The Greek E-Journal of Perioperative Medicine 2013; 11: 22-30 (ISSN 1109-6888) www.anesthesia.gr/ejournal
Ελληνικό Περιοδικό Περιεγχειρητικής Ιατρικής 2013; 11: 22-30 (ISSN 1109-6888) www.anesthesia.gr/ejournal

Review of statistical methods reported in oral presentations during seven Greek national anesthesiology conferences

Aslanidis T[1] MD, Charitidou S[1] MD, Matiaki E[1] MD, Milonas A[2] MD, Georgopoulou E[2] MD, Anagnostara E[1] MD, Aslanidis K[3]

ABSTRACT

Review of statistical methods reported in oral presentations during seven Greek national anesthesiology conferences

Aslanidis T, Charitidou S, Matiaki E, Milonas A, Georgopoulou E, Anagnostara E, Aslanidis K

The application of statistical methods in order to extract safer conclusions from samples of medical data has become a key methodology for synthesis and evaluation in any medical research. This study makes a retrospective overview of statistical methods used for oral presentations in the summaries of Greek anesthesia conferences and tries to “capture” the change in the use of statistics in recent years. Nine hundred and twenty five oral presentations from seven Greek anesthesia conferences were included for further analysis. The results, recorded an increase of randomized studies, the majority of which, had the character of single-blind study, using relatively small samples (n=31-51). Nevertheless, there is a trend towards studies with larger samples. References to the software used for processing of data also increase over the years. The majority of statistic analyses are done by using descriptive statistical methods.Tests for normality of the data are presented in the last conferences, but there is no reference to the power of any study. Extrapolations are mainly based on p value, although, over the years, the use of confidence intervals is appearing also as an alternative. Finally, more and more and different methodologies and statistical analysis tests are being selected. Although these findings refer to only 25% of all abstracts of the 28 conferences carried out until the beginning of this study (2009), they give us a first insight look on the use of statistics among Greek anesthesiologists. The last seems to ameliorate as the education in that field getting better with time.

(1)1st Anesthesia Department,
(2) 2nd Anesthesia Department ,
G.H. “Hippokratio”, Thessaloniki, Greece,
(3) Mathematician

INTRODUCTION

Scientific knowledge begins from experience. Thus in early stage, science is a systematic gathering-recording of primary data. Further data analysis creates relations between the observed and the expected,therefore creating the concept of prognosis. Statistics is the mathematical science of stochastic model application in order to describe natural phenomena and create more objective relations between primary "real-experience-induced" recorded data. Why use statistics? Because natural phenomena have an intense probabilistic character; and because the same causes (data) can have a variety of results[1].

Biostatistics (referred also as biometry or biometrics) is the application of statistics to a wide range of topics in biology and medicine. Its utility lies upon the fact that it quantifies the uncertainty and, in that way, helps us making the best of the available decisions[2].

Furthermore,anesthesiology practice includes a continuous gathering, evaluation and analysis of data out of a minor sample (one patient) and management of that information in accord to previous statistical conclusions from previous samples (i.e. scientific knowledge).

Hence, statistical education of anesthesiologysts today is becoming an essential aspect of their training. This study is a retrospective overview of statistical methods as reflected in the summaries of oral presentations in Greek anesthesia conferences and to detect the change in the use of statistics in recent years.

MATERIAL AND METHODS

Initially, all oral presentations from the eighteen first conferences of Greek Anesthesiology Society and the first ten conferences of the Greek Society of Anesthesiology and Intensive Care of Northern Greece were gathered. Then with help of a web-based random number generator(www.psychicscience.org/random.aspx) 2401 sequences of 28 numbers were created. The seven more often appeared numbers were chosen. The latter corresponded to 2nd, 6th,7th,8th,9th and 10th conference of Greek Society of Anesthesiology and Intensive Care of Northern Greece and the 17th conference of Greek Anesthesiology Society.

Oral presentations from the aforementioned seven conferences were randomly appointed to five researchers. Parameters about statistical methods (like study design, sampling, descryptive and inferential statistics) used in each presented study were recorded. The aim of the study was to record the use of statistic and not to assess the methodology of each study, so it was considered a priori that the choice and use of recorded statistical methods were the appropriate one.

