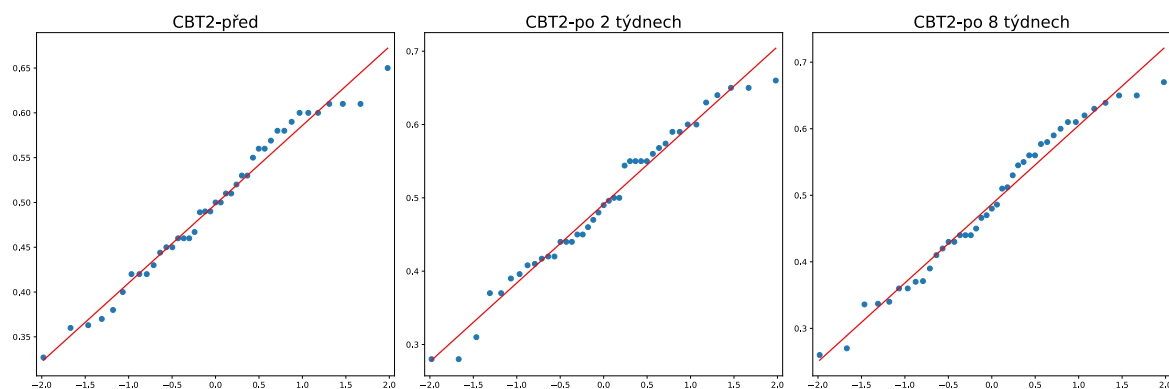


Obrázek 3: QQ plot pro CBT1

Tabulka 2: Výsledky Lillieforsova testu pro CBT1

týden	KS-statistika	p-hodnota
W0	0.1314	0.0718
W2	0.1278	0.0901
W8	0.1364	0.0525

Z QQ plotu lze vidět mírné odchylky od normálního rozdělení, především u první veličiny W0CBT1. Pro upřesnění použijeme Lillieforsův test. Z něj vidíme, že na hladině 0.05 nezamítáme hypotézu, že data W0CBT1, W0CBT2 a W0CBT3 mají normální rozdělení.

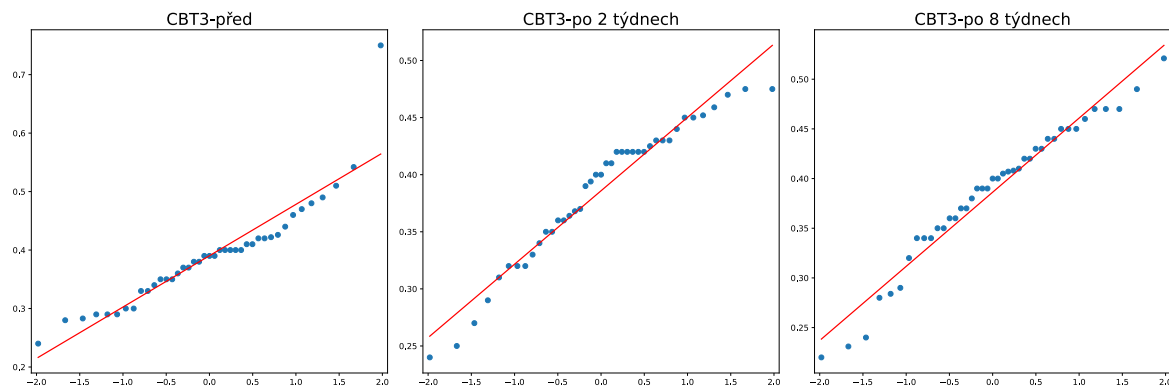


Obrázek 4: QQ plot pro CBT2

Tabulka 3: Výsledky Lillieforsova testu pro CBT2

týden	KS-statistika	p-hodnota
W0	0.0899	0.5466
W2	0.1132	0.2116
W8	0.0897	0.5489

Pro CBT2 nejsou na QQ plotu zřejmé velké výchylky oproti normálnímu rozdělení, snad kromě pravého chvostu. Z Lillieforsova testu opět vidíme, že na hladině 0.05 normalitu nezamítáme.

