NON-PARAMETRIC TESTS

Sorana D. Bolboacă

Outline

Non-parametric tests by example:

- Mann-Whitney test
- Wilcoxon test
- Kruskal-Wallis test

Recall on variable types and scales of measurements …

- Quantitative variable (also known as cardinal data) that could be on:
  - Interval scale: zero point is arbitrary (degree)
  - Ratio scale: zero point is fixed
- Qualitative data on:
  - Ordinal scale: condition after treatment as 1 = much improved, 2 = slightly improved, 3 = stays the same; 4 = slightly worse; 5 = much worse
  - Nominal scale: data can be classified into categories but the categories have no specific order

Parametric vs. non-parametric tests

- Parametric statistical methods: parametric form of the distribution is assumed to be known
- Nonparametric statistical methods:
  - Assumptions about the shape of the distribution are not made
  - Central limit theorem seems inapplicable because of small sample size
  - ←make fewer assumptions about the distribution shape

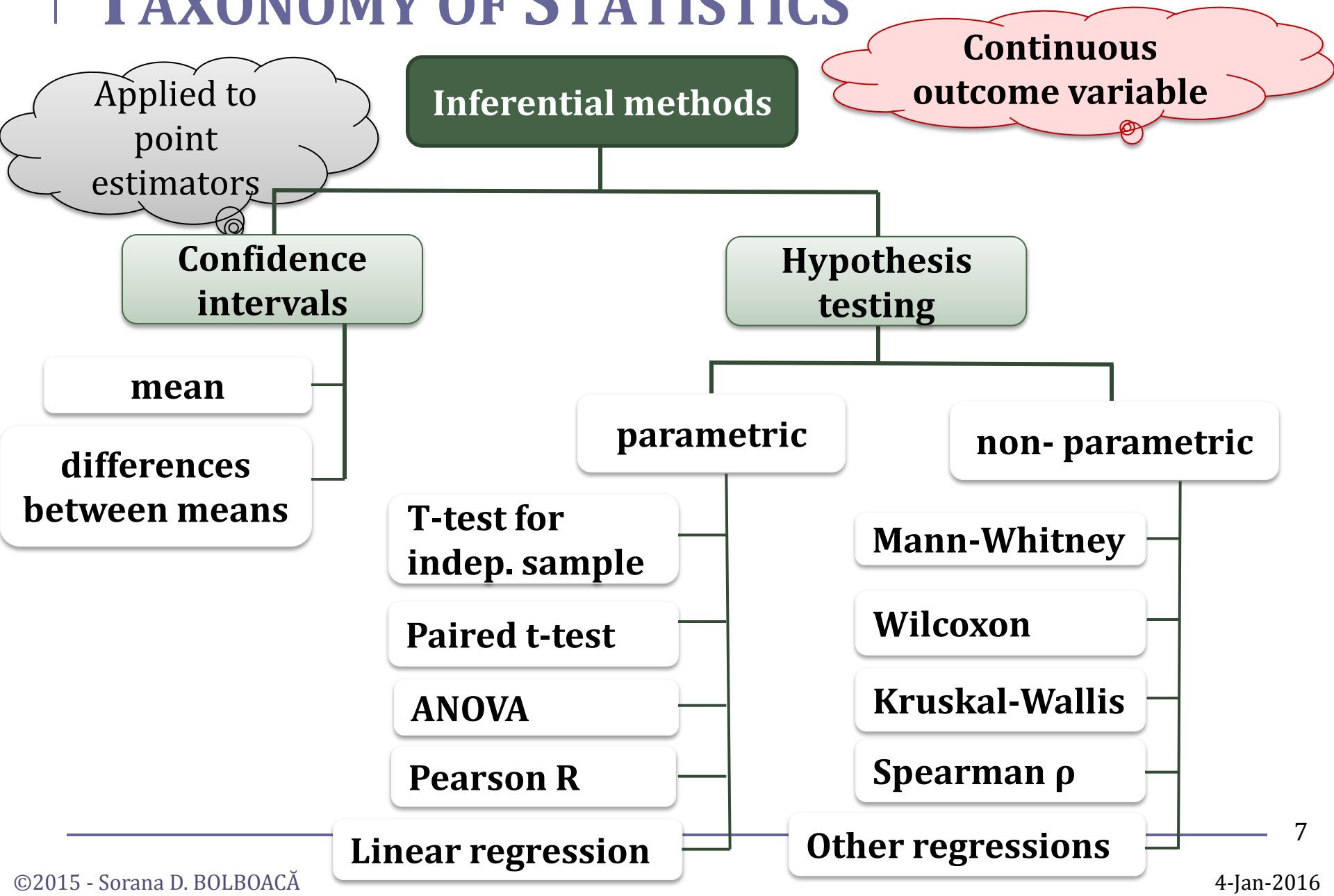
Non-parametric tests: Advantages

- Do Not Involve Population Parameters
  - Example: Probability Distributions, Independence
- Data Measured on any Scale
  - Ratio or Interval
  - Ordinal
    - Example: Good-Better-Best
  - Nominal
    - Example: Male-Female

Non-parametric tests: Disadvantages

- May waste information
  - If data permit using parametric procedures
  - Example: converting data from ratio to ordinal scale
- Difficult to compute by hand for large samples
- Tables not widely available

TAXONOMY OF STATISTICS



Non-parametric tests by examples

Variable scale	Two-samples		K-samples	
	Related samples	Independ samples	Related samples	Independ samples
Nominal	McNewman	Fisher exact Chi-Square	Cochran Q (dichotom)	Chi-Square
Ordinal	Sign Wilcoxon	Mann-Whitney	Friedman Kendall's	Kruskal-Wallis one-way analysis of variance

Mann-Whithey test

- Also known as Wilcoxon rank-sum test
- Tests two independent population
- Corresponds to t-test for 2 independent means
- Assumptions
  - Independent, random samples
  - Populations are continuous
- Can use normal approximation if $n_i \geq 10$