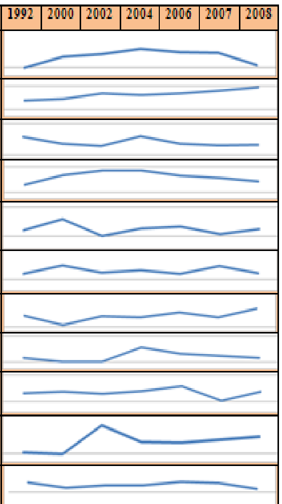
Final data records were rechecked from a 6th independent researcher and then further analysis was conducted with Microsoft Office Excel 2007,modified with Deakin University XLStatistics® 10.30.10 and Analyse-it Software Ltd. Analyse-it® SE 2.22 tools. Results are presented as descriptive statistics,while choice of measures of dispersion was based on the most reliable representation of the sample.

RESULTS

In the final data analysis there were included records from 925 studies. The average number of studies presented per conference (s.p.c) was 132 (range 87-174); the thematic distribution of which is showed on table 1.

Table 1. Coverage distribution per subject section per conference. A trend with relative change through time is also displayed.

YEAR	1992	2000	2002	2004	2006	2007	2008	1992	2000	2002	2004	2006	2007	2008
Emergency Medicine	0,00	5,75	7,14	10,00	8,20	8,05	1,50							
Regional anesthesia	7,75	9,20	13,57	12,14	13,93	16,09	18,80							
General anesthesia	29,46	18,39	14,29	30,00	18,03	14,94	15,79							
Pediatric anesthesia	1,55	3,45	4,29	4,29	3,28	2,87	2,26							
Cardio anesthesia	3,88	11,49	0,00	5,00	6,56	1,15	4,51							
Pain management	10,08	24,14	11,43	15,71	9,84	22,99	10,53							
Obstetric anesthesia	6,98	1,15	6,43	5,71	9,02	5,75	11,28							
Neuro anesthesia	0,78	0,00	0,00	2,86	1,64	1,15	0,75							
ICU	16,28	19,54	15,00	20,00	29,51	1,72	19,55							
Special anesthetic Techniques	0,78	0,00	15,00	6,43	5,74	7,47	9,02							
Other	18,60	8,05	12,86	12,86	19,67	17,82	6,02							
TOTAL	100	100	100	100	100	100	100							



More than half of the studies (480) were experimental and 445 were observational. The majority of the latter were retrospective (average 18, 84 s.p.c.), while on the other hand most of the experimental ones were with serial control (average 27,42 s.p.c.). The rest types of study designs are presented in table 2.

Table 2. Study design distribution.

A)Obs. **Year of Conference**	Case report	Meta-Analysis	Obs. Descriptive	Obs. Cohorts	Obs Retro-spective	Historical Retro-spective	Survey
1992	6	0	42	23	6	0	2
2000	3	1	15	3	8	0	1
2002	3	1	13	10	28	2	1
2004	1	0	25	13	26	3	0
2006	14	0	3	1	27	0	7
2007	14	3	6	23	22	3	11
2008	16	2	15	21	12	18	4

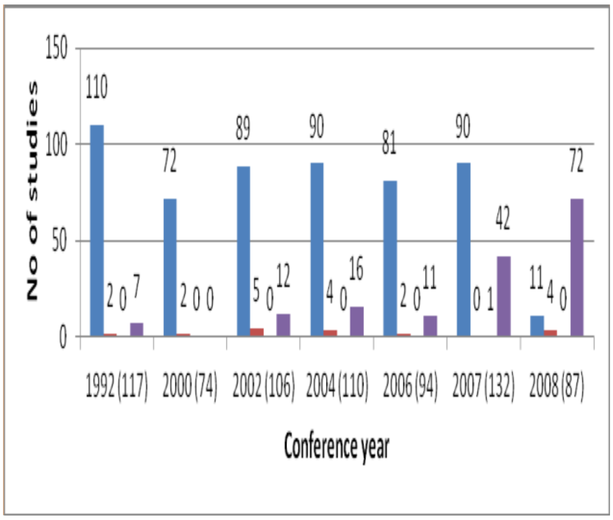
B)Exp. **Year of Conference**	Parallel Control-Randomised	Parallel Control-Non Randomised	Serial Control-Self Control	Serial Control-Cross-Over	With External Control (historical data)	Without Contol
1992	4	33	5	0	1	7
2000	18	24	9	0	0	4
2002	37	26	5	1	0	13
2004	38	20	8	0	0	5
2006	43	19	16	0	0	5
2007	64	18	7	0	0	3
2008	19	21	5	0	0	2