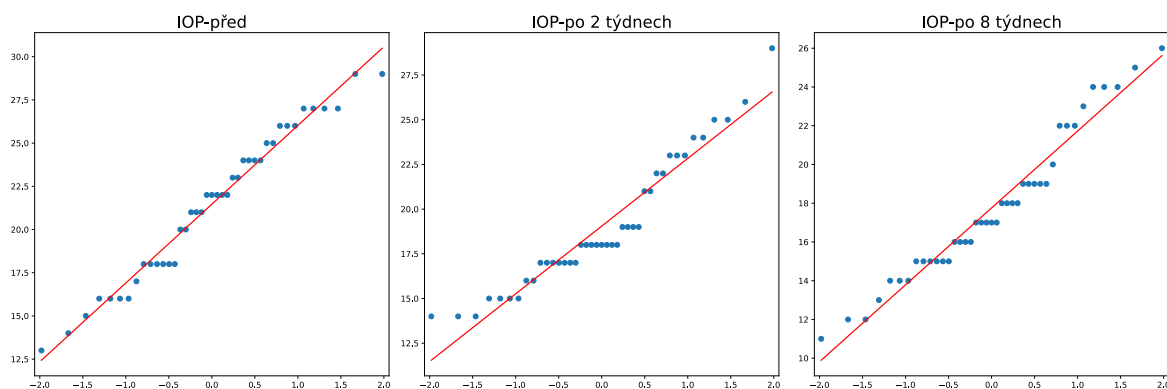


Obrázek 5: QQ plot pro CBT3

Tabulka 4: Výsledky Lillieforsova testu pro CBT3

týden	KS-statistika	p-hodnota
W0	0.1482	0.0238
W2	0.1494	0.0223
W8	0.1064	0.2858

V QQ plotu pro W0CBT3 vidíme jedno velmi se vychylující pozorování. Obě p-hodnoty pro W0CBT3 i pro W2CBT3 jsou 0.02, tudíž Lillieforsův test zamítá na hladině 0.05 normalitu pro W0CBT3 i W2CBT3. Pro CBT3 tedy nebude možné použít ANOVA metodu.



Obrázek 6: QQ plot pro IOP

Tabulka 5: Výsledky Lillieforsova testu pro IOP

týden	KS-statistika	p-hodnota
W0	0.1310	0.0734
W2	0.1979	0.0009
W8	0.1269	0.0946

Pro veličinu W2IOP opět zamítáme hypotézu o normálním rozdělení na hladině 0.05. Pro veličiny IOP a CBT3 bude tedy nutné použít neparametrickou verzi ANOVA, tedy Kruskalův-Wallisův test.

Jako poslední je potřeba otestovat rovnost rozptylů v daných skupinách. Pro normálně rozdělená data bude pro test rovnosti rozptylů použit Bartlettův test, který se nachází v balíku *scipy.stats*.

Tabulka 6: Výsledky Bartlettova testu

veličiny	testovací statistika	p-hodnota
W0CBT1, W2CBT1, W8CBT1	2.3455	0.3095
W0CBT2, W2CBT2, W8CBT2	3.5366	0.1706
W0CBT3, W2CBT3, W8CBT3	5.3946	0.0673
W0IOP, W2IOP, W8IOP	1.1761	0.5554

Podle Bartlettova testu nezamítáme pro žádné veličiny na hladině 0.05 hypotézu o rovnosti rozptylů. Pro data, pro něž jsme zamítli hypotézu o normálním rozdělení bude použit ještě Levene test, který slouží pro porovnání rozptylů dat, jež nejsou normálně rozdělena. Opět se nachází v balíku *scipy.stats*.

Tabulka 7: Výsledky Leveneho testu

veličiny	testovací statistika	p-hodnota
CBT3	0.4516	0.6376
IOP	1.1727	0.3130

Ani tento test nezamítá na hladině 0.05 rovnost rozptylů, tudíž předpoklad rovnosti rozptylů je splněn u všech 4 veličin.

4 ANOVA

Pro ověření vlivu léku na dané veličiny (CBT1, CBT2, CBT3, IOP) použijeme metodu ANOVA, kterou vždy porovnáme, jestli má trojice dat (před, 2 týdny po, 8 týdnů po) stejné střední hodnoty. Hypotézou je rovnost všech tří hodnot, alternativou je alespoň jedna nerovnost.

Pro provedení samotné ANOVA metody lze v Pythonu použít funkce **f_oneway** z balíku *scipy.stats*, která ovšem vrací jenom hodnotu F statistiky a p-hodnotu. Pro získání celé tabulky analýzy rozptylu je nutné použít funkci **ols** z *statsmodels.formula.api* pro tvorbu modelu a následně na ten model použít funkci **stats.anova_lm** z *statsmodels.api*.

Výsledkem metody je tabulka analýzy rozptylu, která obsahuje počet stupňů volnosti, hodnotu F statistiky, součet čtverců(SS) a p-hodnoty.

zdroj	SS	df	F	p-hodnota
mezi skupinami	0.0554	2	1.603	0.205
v rámci skupin	2.076	120	-	-

Obrázek 7: ANOVA CBT1

zdroj	SS	df	F	p-hodnota
mezi skupinami	0.0027	2	0.1378	0.8713
v rámci skupin	1.184	120	-	-

Obrázek 8: ANOVA CBT2

Z tabulek analýzy rozptylu vidíme, že p-hodnoty jsou vyšší než 0.05 a tedy ani v jednom případě nezamítáme hypotézu o shodných středních hodnotách.

