İSTANBUL UNIVERSITY PRESS

ARAŞTIRMA MAKALESİ / RESEARCH ARTICLE
Journal of Advanced Research in Health Sciences /
Sağlık Bilimlerinde İleri Araştırmalar Dergisi 2023
DOI: 10.26650/JARHS2023-1224130

COMPARISON OF FREQUENTIST AND BAYESIAN APPROACHES ON SAMPLE SIZE: METHODOLOGIC STUDY

İSTATİSTİKTE FREKANSÇI VE BAYESYEN YAKLAŞIMIN ÖRNEKLEM BÜYÜKLÜĞÜ ÜZERİNDEKİ ETKİLERİNİN KARŞILAŞTIRILMASI: METODOLOJİK ÇALIŞMA

Cennet YILDIZ[1], Eray YURTSEVEN[2], Fatma Nihan TURHAN ÇAĞLAR[1], Mehmet Güven GÜNVER[2]

[1]University of Health Sciences, Bakırköy Dr. Sadi Konuk Education and Research Hospital, Department of Cardiology, Istanbul, Turkiye
[2]Istanbul University, Istanbul Faculty of Medicine, Department of Biostatistics, Istanbul, Turkiye

ORCID ID: C.Y. 0000-0003-2456-3206; E.Y. 0000-0003-0565-6407; F.N.T.Ç. 0000-0001-7925-2398; M.G.G. 0000-0002-4628-8391

Citation/Atıf: Yildiz C, Yurtseven E, Turhan Caglar FN, Gunver MG. Comparison of frequentist and Bayesian approaches on sample size: methodologic study. Journal of Advanced Research in Health Sciences 2023;6(2):122-128. https://doi.org/10.26650/JARHS2023-1224130

ABSTRACT

Objective: In the present study, we aimed to evaluate the effects of sample size on results of study by using frequentist and Bayesian approaches.
Material and Methods: The small sample consisted of 32 patients with ischemic heart disease (IHD) and 37 control subjects. In order to compare the statistical differences between small and large sample sizes, two samples were constituted. All the patients included in the study were male and between 40-50 years old. The large sample consisted of 355 IHD patients and 545 controls. Patients' biochemical variables including glucose, triglyceride (TG), total cholesterol (TC), high density lipoprotein cholesterol (HDL), low density lipoprotein cholesterol (LDL), urea, creatinine, hemoglobin, hematocrit, (HCT), red cell distribution width (RDW), White blood cell (WBC), platelet (PLT), mean platelet volume (MPV), neutrophil (NEUT), lymphocyte (LYM) were recorded. Patients in the small and large samples were compared with both frequentist and Bayesian approaches.
Results: Except for glucose levels there were no statistical differences with respect to the biochemical variables of two groups in a small sample size when the variables were analyzed by the frequentist approach. Similarly, we did not find any differences between biochemical variables when the data were analyzed by the Bayesian approach. When the large sample size data were analyzed by the frequentist approach, glucose, TG, TC, HDL, LDL, urea, creatinine, hemoglobin, HCT, WBC, NEUT, LYMP levels were found to be statistically significantly different between patients who had IHD and the controls. Similarly, there were significant differences between two groups with respect to glucose, TG, TC, HDL, LDL, urea, creatinine, hemoglobin, HCT, WBC, NEUT, LYMP levels when the data analyzed by Bayesian approach.
Conclusion: Our study results suggested that there were no differences between the frequentist and Bayesian approach results when the sample size is large and the power of the study is high.
Key words: Frequentist, bayesian, sample size