A) Obs. (Observationals) B) Exp. (Experimentals)

Sampling randomization and blinding were also recorded.In regard to sampling, in the first 2 conferences the ratio of randomized to non-randomized experimental studies was 11:28,5 per conference

This relationship is reversed in the rest, more recent, conferences to 40,2:20,5s.p.c. Apropos blinding, though the majority of the presentations referred to single blind studies (average 77,5s.p.c.) and only an attempt for triple blinding was recorded, a tendency towards increa-se in open study design was revealed (Figure1).

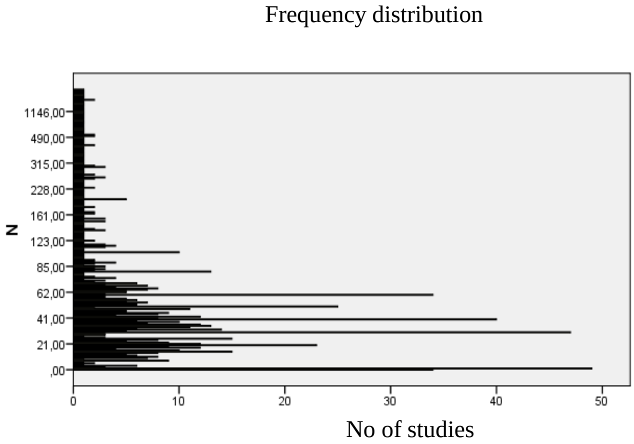
Figure 1. Distribution of prospective studies according to blindness design (blue-single blind, red-double blind, green-triple blinde, puple-open).



Sampling records revealed a tendency towards larger studies. From 1992 to 2008 the range of the means for sampling was 90 to 252 with tendency equation $f(x)= 6{,}692x^2-30{,}609x+132{,}49$ ($R^2=0.833$). On the other hand, the range of the medians of the studies was 31 to 50 (tendency equation $f(x)=-1{,}208x^2-10{,}57x+23{,}35$ ($R^2=0.709$); the difference is probably due to conduction of larger studies in more recent years (Figure 2).

There was no report about power (i.e. minimi-zing error type II) calculation. Reference to the software used for data analysis is more of-ten from 2004 and on. Interestingly, where the-re is mention to software, 59,6% of the authors preferred IBM SPSS for their analysis, while the rest 40,4% of them referred to MS Office Excel.

Figure 2. Bar chart of distribution of samples (n) used in the studies presented.

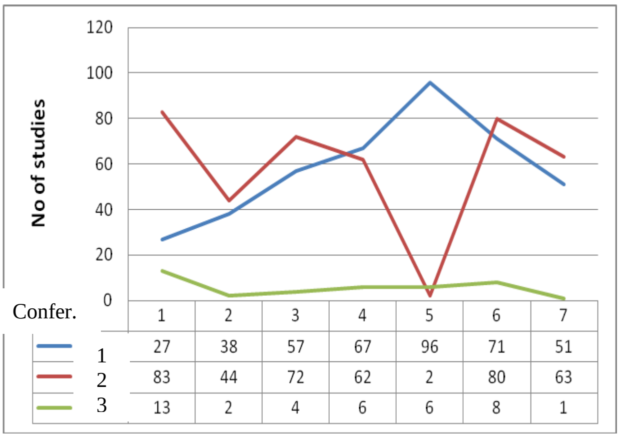


Descriptive statistics appeared in 627 studies. Measures of dispersion were the most popular way (used from 83,2% of the authors) to de-scribe the sample, while other used graphical (7.9%) or mixed (8,9%) summaries.

Reference to normality tests appear also more often from 2006 and forward. Nonetheless, ve-ry few studies actually specify the test they u-sed. Whenever it is specified, Kolmogorov-Smirnov criterion is the one of the two methods (in 75% of the studies) mentioned and Shapiro -Wilk (W-test) is the other.

Hypothesis testing was mentioned in 857 stu-dies. Yet, in the majority of them conclusion a-re extracted from descriptive analysis, while few are those who based their conclusion on p value calculation (yet fewer studies specify that value); only from 2007 and forwards, con-fidence intervals appeared in oral presentations. Unfortunately, there is still a scarce minority of studies that did not clarify the way authors reached their conclusions. (Figure 3).