Wilcoxon Signed-Rank Test (Normal Approximation Method for Two-Sided Level α Test)

(1) Rank the differences as shown in Equation 9.4.

(2) Compute the rank sum R_1 of the positive differences.

(3) (a) If $R_1 \neq \frac{n(n+1)}{4}$ and there *are no ties* (no groups of differences with the same absolute value), then

$$T = \left[\left|R_1 - \frac{n(n+1)}{4}\right| - \frac{1}{2}\right] \Big/ \sqrt{n(n+1)(2n+1)/24}$$

(b) If $R_1 \neq \frac{n(n+1)}{4}$ and there *are ties*, where t_i refers to the number of differences with the same absolute value in the ith tied group and g is the number of tied groups, then

$$T = \left[\left|R_1 - \frac{n(n+1)}{4}\right| - \frac{1}{2}\right] \Big/ \sqrt{n(n+1)(2n+1)\Big/24 - \sum_{i=1}^{g}(t_i^3 - t_i)/48}$$

(c) If $R_1 = \frac{n(n+1)}{4}$, then $T = 0$.

(4) If

$$T > z_{1-\alpha/2}$$

then reject H_0. Otherwise, accept H_0.

(5) The p-value for the test is given by

$$p = 2 \times [1 - \Phi(T)]$$

(6) This test should be used only if the number of nonzero differences is ≥ 16 and if the difference scores have an underlying continuous symmetric distribution. The computation of the p-value is illustrated in Figure 9.5.

Ranking Procedure for the Wilcoxon Signed-Rank Test

(1) Arrange the differences d_i in order of *absolute value* as in Table 9.1.

(2) Count the number of differences with the same absolute value.

(3) Ignore the observations where $d_i = 0$, and rank the remaining observations from 1 for the observation with the lowest absolute value, up to n for the observation with the highest absolute value.

(4) If there is a group of several observations with the same absolute value, then find the lowest rank in the range = 1 + R and the highest rank in the range = $G + R$, where R = the highest rank used prior to considering this group and G = the number of differences in the *range of ranks* for the group. Assign the *average rank* = (lowest rank in the range + highest rank in the range)/2 as the rank for each difference in the group.

Mann-Whitney test

- 10 subjects followed Atkin's diet vs. 10 subjects followed DASH diet
  - Atkin's group lost 15.65 kg
  - DASH group lost 8.39 kg

  Conclusion: Is Atkin's better?