ÖZ

Amaç: Bu çalışmada, frekantist ve Bayesyen yaklaşımlar kullanılarak örneklem büyüklüğünün araştırma sonuçları üzerindeki etkilerinin araştırılması amaçlanmıştır.
Gereç ve Yöntem: Çalışmamızda küçük ve büyük örneklem büyüklüğünde istatistiksel farklılıkları karşılaştırmak amacı ile küçük ve büyük olmak üzere iki örneklem oluşturulmuştur. Çalışmaya alınan tüm hastalar erkek ve 40 ile 50 yaş aralığındadır. Küçük örneklem için iskemik kalp hastalığı (İKH) olan 32, İKH olmayan 37 kişi çalışmaya dahil edilmiştir. Büyük örneklem için İKH olan 355, olmayan 545 kişi çalışmaya alınmıştır. Tüm hastaların glukoz, trigliserid (TG), total kolesterol (TKOL), yüksek yoğunluklu lipoprotein kolesterol (HDL), düşük yoğunluklu lipoprotein kolesterol (LDL), üre, kreatinin, hemoglobin, hematokrit (HCT), kırmızı kan hücresi dağılım genişliği (RDW), lökosit (WBC), trombosit (PLT), ortalama trombosit hacmi (MPV), nötrofil (NÖT), lenfosit (LYM) değerleri kaydedilmiştir. Küçük ve büyük örneklemler frekansçı ve Bayesyen yaklaşımla karşılaştırılımıştır.
Bulgular: Küçük örneklem büyüklüğünde frekantist yaklaşım ile yapılan analizde tüm biyokimyasal veriler İKH olan ve olmayan kişilerde karşılaştırılmış ve glukoz seviyeleri dışında diğer parametrelerde anlamlı fark saptanmamıştır. Yine grupların Bayesyen yaklaşımla yapılan karşılaştırmalarında parametreler arasında anlamlı istatistiksel fark elde edilmemiştir. Buna karşın büyük örneklem büyüklüğünde frekantist yaklaşım ile yapılan karşılaştırmalarda glukoz, TG, TKOL, HDL, LDL, üre, kreatinin, hemoglobin, HCT, WBC, NÖT ve LYM değerleri her iki grup arasında anlamlı olarak farklı çıkmıştır. Aynı şekilde Bayesyen yaklaşım ile yapılan karşılaştırmalarda glukoz, TG, TKOL, HDL, LDL , üre, kreatinin, hemoglobin, HCT, WBC, NÖT ve LYM değerleri iki grup arasında istatistiksel olarak anlamlı çıkmıştır.
Sonuç: Büyük örneklem büyüklüğünde ve yüksek bir güçte çalışmada verinin frekansçı ya da Bayesyen istatistik ile değerlenirilmesi açısından fark bulunmamamaktadır.
Anahtar Kelimeler: Frekantist, bayesyen, örneklem büyüklüğü

Corresponding Author/Sorumlu Yazar: Cennet YILDIZ **E-mail:** cennet_yildiz@live.com

Submitted/Başvuru: 25.12.2022 • **Revision Requested/Revizyon Talebi:** 02.03.2023 • **Last Revision Received/Son Revizyon:** 09.03.2023 • **Accepted/Kabul:** 09.03.2023

INTRODUCTION

The main purpose of statistical interference is to make predictions about population based on the data that are derived from a sample. Population is defined as the entire group that we want to draw conclusions about (1).

The field of statistics, as a branch of science, has been influenced by different ideas during its development. These ideas have become evident over time and are now polarized as the frequentist and Bayesian approaches. Deductive and inductive methods have been adopted in frequentist and Bayesian approaches, respectively. In the frequentist approach, data are random, while parameters are unknown and fixed. In contrast, according to the Bayesian approach, data are fixed, and do not change after observations, whereas parameters are accepted as a random variable (1). Statistical hypothesis testing also differs between the two approaches. While deterministic rules are followed by the frequentist approach, the Bayesian methods are closer to a probability based interpretation. Bayesian statistics involves updating prior beliefs as more evidence becomes available (2). However, the frequentist statistics are interested in with whether an event (hypothesis) occurs or not. In this method, results of repeated experiments are examined under the same conditions and analysis of external information other than the sample data are not made.

Sample size estimation is one of the important steps in scientific studies. In statistics, the universe is the of set all experimental units, from which a sample is to be drawn. In order for the results of the study to be reliable, sample size should be sufficient in number and represent the universe appropriately. It is among one of the factors that directly affects the strength of a study. As the number of observations related to the research increases, the reliability of the data also increase (3). In large sample sizes, meaningless effects can be found to be statistically significant, whereas in small sample sizes these differences may not be detected. For these reasons, it is recommended to keep the sample size at the optimum level (4,5). Increasing the sample size reduces the standard error, resulting in more concentrated distributions around the mean (6).