Figure 3. Way of conclusion reaching. (1-p value, 2-descriptive statistics, 3-not mentioned).



Confer.	1	2	3	4	5	6	7
1	27	38	57	67	96	71	51
2	83	44	72	62	2	80	63
3	13	2	4	6	6	8	1

Finally, even though a variety of tests was used for inferential statistical analysis, a clear preference to t test, Mann -Whitney U test and chi-square (χ^2) test was recorded.

DISCUSSION

Statistics in medical research first appeared in 1917 when A. Cohn used them for studying the effect of digitalis in pneumonia, but it is not until David Sackett invented the term “evidence-based medicine” in 1992 that interest about statistics increased. A query in Pubmed about statistics in medical research retrieves 1917 results for year 1991 and 8133 for year 2011, while query with term “statistics in anesthesiology” reveals the same tendency (45 results in 1991 and 307 in 2011)[3].

In regard to anesthesiology in Greece, no previous attempt had been made to photograph the use of statistics either in published articles or during meetings.

The results of our analysis show that the majority of presented studies were based on observational designs, mainly retrospective, although there is a tendency towards increase of surveys or meta-analyses. An attempt in recent years for larger samples is also defined, even though the average number of the sample remains the same. The inversion of the randomised to non-randomised studies ratio may be explainned by an effort to minimize systemic bias. Yet, lack of reporting calculation of power remains an issue. Open label studies seem to increase versus a decrease in single blind designs and meanwhile, number of double blind studies remains constant. An also interesting discovery is that statistical software used for data analysis is rarely mentioned. However when a report was done about it, only two packages are recorded. The majority of authors chose numerical summaries for describing their sample, and very few mentioned conduction of normality tests. The latter may be considered a form of word/space economy. Finally, inferential statistics are mainly based on p value calculation; maybe because this way of hypothesis testing is simpler or because most physicians are familiar with this method[4].

Even though, the aforementioned conclusions are in regard to the quarter of the total number of studies presented in the 28 Greek anesthesiology conferences conducted till the start of this study (2009), yet they give us an overall impression of the use of statistics throughout time among anesthesiologists and intensivists in Greece.

There are few studies in the literature about use of statistics in oral presentations or posters during anesthesiology conferences[5]. Most of the published studies record the use and evaluate the adequacy of statistical methodology appeared in anesthesiology journals[6-10]. Yet, all of them highlight the frequency of erroneous use of statistics in anesthesiology articles and the fact that as a result of marked advances in the use of statistics in anesthesiology journals, readers are seeing more applications of higher-level statistics. Readers must therefore acquire greater knowledge of statistics in order to understand the methods used in original research publiccations.Moreover, higher level statistics ensure a better "fate" for the study[11].

Based on the above-mentioned, there are several suggestions that could improve future use of statistics in studies presented in national conferences. Report to statistical software used for data analysis is an easy but essential first step. The variety of available tools is increasing. But, when choosing an analytical tool to use, there are many factors to consider. Does it run natively on your computer? Does the software provide all the methods you use? If not, how extensible is it? Does that extensibility use its own language, or an external one (e.g. Python, R, SQL, C) that is commonly accessible from many packages? Does it fully support the style (programming vs. point-and-click GUI) that you like? Are its visualization options (e.g. static vs. interactive) adequate for your problems? Does it provide output the form you prefer (e.g. cut & paste vs. LaTeX integration)? Does it handle large enough data sets? Do your colleagues use it so you can easily share data and programs? Can you afford it? Even though SPSS and MS Office Excel are the most popular packages in our study, this is not the case worldwide, where R and SAS are gaining more users[12-14].

Apart from this, obligatory reference to normality test selected is also essential for better presentation of the results. Specification of test (e.g. Shapiro-Wilk or Anderson-Darling) should also be an aspect the authors should be careful about. The same could apply for power report, yet regulations for specific value (0.8) could be more lenient than in journals[15-16]. Finally, an effort should be made to decrease the use of calculation of p value for hypothesis testing, as the method has a lot of disadvantages, the detailed report of which is beyond the scope of this article. Confidence intervals or

Baye's theorem application seem to be more reliable alternatives[3,17-21].