  Never conclude without looking to the raw data!!!

Mann-Whithey test

- Comparing the mean weight loss of the two groups is not appropriate! Why?
  - Data are not normally distributed
  - There is an extreme value which significantly influence the mean
- Rank the values (put all values in the same pool and give 20 to the least weight lost and 1 the most weigh lost)
- Sum the ranks for each diet
- The better the higher the sum of ranks

Atkin (rank)	DASH (rank)
+1.81 (20)	-3.63 (14)
+1.36 (19)	-4.54 (13)
0 (18)	-5.44 (11)
-1.36 (17)	-7.26 (8)
-1.81 (16)	-8.16 (7)
-2.27 (15)	-9.07 (6)
-4.99 (12)	-9.53 (5)
-6.35 (10)	-10.88 (4)
-6.80 (9)	-11.79 (3)
-136.08 (1)	-13.61 (2)
∑Rank=137	∑Rank=73

Mann-Whitney U Test (Spreadsheet1)
By variable Group
Marked tests are significant at p <.05000

variable	Rank Sum Atkin	Rank Sum DASH	U	Z	p-level
weight-lost	137.0000	73.00000	18.00000	2.418973	0.015565

DASH clearly ranked lower!

WILCOXON SIGNED RANK

- Tests 2 related populations
- Corresponds to t-test for dependent (paired) means
- Assumptions
  - Random samples
  - Both populations are continuous

WILCOXON SIGNED RANK: PROCEDURE

1. Obtain difference scores, $d_i = x_{1i} - x_{2i}$
   - Note: in the text, D1 is what's called X1 here
2. Take absolute value of differences, d_i
3. Delete differences with 0 value
4. Assign ranks, r_i, where smallest = 1
5. Assign ranks same signs as d_i
6. Sum '+' ranks (T_+) & '-' ranks (T_-)
   - Test statistic is T_- (one-tailed test)
   - Test statistic is smaller of T_- or T_+ (2-tail)

WILCOXON SIGNED RANK: EXAMPLE

You work at the financial summary of your office. Is the **new** financial package **faster** (**0.05** level)? You collect the following data entry times:

User	Old software	New software	D_i	$\|D_i\|$	Rank	Sign
1	9.98	9.88	+0.10	0.10	4	+
2	9.88	9.86	+0.02	0.02	1	+
3	9.90	9.83	+0.07	0.07	2/2.5	+
4	9.99	9.80	+0.19	0.19	5	+
5	9.94	9.87	+0.07	0.07	3/2.5	+
6	9.84	9.84	0.00	0.00		

T+ = 15, T- = 0

Wilcoxon Matched Pairs Test (Spreadsheet4)
Marked tests are significant at p <.05000

Pair of Variables	Valid N	T	Z	p-level
Old & New	6	0.00	2.022600	0.043115

Kruskal-Wallis Test

- Non-parametric method used to compare k independent samples
- Ocular anti-inflammatory effects of four drugs on lip closure after administration of arachidonic acid

	1 Var1	2 Var2
1	Indomethacin	2
2	Indomethacin	3
3	Indomethacin	3
4	Indomethacin	3
5	Indomethacin	3
6	Indomethacin	0
7	Aspirin	1
8	Aspirin	3
9	Aspirin	1
10	Aspirin	2
11	Aspirin	2
12	Aspirin	3
13	Piroxicam	3
14	Piroxicam	1
15	Piroxicam	2
16	Piroxicam	1
17	Piroxicam	3
18	Piroxicam	3
19	BW755C	1
20	BW755C	0
21	BW755C	0
22	BW755C	0
23	BW755C	0
24	BW755C	2

Kruskal-Wallis ANOVA by Ranks; Var2 (Spreadsheet6)
Independent (grouping) variable: Var1
Kruskal-Wallis test: H (3, N= 24) =8.860063 p =.0312

Depend.: Var2	Code	Valid N	Sum of Ranks
Indomethacin	101	6	96.00000
Aspirin	102	6	82.00000
Piroxicam	103	6	89.00000
BW755C	104	6	33.00000

Non-parametric tests by example

PLoS One. 2015 Dec 23;10(12):e0145303. doi: 10.1371/journal.pone.0145303.