Cardiovascular diseases account for a third of deaths worldwide (7). Among these diseases, the prevalence of ischemic heart disease (IHD) is the highest (9). Many studies have been conducted to compare the biochemical findings of IHD patients with normal subjects. As a result of these studies, various results have emerged. Leukocytes (WBC), which play a role in atherosclerosis pathogenesis, have been found as a prognostic factor for coronary artery disease (CAD) (9). It has been suggested that an increased number of leukocytes increases the risk of death due to IHD by 65% (10). Hyperlipidemia is a strong and modifiable risk factor for cardiovascular diseases. Patients with IHD have been shown to have higher levels of total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C) and lower levels of high-density lipoprotein cholesterol (HDL-C) compared to healthy subjects (11). Lower levels of platelet (PLT) number and mean platelet volume (MPV) have been found in patients with acute coronary syndromes (12). However, the number of PLT did not differ between patients with chronic coronary syndromes and healthy subjects (13). Both high and low levels of hemoglobin (Hgb) concentrations have been associated with cardiovascular diseases and low levels of Hgb concentrations have prognostic value in patients with IHD (14).

In the present study, we aimed to compare the effects of frequentist and Bayesian approaches on sample size and to find whether any differences exist between two methods.

MATERIAL and METHOD

The present study compared the biochemical variables of IHD patients with healthy subjects. For comparison two sample sizes were constituted. The small sample size was composed of 32 consecutive patients with IHD and 37 consecutive healthy subjects who applied to our cardiology outpatient clinic between December 2021 and January 2022. The large sample size was composed of 355 consecutive patients with IHD and 545 consecutive healthy subjects who applied to our cardiology outpatient clinic between January 2021 and September 2022. Patients with chronic renal failure, hepatic diseases, thyroid function abnormalities or malignancy were excluded from the study. Ethical approval of the study was obtained from Bakırköy Dr. Sadi Konuk Training and Education Hospital Ethical Committee and the study was conducted in accordance with the Helsinki declaration (approval date: 04/10/2021, approval number: 2021-19). All patients gave written informed consent before study enrollment. Blood samples of the participants were drawn from the antecubital vein after 12-hour fasting. Patients' TC, HDL-C, LDL-C, triglyceride (TG), glucose, urea, creatinine and hemogram values were determined. Small and large sample sizes were compared both with frequentist and Bayesian approaches.

H_0 (null) and H_1 (alternative) hypotheses about the θ parameter were established and priori and posterior probabilities related to this parameter were calculated. The final decision was made by dividing the posterior distribution of the alternative hypothesis to posterior distribution of the null hypothesis.

if $\frac{P\,(H1\backslash Data)}{P\,(H0\backslash Data)} > 1$, then the H_1 hypothesis was selected,

if $\frac{P\,(H1\backslash Data)}{P\,(H0\backslash Data)} < 1$, then the H_0 hypothesis was selected.

Posterior distribution of the null hypothesis was calculated as follows;

$$P\,(H_0\backslash Data) = \frac{P\,(Data\backslash Ho)P(Ho)}{P\,(Data\backslash Ho)P(Ho)+P(Data\backslash H1)P\,(H1)}$$

Posterior distribution of the alternative hypothesis was calculated as follows;

$$P\ (H_1\backslash Data) = \frac{P\ (Data\backslash H1)P(H1)}{P\ (Data\backslash Ho)P(Ho)+P(Data\backslash H1)P\ (H1)}$$

The result of the division of two posterior distributions was: $\frac{P\ (Data\backslash H1)}{P\ (Data\backslash H0)}$. This ratio was called the Bayes factor. If the Bayes factor was between 1 and 3 or 3 and 10, then there was anecdotal or moderate evidence for the alternative hypothesis, respectively. If the Bayes factor was in between 10 and 30 or 30 and 100, then there was strong and very strong evidence for the alternative hypothesis, respectively. If the Bayes factor was greater than 100, then there was extreme evidence for the alternative hypothesis. Alternatively, if Bayes factor was in between 1/3 and 1/10, 1/10 and 1/30, 1/30 and 1/100, or less than 1/100, then there was anecdotal, moderate, strong and extreme evidence for thr null hypothesis respectively (15).