Protože se nám nepodařilo prokázat normální rozdělení veličin CBT3 a IOP, je nutné použít neparametrickou verzi ANOVA metody, a to Kruskalův-Wallisův test. Čistě pro ilustraci můžeme ale vyzkoušet i ANOVA metodu.

zdroj	SS	df	F	p-hodnota
mezi skupinami	0.0004	2	0.0446	0.9563
v rámci skupin	0.6654	120	-	-

Obrázek 9: ANOVA CBT3

zdroj	SS	df	F	p-hodnota
mezi skupinami	290	2	9.4289	0.0001
v rámci skupin	1847	120	-	-

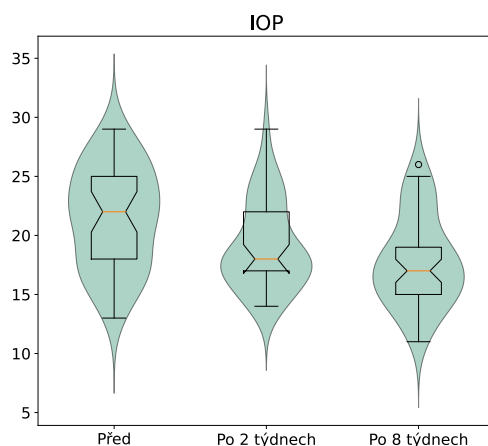
Obrázek 10: ANOVA IOP

Vidíme že ANOVA zamítá rovnost středních hodnot pro nitrooční tlak. Nyní pro tyto dvě veličiny použijeme neparametrický Kruskal-Wallisův test. Tento test je opět implementován jako **kruskal** v knihovně *scipy.stats*. Hypotéza a alternativa zůstávají stejné jako v předešlých případech. Výsledky lze vidět v tabulce níže.

Tabulka 8: Výsledky Kruskal-Wallisova testu

veličiny	testovací statistika	p-hodnota
CBT3	0.2771	0.8706
IOP	14.9622	0.0005

P-hodnota pro veličinu CBT3 je vyšší než 0.05, tudíž nezamítáme hypotézu shodnosti středních hodnot. Avšak p-hodnota pro veličinu IOP je pouze 0.0005 a tedy zamítáme hypotézu, že střední hodnoty nitroočního tlaku se s podáváním léku nemění. Vidíme, že výsledky tohoto testu se neliší od výsledků ilustrativní ANOVA metody. Můžeme se znovu podívat na krabicové grafy nitroočního tlaku, tentokrát bez odlehlých pozorování.



Obrázek 11: Krabicový graf nitroočního tlaku po vyřazení odlehlých dat

Zdá se, že střední hodnota před podáváním léku je vyšší než po začátku podávání.

Pro podrobnější analýzu rozdílů mezi středními hodnotami lze použít pro normální data Tukeyeho vícenásobné porovnávání. Ačkoliv naše data nejsou normální, pro ilustraci použijeme i tento test. Opět lze použít funkci `tukey_hsd` z balíku `scipy.stats`.

Tabulka 9: Tukeyeho vícenásobné porovnávání

porovnání	testovací statistika	p-hodnota	dolní CI	horní CI
W0 - W2	2.415	0.017	0.358	4.471
W0 - W8	3.707	0.000	1.651	5.764
W2 - W8	1.293	0.299	-0.764	3.349

Z Tukeyeho tabulky vyplývá, že dvojice (W0, W8) i dvojice (W0, W2) mají jiné střední hodnoty. U dvojice (W2, W8) se naopak nepodařilo prokázat rozdílnost středních hodnot.

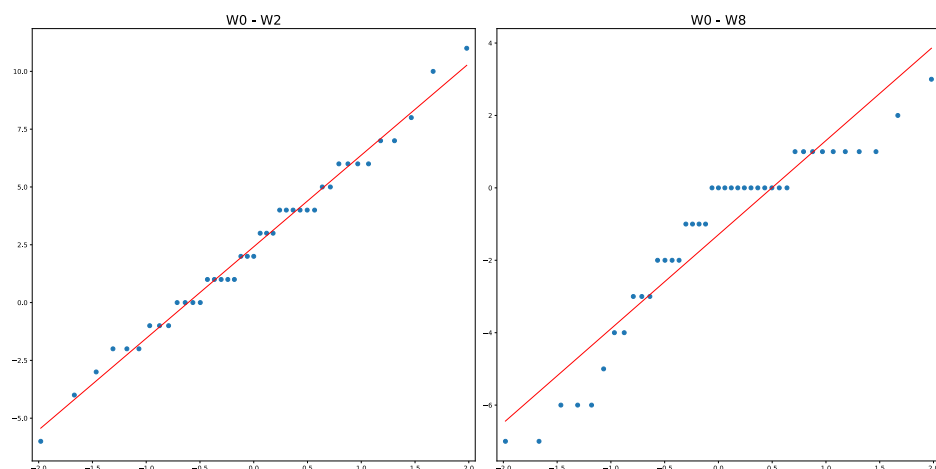
Vzhledem k tomu, že se neprokázala normalita IOP dat, je vhodnější použít neparametrický Dunnův test.