Effects of Pre-Natal Vitamin D Supplementation with Partial Correction of Vitamin D Deficiency on Early Life Healthcare Utilisation: A Randomised Controlled Trial.

Griffiths M[1], Goldring S[1], Griffiths C[2], Shaheen SO[2], Martineau A[2], Cross L[2], Robinson S[3], Warner JO[1], Devine A[2], Boyle RJ[1].

Author information

Abstract

BACKGROUND: Some observational studies have suggested that higher prenatal Vitamin D intake may be associated with improved health outcomes in childhood. However there have been mixed results in this area with some negative studies, especially for effects on atopic and respiratory outcomes. We examined the effect of prenatal Vitamin D on healthcare utilisation in the first three years of life.

METHODS: In an ethnically stratified randomised controlled trial conducted at St Mary's Hospital London, 180 women at 27 weeks gestation were allocated to no Vitamin D, 800 IU ergocalciferol daily until delivery, or a single oral bolus of 200,000 IU cholecalciferol. Participants were randomised in blocks of 15 using computer-generated numbers and investigators were blinded to group assignment. Supplementation increased maternal and cord blood 25(OH) vitamin D concentrations, but levels remained lower than current recommendations. Primary health economic outcome was overall cost of unscheduled healthcare utilisation in the first three years of life as documented in the child's electronic health record. Secondary outcomes included cost attributable to: primary and secondary healthcare visits, respiratory and atopic complaints, cost in years 1, 2 and 3 of life and cost and frequency of prescribed medication. All costs were calculated as pounds sterling. Differences between groups were analysed using unpaired t-test or Mann-Whitney U test, and analysis of variance for adjusted analyses.

RESULTS: We assessed 99/180 (55%) complete electronic health records, control (n = 31), daily (n = 36) and bolus (n = 32). We found no difference in total healthcare utilisation costs between the control and daily (mean difference in costs in pounds sterling 1.02, 95%CI -1.60, 1.65; adjusted 1.07, 95%CI -1.62, 1.86) or control and bolus groups (mean difference -1.58, 95%CI -2.63, 1.06; adjusted -1.40, 95%CI -2.45, 1.24). There were no adverse effects of supplementation reported during the trial.

CONCLUSIONS: We found no evidence that prenatal vitamin D supplementation from 27 weeks gestation to delivery, at doses which failed to completely correct maternal vitamin D deficiency, influence overall healthcare utilisation in children in the first 3 years.

TRIAL REGISTRATION: Controlled-Trials.com ISRCTN68645785.

Non-parametric tests by example

PLoS One. 2015 Dec 23;10(12):e0145303. doi: 10.1371/journal.pone.0145303.

Effects of Pre-Natal Vitamin D Supplementation with Partial Correction of Vitamin D Deficiency on Early Life Healthcare Utilisation: A Randomised Controlled Trial.

Griffiths M[1], Goldring S[1], Griffiths C[2], Shaheen SO[2], Martineau A[2], Cross L[2], Robinson S[3], Warner JO[1], Devine A[2], Boyle RJ[1].