Statistical analysis

Normality testing of the data was made by the Kolmogorow-Smirnow test. Parametric and non-parametric data were expressed as mean±SD and median and interquartile range (25-75), respectively. A comparison of the two groups was made by the student's t test or Mann-Whitney U test. A p value of less than 0.05 was considered as significant. In order to make Bayesian comparisons, informative prior distributions were used. All of the statistical analyses were done by using IBM SPSS version 26 software (the Statistical Package for the Social Sciences).

RESULTS

Small sample size

Except for serum glucose levels, there were no differences between biochemical variables of the two groups when analyzed by frequentist methods. Serum glucose levels were found to be significantly higher in patients with IHD than that of controls (117.50 (96.50-160.00) mg/dl vs 102.20 (93.40-115.00) mg/dl, p=0.031, respectively). Table 1 shows the comparison of the two groups with frequentist approach.

Informative Bayesian t-test results showed no differences between the biochemical variables of IHD patients and controls. Values of Bayes factors were found to be between zero and one, indicating no significant difference. Since the results were not statistically significant, all of the 95% confidence intervals covered zero. Table 2 and Table 3 show informative Bayesian t test results and posterior distribution statistics, respectively.

Large sample size

According to the results of the frequentist methods, all biochemical variables except for red cell distribution width (RDW), PLT (platelet), and mean thrombocyte volume (MPV) were significantly different between two groups. Table 4 shows comparison results of the two groups according to the frequentist approach.

Informative Bayesian t-test results showed that levels of glucose, TG, TC, HDL-C, LDL-C, urea, creatinine, hematocrit (HCT), WBC, neutrophil (NEUT) and lymphocyte (LYMP) were

Table 1: Comparison of patients with IHD and controls by the frequentist approach (small sample size).

Parameter	Control	IHD	Z/t score	p
Glucose (mg/dL)	102.20 (93.40-115.00)	117.50 (96.50-160.00)	-2.154	**0.031**
Triglyceride (mg/dL)	169.75 (117.96-229.00)	141.00 (112.00- 303.00)	-1.294	0.196
TC (mg/dL)	190.091±33.672	189.648±58.088	0.039	0.969
HDL-C (mg/dL)	43.00 (41.80-47.80)	41.00 (36.32-51.50)	-1.071	0.284
LDL-C (mg/dL)	99.10 (88.30-132.68)	111.00 (62.52-152.60)	-0.102	0.919
Urea (mg/dL)	30.90 (26.05-34.275)	29.00 (26.25-36.00)	-0.614	0.539
Creatinine (mg/dL)	0.87 (0.825-0.940)	0.875 (0.765-0.98)	-0.175	0.861
Hemoglobin (g/dL)	15.45 (14.60-16.10)	14.95 (13.85-15.425)	-1.680	0.093
Hematocrit (%)	45.633±2.696	43.869±4.972	1.792	0.080
RDW (%)	13.125 (12.90-13.50)	13.10 (12.606-13.50)	-0.458	0.647
WBC (103/µL)	7.985±1.701	8.517±2.276	-1.108	0.272
Platelet (10e3/uL)	196.00 (20.895-287.00)	244.00 (166.27-289.25)	-0.963	0.336
MPV (fL)	10.081(9.90-10.30)	10.05 (9.225-10.425)	-0.441	0.659
NEUT (103/µL)	4.885 (4.205-4.885)	5.415 (3.89-6.21)	-1.280	0.200
LYMP (103/µL)	3.180 (1.180-4.02)	2.82 (0.77-4.67)	-1.280	0.200

TC: Total cholesterol, HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC: White blood cell, MPV: Mean platelet volume, NEUT: Neutrophil, LYMP: Lymphocyte.

Table 2: Results of informative Bayesian t-tests (small sample size).