So, should we become statisticians? Education of every physician in biostatistics is essential not only because conducting trials is easer, but because critical evaluation of different information is better in every day practice. Nevertheless, the reply to the question is no; the answer lies within better understanding of the concepts behind statistics and closer cooperation with a statistician. Proficiency in the use of computational tools helps, but the "mathematical physiology" under the calculation is what makes the difference. And, in one way or another, the question will be always the physiology of the patient standing in front of us[21].

REFERENCE

1. Δεραμάνης Α.: Συνορθώσεις Παρατηρήσεων και Θεωρία Εκτίμησης, Εκδ. Ζήτη 1992.
2. Τριχόπουλος Δ, Τζώνου Α, Κατσουγιάννη Κ: Βιοστατιστική. Εκδ. Γρ. Παρισιάνος 2000.
3. http://www.ncbi.nlm.nih.gov/pubmed?term=statistics in medical research. Accessed at 1 Sep 2012.
4. Γαλάνης Π. Η λανθασμένη εφαρμογή των τιμών P και του ελέγχου των υποθέσεων στη βιοϊατρική έρευνα. Αρχεία Ελλ. Ιατρικής 2010;27:691-707.
5. Lee FJ, Wu CC, Peng SY, Fan KT. Are P values and statistical assessments in poster abstracts presented at annual meetings of Taiwan Society of Anesthesiologists relative to the characterristics of hospitals? Acta Anaesthesiol Taiwan 2007 ; 45: 163-167.
6. Barbosa FT, de Souza DA. Frequency of the adequate use of statistical tests of hypothesis in original articles publicshed in the Revista Brasileira de Anesthesiologia between January 2008 and December 2009. Rev Bras Anestesiol 2010; 60 :528-536.
7. Greenfield ML, Mhyre JM, Mashour GA, et al. Improvement in the quality of randomized controlled trials among general anesthesiology journals 2000 to 2006: a 6-year follow-up. Anesth Analg 2009;108:1916-1921.
8. Rubio García B, Rodríguez Zazo A, Martínez Terrer T, et al. Use of statistics and the accessibility of original articles published in 3 anesthesiology journals. Rev Esp Anestesiol Reanim 2010; 57: 281-287.
9. García-Alamino JM, Parera A, Ollé G, et al. Clinical trials published in Revista Española de Anestesiología y Reanimación: characteristics and quality of design. Rev Esp Anestesiol Reanim 2007 ; 54: 333-339.

10. Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. Br J Anaesth 2003; 90 :514-516.
11. Meininger D, Bück M, Bohlmann S,et al. Gone with the wind or... Fate of scientific articles presented at large anesthesia congresses-an update. Anaesthesist 2011 ;60 :118-124.
12. Texeira A, Rosa A, Calapez T. Statistical Power Analysis with Microsoft Excel: Normal Tests for One or Two Means as a Prelude to Using Non-Central Distributions to Calculate Power, Journal Stat. Ed. 2009;17:1-21
13. Marcelo GA, Bruno L, Alyson LCO,et al. On the Numerical Accuracy of Spreadsheets. Journal Stat. Soft. 2010;34:1-29
14. Muenchen RA. The popularity of data analysis software 2012. Available at http://r4stats.com/articles/popularity. Accessed at 9 Sep 2012.
15. Fox J, Andersen R. Using the R statistical computing environment to teach social statistics courses. McMaster University ed. 2005;1-36
16. Ward MM. Primer: measuring the effects of treatment in clinical trials. Nat Clin Pract Rheumatol 2007;3:291-297
17. David MC. How to make clinical decisions from statistics. Clin Exp Optom 2006;89:176-183
18. Nishida T. Data transformation and normalization. Rinsho Byori 2010; 58:990-997
19. Henderson AR. Testing experimental data for univariate normality. Clin Chim Acta 2006;366:112-129
20. Tenenbaum JB, Kemp C, Griffiths TL, et al. How to grow a mind: statistics, structure, and abstraction. Science 2011; 331:1279-1285.
21. Song F, Parekh S, Hooper L,et al. Dissemination and publication of research research findings: an updated review of related biases. Health Technol Assess 2010; 14:1-193.

Key words: statistics, anesthesiology, oral presentations

Corresponding author:

Theodoros Aslanidis

Address: Doridos str.4,

P.C. 54633

Thessaloniki

Tel.: +30-6972477166

e-mail:thaslan@hotmail.com