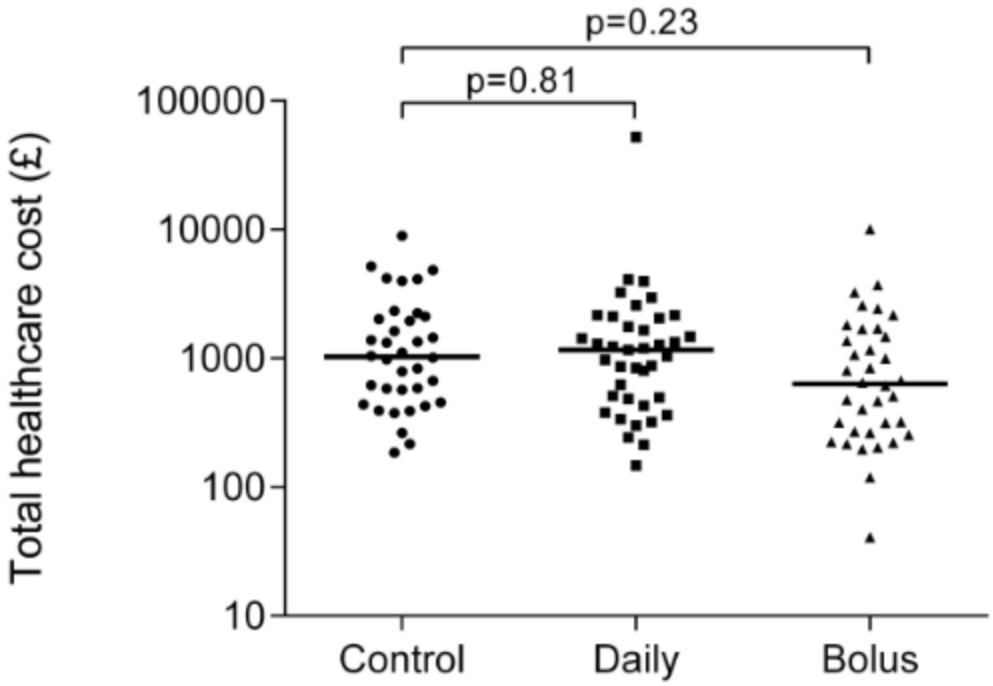


Fig 2. Effect of prenatal vitamin D randomisation on total healthcare utilisation in the first three years of life. Horizontal bars represent medians and p values are for adjusted analyses.

doi:10.1371/journal.pone.0145303.g002

Non-parametric tests by example

J Tehran Heart Cent. 2015 Jul 3;10(3):122-8.

Evaluating the Potential Effect of Melatonin on the post-Cardiac Surgery Sleep Disorder.

Dianatkhah M[1], Ghaeli P[1], Hajhossein Talasaz A[2], Karimi A[3], Salehiomran A[3], Bina P[3], Jalali A[3], Ghaffary S[1], Shahmansouri N[3], Vejdani S[1].

Author information

Abstract

BACKGROUND: Postoperative neurological injuries, including cognitive dysfunction, sleep disorder, delirium, and anxiety, are the important consequences of coronary artery bypass graft surgery (CABG). Evidence has shown that postoperative sleep disturbance is partly due to disturbed melatonin secretion in the perioperative period. The aim of this study was to evaluate the effect of melatonin on postoperative sleep disorder in patients undergoing CABG.

METHOD: One hundred forty-five elective CABG patients participated in a randomized double-blind study during the preoperative period. The patients were randomized to receive either 3 mg of melatonin or 10 mg of Oxazepam one hour before sleep time. Each group received the medication from 3 days before surgery until the time of discharge. Sleep quality was evaluated using the Groningen Sleep Quality Score (GSQS), and the incidence of delirium was evaluated by nursing records. Sleep quality and anxiety scores were compared before and after surgery through the Wilcoxon signed-rank test. The analysis of covariance (ANCOVA) and independent t-test were used to compare the sleep and anxiety scores between the groups. P values ≤ 0.05 were considered statistically significant.

RESULTS: Totally, 137 patients at a mean age of 60 years completed the study (76% male). The analysis of the data showed that sleep was significantly disturbed after surgery in both groups. The patients in the Oxazepam group demonstrated significantly higher disturbance in their mean postoperative GSQS score than did their counterparts in the melatonin group (p value < 0.001). A smaller proportion of the participants experienced delirium in the melatonin group (0.06%) than in the Oxazepam group (0.12%); however, this difference was not statistically significant.

CONCLUSION: The result of the present study revealed that melatonin improved sleep in post-cardiac surgery patients more than what was observed with Oxazepam. Therefore, melatonin may be considered an effective alternative for Benzodiazepines in the management of postoperative sleep disorder.

KEYWORDS: Coronary artery bypass; Melatonin; Sleep disorders

Non-parametric tests by example

Table 2. Effect of melatonin on postoperative sleep, anxiety, and delirium

	Melatonin	Oxazeoam	P Value
Delirium			
After intervention	4 (6.1)	9 (12.7)	0.187*
Sleep			
Before intervention	2.0 (1.0-3.0)	2.0 (1.5-3.5)	0.436**
After intervention	2.5 (1.5-5.6)	8.0 (3.0-11.0)	< 0.001**
P Value	0.001***	< 0.001***	
Anxiety			
Before intervention	3.5 (2.0-6.2)	6.0 (2.0-10.0)	0.055**
After intervention	5.0 (3.7-8.2)	9.0 (5.0-13.0)	0.171****
P Value	0.013***	< 0.001***	

Variables are presented as n (%), mean with±SD, or median with interquartile range (IQR) boundaries.
*Chi-square test. **Independent T-test. ***Wilcoxon signed-ranked test. ****Analysis of covariance.