Parameter	MD	PSD	Bayes factor	t score	df	p
Glucose (mg/dL)	22.834	13.184	0.882	1.732	67	0.088
Triglyceride (mg/dL)	-13.666	29.743	0.271	-0.459	67	0.647
TC (mg/dL)	-0.442	11.247	0.248	-0.039	67	0.969
HDL-C (mg/dL)	-0.465	2.475	0.252	-0.188	67	0.852
LDL-C (mg/dL)	3.418	10.186	0.260	0.336	67	0.738
Urea (mg/dL)	4.055	2.958	0.551	1.371	67	0.175
Creatinine (mg/dL)	0.032	0.142	0.254	0.227	67	0.821
Hemoglobin (g/dL)	-0.447	0.405	0.417	-1.104	67	0.273
Hematocrit (%)	-1.764	0.946	1.078	-1.866	67	0.066
RDW (%)	-0.268	0.388	0.304	-0.690	67	0.493
WBC (103/µL)	0.532	0.480	0.418	1.108	67	0.272
Platelet (10e3/uL)	32.793	26.998	0.465	1.215	67	0.229
MPV (fL)	-0.236	0.198	0.454	-1.192	67	0.237
NEUT (103/µL)	0.349	0.343	0.386	1.018	67	0.313
LYMP (103/µL)	-2.256	0.235	0.411	-1.089	67	0.280

MD: Mean difference, PSD: Pooled standard error difference, df: Degree of freedom, TC: Total cholesterol, HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC: White blood cell, MPV: Mean platelet volume, NEUT: Neutrophil, LYMP: Lymphocyte.

Table 3: Posterior distribution statistics (small sample size).

Parameter	Posterior			%95 CI	
	Mode	Median	Variance	Lower limit	Upper limit
Glucose (mg/dL)	23.124	23.124	173.918	-2.723	48.972
Triglyceride (mg/dL)	-12.226	-12.226	844.040	-69.167	44.716
TC (mg/dL)	-1.427	-1.427	129.644	-23.743	20.890
HDL-C (mg/dL)	-0.836	-0.836	6.130	-5.689	4.017
LDL-C (mg/dL)	2.609	2.609	105.487	-17.521	22.739
Urea (mg/dL)	3.992	3.992	9.129	-1.930	9.914
Creatinine (mg/dL)	0.095	0.095	0.016	-0.152	0.341
Hemoglobin (g/dL)	-0.447	-0.447	0.162	-1.235	0.342
Hematocrit (%)	-1.819	-1.819	0.941	-3.721	0.082
RDW (%)	-0.224	-0.224	0.150	-0.983	0.535
WBC (103/µL)	0.541	0.541	0.236	-0.411	1.494
Platelet (10e3/uL)	31.130	31.130	650.090	-18.843	81.102
MPV (fL)	-0.240	-0.240	0.038	-0.623	0.143
NEUT (103/µL)	0.367	0.367	0.121	-0.314	1.047
LYMP (103/µL)	-0.256	-0.256	1.347	-2.530	2.018

CI: Confidence interval, TC: Total cholesterol, HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC: White blood cell, MPV: Mean platelet volume, NEUT: Neutrophil, LYMP: Lymphocyte.

significantly different between the two groups. The Bayes factors were found to be extremely high which supported the alternative hypothesis. Similar to the frequentist approach, Bayesian methods also did not find any differences in RDW, PLT and MPV values between the two groups of subjects. Bayesian factors of these variables were found to be between

Table 4: Comparison of patients with IHD and controls by the frequentist approach (large sample size)

Parameter	Control	IHD	t score	p
Glucose (mg/dL)	100.75 (92.00-117.00)	111.75 (95.00-146.00)	-5.788	**<0.001**
Triglyceride (mg/dL)	155.25 (115.00-219.50)	185.00 (126.00-276.50)	-3.028	**0.003**
TC (mg/dL)	192.00 (172.50-220.125)	178.00 (146.50-207.50)	5.092	**<0.001**
HDL-C (mg/dL)	42.00 (37.00-47.50)	38.00 (34.00-44.00)	5.614	**<0.001**
LDL-C (mg/dL)	118.00 (99.00-141.50)	98.00 (72.81-122.75)	7.811	**<0.001**
Urea (mg/dL)	29.00 (25.00-34.00)	30.00 (25.00-35.00)	-2.385	**0.017**
Creatinine (mg/dL)	0.84 (0.76-0.94)	0.85 (0.76-0.95)	-2.387	**0.017**
Hemoglobin (g/dL)	15.20 (14.40-15.90)	14.80 (13.90-15.80)	2.927	**0.004**
Hematocrit (%)	44.70 (42.25-47.00)	43.95 (41.00-46.47)	2.998	**0.003**
RDW (%)	13.10 (12.60-13.50)	13.00 (12.60-13.70)	-0.389	0.698
WBC (103/µL)	7.92 (6.69-9.40)	8.75 (7.37-10.39)	-5.634	**<0.001**
Platelet (10e3/uL)	248.00 (212.50-290.00)	245.50 (209.25-290.50)	-1.181	0.238
MPV (fL)	10.10 (9.40-10.70)	10.10 (9.50-10.80)	-1.065	0.287
NEUT (103/µL)	4.74 (3.69-5.75)	5.20 (4.17-6.36)	-5.005	**<0.001**
LYMP (103/µL)	2.50 (1.96-3.03)	2.50 (2.09-3.04)	-2.186	**0.029**

