

# Department of Epidemiology and Biostatistics

**513-607A**

Principles of Inferential Statistics

Final Examination: In-class portion

(2 hours)

Q..... Points

Q1 ...../2

Q2 ...../2

Q3 ...../2

Q4 ...../2

Q5 ...../3

Q6 ...../3

Q7 ...../3

Q8 ...../4

Q9 ...../4

December 3, 1986

Name:

---





**THE COST IMPLICATIONS OF ACADEMIC GROUP PRACTICE**  
A Randomized Controlled Trial

DAVID I. COHEN, M.D., M.Sc., DANIEL BRESLAU, B.A., DAN K. PORTER, M.S.S.A., HAROLD I. GOLDBERG, M.D., NEAL V. DAWSON, M.D., CHARLES O. HERSHEY, M.D., JAR-CHI LEE, M.S., CHRISTINE E. MCLAREN, PH.D., AND NAOMI BRESLAU, PH.D.

**Abstract** We evaluated the reorganization of a general medical clinic into several group practices, using equivalent groups of patients and physicians in a randomized controlled trial. The group practice, unlike the traditional clinic, provided decentralized registration, clinic coverage five days a week, and telephone coverage at night and on weekends. Residents worked in small groups with an attending physician, nurse practitioner, and receptionist. All financial activity involving a sample of 2299 patients was followed during the 11-month intervention.

The total hospital charges per patient were 26

percent lower for the patients seen in the group practice than for those seen in the traditional clinic ( $P=0.003$ ). This difference was primarily attributable to inpatient charges, which were 27 percent lower per patient hospitalized ( $P=0.004$ ). The mean length of stay was 8.3 days among group-practice patients and 10.5 days among traditional-clinic patients ( $P=0.011$ ).

We concluded that organizational changes to improve outpatient access and to integrate inpatient and outpatient services can decrease medical charges. (N. Engl. J. Med. 1986; 314:1553-7.)

**BACKGROUND**

**The Firms**

The Department of Medicine at Cleveland Metropolitan General Hospital is organized into four general medical services, called firms, each consisting of a 28-bed inpatient unit and an outpatient service.<sup>20</sup> House officers are randomly assigned to one of the firms when they enter the program, and remain with that firm for the duration of their training. Patients are also randomly assigned to one of the four firms upon first contact with the Department of Medicine, either in the outpatient clinic or at the time of admission to the hospital. They are linked by computerized records to a single resident, who serves as their primary care

physician. When that house officer completes the program, the patient is assigned to another resident on the same firm. In this manner, the firm system ensures continuity of care. Moreover, the ongoing randomization into the firms has created four similar groups of patients and physicians, enabling us to study the effects of an intervention in one or two firms while using the others as controls.<sup>21</sup> For this study, two firms were randomly assigned to receive the experimental intervention, and the other two served as controls.

In the “study design and analysis” section, the authors say that “the analysis was carried out separately with both patients and residents as the units of analysis”.

- Which analysis do you think is more appropriate, and why? (You might also consider the firm as a third possible unit of analysis).

## REPEATABILITY OF ASEPSIS WOUND SCORING METHOD

Sir, In *The Lancet* of Feb 8 (p. 311) we described a scoring method (ASEPSIS) for postoperative wound infections for use in clinical trials of antibiotic prophylaxis. The purpose was to reveal differences between the efficacy of antibiotic regimens for the whole range of abnormal wound healing, from minor erythema and serous discharge to severe purulent infection. The power of a clinical trial using this method would be greater than one that registered only severe infections.

The score was composed partly of points awarded daily for the clinical appearance of the wound during the first week and partly of points given for any additional treatment and prolonged hospital stay required as the result of infection. We indicated that the allocation of points from the clinical appearance of the wound might be open to bias, despite being based on objective criteria. We have now assessed that risk using a second independent observer on a further 51 patients.

One of us (A. P. R. W.) had scored all the wounds from the start of the current trial of antibiotic prophylaxis in cardiac surgery but the second observer (A. W.) had no other connection with the trial and was not aware of the antibiotic regimen used in each patient. After an introductory period of

one week, the wounds of 51 patients were examined independently by both observers and, at the end of the study, a comparison was made only between the points given for the clinical appearance of the wound. The mean and standard deviation of the differences in score was calculated and a coefficient of repeatability was calculated.<sup>1</sup>

Of the 51 patients examined all had sternal wounds and 34 had leg wounds following vein harvest. The ranges of sternal wound scores were 0-22 points for one observer and 0-18 points for the other. For the leg wounds, the scores were between 0 and 13 points and 0 and 15 points. The differences between each pair of scores are shown in the figure. The mean difference was 0 and 0-1 points from sternal and leg wound scores, respectively. The coefficient of repeatability was 4-1 points for the sternal wound and 3-2 points for the leg wound. Thus, 95% of the differences in the wound scores by two independent observers would be expected to be less than these values.

The scoring system used to assess the clinical appearance of a wound is reproducible, and its potential usefulness in clinical research or infection surveillance is confirmed.