Tabulka 10: Výsledky Dunnova testu

porovnání	p-hodnota
W0 - W2	0.0147
W0 - W8	0.0001
W2 - W8	0.1674

Podle výsledných p-hodnot vidíme, že se Dunnův test shoduje s Tukeyeho a tedy zamítáme hypotézu o rovnosti středních hodnot (W0, W2) a (W0, W8).

Jak již bylo zmíněno, nakonec je nutné ještě výsledky ověřit pomocí t-testu. Klasický t-test lze použít pokud máme shodné rozptyly a nezávislá data. Vzhledem k závislosti daných veličin je nutné použít párový t-test na dvojice (W0IOP, W2IOP) a (W0IOP, W8IOP). Aby byly splněny předpoklady párového t-testu, je nutné ověřit normální rozdělení rozdílů testovaných veličin. K tomu lze opět použít QQ ploty a Lillieforsův test.



Obrázek 12: QQ plot pro rozdíly W0-W2 a W0-W8

Tabulka 11: Výsledky Lillieforsova testu

veličiny	KS statistika	p-hodnota
W0 - W2	0.0871	0.5980
W0 - W8	0.2296	0.0009

Vidíme, že Lillieforsův test zamítá na hladině 0.05 normalitu rozdílů (W0IOP, W8IOP). Pro tuto dvojici bude tedy nutné použít neparametrickou verzi t-testu, v tomto případě Wilcoxonův test.

Pro ilustraci opět použijeme párový test i na dvojici (W0IOP, W8IOP). Výsledky vidíme v tabulce níže. Hypotézou je opět shodnost středních hodnot a alternativou jejich odlišnost. Pro tento test byla použita funkce **ttest_rel**, opět z knihovny *scipy.stats*.

Tabulka 12: Výsledky párového t-testu

dvojice	t-statistika	p-hodnota
(W0IOP, W2IOP)	4.1546	0.0.0001
(W0IOP, W8IOP)	5.8685	7.2334e-7

Pro (W0IOP, W8IOP) použijeme i Wilcoxonův test.

Tabulka 13: Výsledky Wilcoxonova testu

dvojice	statistika	p-hodnota
(W0IOP, W8IOP)	65	9.0315e-6

Vidíme tedy, že ani párový t-test ani Wilcoxonův test na hladině 0.05 zamítají shodnost středních hodnot, a tedy potvrzují výsledek Tukeyeho porovnávání i Dunnova testu.

5 Závěr

Cílem bylo ověřit vliv léku na hodnoty CBT1, CBT2, CBT3 a nitrooční tlak. Již na začátku byl pozorován vyšší nitrooční tlak než je běžný v populaci. Nejprve jsme na základě interpretace veličin smazali odlehlá pozorování s hodnotami CBT veličin vyššími než 1. Následně byly ověřeny předpoklady pro použití ANOVA metody, přičemž se nepovedlo prokázat normalitu veličin CBT3 a IOP.

Použili jsme tedy ANOVA metodu pro CBT1 a CBT2 a neparametrický Kruskalův-Wallisův test pro CBT3 a IOP. U veličin CBT1, CBT2 a CBT3 se nepodařilo prokázat odlišnost středních hodnot, a tedy zamítáme vliv používání léku na tyto veličiny. U veličiny IOP se nám podařilo prokázat odlišnost středních hodnot a pomocí Tukeyeho porovnávání a Dunnova testu jsme zjistili, že střední hodnota nitroočního tlaku před podáváním léku je odlišná od středních hodnot při podávání.

Vzhledem k tomu, že hodnoty nitroočního tlaku u zdravého člověka se pohybují mezi 10 a 20, rozdíl středních hodnot by mohl znamenat, že lék byl úspěšný a snížil nitrooční tlak nemocných pacientů. Z krabicového grafu vidíme, že hodnoty po 8 týdnech jsou opět nižší než hodnoty po 2 týdnech užívání, ačkoliv se v tomto případě neprokázal rozdíl středních hodnot.

Závěrem tedy je, že lék pomohl ke snížení nitroočního tlaku nemocných pacientů a zároveň neměl žádné vedlejší účinky na spektrální citlivost jejich očí.