TC: Total cholesterol, HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC: White blood cell, MPV: Mean platelet volume, NEUT: Neutrophil, LYMP: Lymphocyte.

zero and one. Table 5 and Table 6 show Bayesian t-test results and posterior distribution statistics of the two groups.

DISCUSSION

In many experimental studies, significance of scientific results should be supported by a p value which belongs to the frequentist paradigm. The frequentist approach consists of a combination of two approaches: the null hypothesis that was put forward by Fisher's inductive approach and Neyman and Pearson's deductive alternative hypothesis, and the concept of power (16, 17). According to the Fisher approach, the p-value is evaluated as the strength of the evidence

Table 5: Results of informative Bayesian t-tests (large sample size).

Parameter	MD	PSD	Bayes factor	t score	df	p
Glucose (mg/dL)	22.657	3.559	2.311e+7	6.366	898	**<0.001**
Triglyceride (mg/dL)	31.815	10.603	6.256	3.001	898	**0.003**
TC (mg/dL)	-19.091	3.750	22351.910	-5.092	898	**<0.001**
HDL-C (mg/dL)	-3.417	0.609	321857.523	-5.614	898	**<0.001**
LDL-C (mg/dL)	-21.231	2.634	2.105e+12	-8.062	898	**<0.001**
Urea (mg/dL)	1.769	0.697	1.801	2.540	898	**0.011**
Creatinine (mg/dL)	0.121	0.042	4.355	2.874	898	**0.004**
Hemoglobin (g/dL)	-0.315	0.104	6.540	-3.016	898	**0.003**
Hematocrit (%)	-0.853	0.271	9.689	-3.147	898	**0.002**
RDW (%)	0.040	0.103	0.082	0.389	898	0.698
WBC (103/µL)	0.915	1.156	1.228e+6	5.859	898	**<0.001**
Platelet (10e3/uL)	6.029	5.106	0.151	1.181	898	0.238
MPV (fL)	0.097	0.091	0.133	1.065	898	0.287
NEUT (103/µL)	0.639	0.121	54301.389	5.271	898	**<0.001**
LYMP (103/µL)	0.119	0.055	0.795	2.186	898	**0.029**

MD: Mean difference, PSD: Pooled standard error difference, df: Degree of freedom, TC: Total cholesterol, HDL-C: High density lipoprotein cholesterol, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC:White blood cell, MPV, Mean platelet volume, NEUT: Neutrophil; LYMP: Lymphocyte.

against the null hypothesis. The closer the p-value gets to 0, the lower the probability that the null hypothesis is true (18). The biggest problem of the Fisher approach is that it does not give a critical p-value to reject the null hypothesis. Although the p-value is a quantitative measure against the null hypothesis, it does not provide an idea of how strong the evidence is (19). Neyman-Pearson solved this problem by putting forward a critical value.

The Bayesian approach criticizes the p-value, as it does not indicate the probability of truth of the null hypothesis and is incorrectly interpreted as the probability of the truth alternative hypothesis. However, there are also some limitations of the Bayes factor (20). It is not clear at what point the Bayesian factor should be accepted as a confirmation of one of the two hypotheses. Moreover, the Bayes-Factor is influenced by the priori distribution of effect size. In general, reaching a Bayes factor of 10 or greater is sufficient for early completion of a study- (15). The Bayesian factor depends on the t-statistic, the degree of freedom and the a priori distribution of the parameter. The Bayes factor and p value are interrelated to each other. Although the p value has the identical meaning for disparate samples, for the same t value, the Bayes factor differs with the changes of sample sizes. In small sample sizes, the Bayes factor allows us to obtain proof for the null hypothesis. In large sample sizes it allows for the detection of even small deviations. As the sample size increases, the statistical power of a study also increases.