Non-parametric tests by example

J Res Pharm Pract. 2015 Oct-Dec;4(4):187-92. doi: 10.4103/2279-042X.167050.

Comparison of the effects of intravenous premedication: Midazolam, Ketamine, and combination of both on reducing anxiety in pediatric patients before general anesthesia.

Sajedi P[1], Habibi B[1].

⊕ Author information

Abstract

OBJECTIVE: In some medical circumstances, pediatric patients may need premedication for transferring to the operating room. In these situations, using intravenous premedication is preferred. We assessed the efficacy and safety of intravenous midazolam, intravenous ketamine, and combination of both to reduce the anxiety and improve behavior in children undergoing general anesthesia.

METHODS: In a double-blind randomized clinical trial, 90 pediatric patients aged 6 months to 6 years with American Society of Anesthesiologist grade I or II were enrolled. Before anesthesia, children were randomly divided into three groups to receive intravenous midazolam 0.1 mg/kg, or intravenous ketamine 1 mg/kg, or combination of half doses of both. Behavior types and sedation scores were recorded before premedication, after premedication, before anesthesia, and after anesthesia in the postanesthesia care unit. Anesthesia time, recovery duration, blood pressure, and heart rate were also recorded. For comparing distribution of behavior types and sedation scores among three groups, we used Kruskal-Wallis test, and for comparing mean and standard deviation of blood pressure and heart rates, we used analysis of variance.

FINDINGS: After premedication, children's behavior was significantly better in the combination group ($P < 0.001$). After anesthesia, behavior type was same among three groups ($P = 0.421$). Sedation scores among three groups were also different after premedication and the combination group was significantly more sedated than the other two groups ($P < 0.001$).

CONCLUSION: Combination of 0.05 mg/kg of intravenous midazolam and 0.5 mg/kg of intravenous ketamine as premedication produced more deep sedation and more desirable behavior in children compared with each midazolam 0.1 mg/kg or ketamine 1 mg/kg.

KEYWORDS: Ketamine; Midazolam; pediatric; premedication

PMC full text: J Res Pharm Pract. 2015 Oct-Dec; 4(4): 187–192.
doi: 10.4103/2279-042X.167050

Table 4

Frequency distribution of behavioral types in three groups at different times of the study

Time period	Behavior type level	Groups			*P**	*P***
		Midazolam	Ketamine	Midazolam + Ketamine		
Before premedication	1	0 (0)	0 (0)	2 (7)	0.133	0.684
	2	27 (90)	25 (83)	23 (76)		
	3	3 (10)	3 (10)	5 (17)		
	4	0 (0)	2 (7)	0 (0)		
After premedication	1	2 (7)	1 (3)	25 (83)	<0.001	<0.001
	2	27 (90)	1 (3)	5 (17)		
	3	1 (3)	15 (50)	0 (0)		
	4	0 (0)	13 (43)	0 (0)		
Before general anesthesia	1	7 (23)	2 (7)	30 (10)	<0.001	<0.001
	2	23 (76)	1 (3)	0 (0)		
	3	23 (76)	1 (3)	0 (0)		
	4	0 (0)	11 (37)	0 (0)		
After general anesthesia	1	29 (97)	28 (93)	30 (10)	0.421	0.330
	2	1 (3)	1 (3)	0 (0)		
	3	0 (0)	1 (3)	0 (0)		
	4	0 (0)	0 (0)	0 (0)		

Data are presented as number (percentage) of patients. **P* value for desirable behavior; ***P* value for undesirable behavior

Recall!!!

- Descriptive statistic parameters must be calculated according with type of variables and units of measurements
- Inferential statistics is choose based on type of variable and after verification of assumptions for each approach!
- A parametric test has a correspondent non-parametric test