In the present study we evaluated the effect of sample size on Bayesian and frequentist results. According to our results both Bayesian and frequentist approaches had higher power in order to detect small differences with larger sample sizes. We obtained higher values of Bayes factors in the larger sample size, indicating support for the alternative hypothesis relative to the null hypothesis that were not seen in the small sample size. Similarly, p values reached statistical significance with the higher sample size, which were found to be greater than 0.05 in the small sample size (except for glucose levels, p=0.031). One of the most important steps in the planning of scientific studies is to allocate resources in such a way that they have sufficient power to build statistical outcomes when a difference exists. With high statistical power, the p value is expected to be small and give the same information as the Bayes factor. High powered studies are associated with lower levels of type-II error rate. Therefore, analysis of high-powered studies can be applied to both frequentist and Bayesian statistics. Similar to the frequentist approach, the Bayesian approach can falsely support the null hypothesis in small sample sizes.

Kelter R. investigated the effect of sample sizes on Bayesian results. In that study, it was shown that an increase in sample size reduces the type II error rate to zero in both Bayesian and frequentist approaches (21). Bayesian inference required a higher number of sampling data for the same type II error rate compared to frequentist tests. In order to detect little deviations between the two samples, Bayesian inference needed higher sample sizes for the identical type II error

Table 6: Posterior distribution statistics (large sample size).

Parameter	Posterior			%95 CI	
	Mod	Median	Varyans	Lower limit	Upper limit
Glucose (mg/dL)	22.657	22.657	13.569	15.437	29.877
Triglyceride (mg/dL)	31.815	31.815	29.363	21.195	42.436
TC (mg/dL)	-19.091	-19.091	8.038	-24.647	-13.534
HDL-C (mg/dL)	-3.417	-3.417	0.384	-4.632	-2.202
LDL-C (mg/dL)	-21.231	-21.231	5.339	-25.760	-16.702
Urea (mg/dL)	1.770	1.770	0.625	0.221	3.318
Creatinine (mg/dL)	0.121	0.121	0.000	0.080	0.162
Hemoglobin (g/dL)	-0.315	-0.315	0.011	-0.525	-0.105
Hematocrit (%)	-0.853	-0.853	0.103	-1.482	-0.223
RDW (%)	0.040	0.040	0.012	-0.177	0.260
WBC (103/µL)	0.915	0.915	0.133	0.201	1.630
Platelet (10e3/uL)	6.029	6.029	16.596	-1.956	14.013
MPV (fL)	0.097	0.097	0.009	-0.090	0.283
NEUT (103/µL)	0.640	0.640	0.027	0.319	0.960
LYMP (103/µL)	0.119	0.119	0.003	0.013	0.225

CI: Confidence interval, TC: Total cholesterol, HDL-C: High density lipoprotein cholestero, LDL-C: Low density lipoprotein cholesterol, RDW: Red cell distribution width, WBC:White blood cell, MPV: Mean platelet volume, NEUT: Neutrophil, LYMP: Lymphocyte.

rate. For medium or large effect sizes, the situation was less problematic. It was stated that, for small sample sizes, it was necessary to conduct further research and evaluate the accuracy of Bayesian tests (21). Another simulation study which was also performed by Kelter R. showed that non-parametric Bayesian two-sample tests had lower type I error rate compared to the Mann-Whitney U test (22). In contrast, the strength of the Bayesian two-sample tests was found to be slightly lower than the frequentist methods. The ability of Bayesian tests to control type I and II error rates and detect an existing difference depends on the power of a priori modeling.

CONCLUSION

P value and Bayes factor should be interpreted correctly by the researcher. According to our results, both Bayesian and frequentist approaches depend on the proportion of sample errors, which depends on the sample size. Similar to the frequentist approach, the Bayesian approach had low accuracy for the acceptance of the null hypothesis in small sample sizes.

Ethics Committee Approval: This study was approved by Bakırköy Dr. Sadi Konuk Training and Education Hospital Ethical Committee (Date: 04.10.2021, No: 2021-19).

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- C.Y., M.G.G., E.Y.; Data Acquisition- C.Y.; Data Analysis/Interpretation- C.Y., M.G.G.; Drafting Manuscript- C.Y., M.G.G., E.Y., F.N.T.Ç.; Critical Revision of Manuscript- M.G.G., E.Y; Final Approval and Accountability- C.Y., M.G.G., E.Y., F.N.T.Ç.; Material and Technical Support- F.N.T.Ç.; Supervision- M.G.G., E.Y., F.N.T.Ç.

Conflict of Interest: The authors have no conflict of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

1. Rindskopf D. Overview of Bayesian statistics. Eval Rev 2020;44(4):225-37.
2. Beerenwinkel N, Siesbourg J. Probability, statistics, and computational science methods. Methods Mol Biol 2019;1910:33-70.
3. Sami W, Alrukban MO, Waqas T, Asad MR, Afzal K. Sample size determination in health research. J Ayub Med Coll Abbottabad 2018;30(2):308-11
4. Sacco W. Statistical power considerations in the use of cost-effectiveness analysis. Professional Psychology 1982;13(5):752-8.
5. Mazen AM, Graf LA, Kellog KE, Hemmasi M. Statistical power in contemporary management research. The Academy of Management Journal 1987;30(2):369-80.
6. Kim HY. Statistical notes for clinical researcher: Type I and type II errors in statistical decision. Restor Dent Endod 2015;40(3):249-52.
7. Mozaffarian D, Benjamin E, Go A, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics-2016 update. A report from the American Heart Association. Circulation 2016;133(4):e38-360.
8. Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, Abyu G, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. J Am Coll Cardiol 2017;70(1):1-25.
9. Sweetnam PM, Thomas HF, Yarnell JW, Baker IA, Elwood PC. Total and differential leukocyte counts as predictors of ischemic heart disease: The caerphilly and speedwell Studies. Am J Epidemiol 1997;145(5):416-21.
10. Weijenberg MP, Feskens EJ, Kromhout D. White blood cell count and the risk of coronary heart disease and all-cause mortality in elderly men. Arterioscler Thromb Vasc Biol 1996;16(4):499-503.
11. Gupta R, Gupta VP, Sarna M. Bhatnagar S, Thanvi J, Sharma V, et al. Prevalence of coronary heart disease and coronary risk factors in urban Indian population: Jaipur Heart Watch -2. Indian Heart J 2002;54(1):59-66.
12. Ranjith MP, Divya R, Metha VK, Krishnan MG, KamalRaj R, Kavishwar A. Significance of platelet volume indices and platelet count in ischaemic heart disease. J Clin Pathol 2009;62(9):830- 3.
13. Pizzulli L, Yang A, Martin JF, Lüderitz B. Changes in platelet size and count in unstable angina compared to stable angina or non-cardiac chest pain. Eur Heart J 1998;19(1):80-4.
14. Houghton DE, Koh I, Ellis A, Key NS, Douce DR, Howard G, et al. Hemoglobin levels and coronary heart disease risk by age, race, and sex in the reasons for geographic and racial differences in stroke study (REGARDS). Am J Hematol 2020;95(3):258-6.
15. Stefan AM, Gronau QF, Schönbrodt FD, Wagenmakers EJ. A tutorial on Bayes Factor Design Analysis using an informed prior. Behav Res Methods 2019;51(3):1042-58
16. Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: an explanation for new researchers. Clin Orthop Relat Res 2010;468(3):885-92
17. Pernet C. Null hypothesis significance testing: a short tutorial. F1000Res 2015;4:621.
18. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations Eur J Epidemiol 2016;31(4):337-50.
19. Domenech RJ. The uncertainties of statistical "significance" Rev Med Chil 2018;146(10):1184-9.
20. Morey RD, Romeijn JW, Rouder JN. The philosophy of Bayes factors and the quantification of statistical evidence. J Math Psychol 2016;72:6-18.
21. Kelter R. Analysis of type I and II error rates of Bayesian and frequentist parametric and nonparametric two-sample hypothesis tests under preliminary assessment of normality. Computational Statistics 2021;36:1263-88.
22. Kelter R. Power analysis and type I and type II error rates of Bayesian nonparametric two-sample tests for location-shifts based on the Bayes factor under Cauchy priors. Computational Statistics Data Analysis 2022;165(21):107326.