Department of Clinical Microbiology,  
University College Hospital,  
London WC1B 6AU  
and Department of Cardiothoracic Surgery,  
Middlesex Hospital,  
London W1

A.P.R. Wilson  
A. Webster  
R.N. Gruneberg  
T. Treasure  
M. F. Sturridge

The authors do not give a definition for the “coefficient of repeatability” they calculate from the data, but refer to an earlier Lancet article.

- a. From the diagrams and the text, what must be the definition of their “coefficient of repeatability”.
  
  
  
  
  
  
  
  
  
  
- b. What other index of repeatability can you suggest?



In an investigation to determine the influence of sampling variability on the diagnostic yield of liver biopsy, two consecutive samples were obtained from each of 75 patients by redirecting the biopsy needle through a single entry site. The results are shown below:

<b>Number of Patients with this many abnormal specimens</b>	<b>Number of specimens abnormal</b>		
	0	1	2
	12	6	57

a. What is the (average) proportion of abnormal specimens per patient?

b. How well does the binomial distribution fit these data?

**ORAL PENICILLIN PROPHYLAXIS IN THALASSEMIA  
AND IN SICKLE CELL ANEMIA**

*To the Editor:* Gaston et al. (June 19 issue)<sup>1</sup> showed that in sickle cell anemia the incidence of infections due to *Streptococcus pneumoniae* is reduced by daily use of oral penicillin. Increased susceptibility to bacterial infections is common in children with other severe hemoglobinopathies, such as  $\beta$ -thalassemia, when transfusion is poor; the most likely organisms are pneumococci, staphylococci, and streptococci.<sup>2</sup> In this case, prophylaxis with oral penicillin can be as useful as it is in sickle cell anemia.

The benefit of prophylactic oral penicillin therapy, 1,000,000 U daily, added to a moderate transfusion regimen – i.e., 1 U (400 to 500 ml) of whole blood monthly, was tested for in a group of patients 3 to 15 years old in whom  $\beta$ -thalassemia had been diagnosed according to classic criteria.<sup>3</sup> Episodes of infection of any cause were recorded on a monthly basis in 12 patients, first during a regimen of transfusion alone and then during transfusion plus penicillin; a comparison was made with a group of 39 normal children matched for age and sex and regularly examined at a pediatric outpatient clinic.<sup>4</sup> The results are summarized in Table 1; clearly, oral penicillin therapy reduced the number of episodes of infection to a nearly normal level.

Table 1. Infections in Children with $\beta$ -Thalassemia Treated with Transfusion and Prophylactic Penicillin and in Normal Children		
	MONTHS OF FOLLOW-UP	EPISODES OF INFECTION/MONTH
Controls (39)	1812	0.062
Patients		
Transfusion alone	310	0.220*
Transfusion plus penicillin	124	0.089 <sup>†</sup>

\*Significantly higher than control (P = 0.01).  
<sup>†</sup>Not significantly higher from control.

Industrialized countries, the main basis of treatment for severe  $\beta$ -thalassemia is frequent blood transfusion with regular provision of chelating agents; except after splenectomy, susceptibility to bacterial infections is a low as in normal children. In most developing countries however, this approach is difficult because of the short supply of adequate blood transfusions and the cost of chelating agents. In these countries, we think that oral (or parenteral) prophylactic penicillin therapy can be recommended for  $\beta$ -thalassemia as well as for sickle cell anemia in children.

P. Colonna, M.D. F.-Z. Ardjoun, M.D. Algiers, Algeria Hopital Mustapha

- a. How does one formally compare the number of episodes of infection/month in the controls (0.062) and in the patients receiving transfusion alone (0.220)?
- b. Use a point estimate to quantify the reduction that the authors call “to a nearly normal level”. How would you calculate an internal estimate for the reduction?
- c. Comment on the comparisons the authors made.d.  
 What type of study would you have proposed to assess the benefit of prophylactic oral penicillin therapy?



**DO SEIZURES IN CHILDREN CAUSE INTELLECTUAL DETERIORATION?**

Jonas H. Ellenberg, PH.D., Deborah G. Hirtz, M.D., and Karin B. Nelson, M.D.

**Abstract** We studied whether the occurrence of seizures in childhood affected intellectual performance. We compared the full-scale IQs at seven years of age of children who had experienced one or more nonfebrile seizures with the IQs of their seizure free siblings who were tested at the same age in a large longitudinal study. Among 98 children with seizures, the mean score on IQ tests at seven years was not significantly different from the mean score of their siblings. Mental retardation was more common among the children with seizures, but the excess was accounted for by children who had neurologic abnormalities before the first seizure. We also examined the IQ before and after the

onset of seizures in 62 children whose first seizure occurred in the interval between psychometric examinations given at four and seven years of age. The IQ at seven years in the children with seizures did not differ significantly from that in controls matched for IQ (as determined at the four-year assessment), sex, race, and socioeconomic status. Thus, in both the sibling-control comparison and the comparisons made between controls and subjects before and after the onset of seizures, the occurrence of nonfebrile seizures was not associated with a significant change in full-scale IQ. (N. Engl. J. Med. 1986; 314:1085-8.)

Table 2. Comparison of IQ at Seven Years between Children with Nonfebrile Seizures and Their Siblings.

COHORT	NO. OF CHILDREN	IQ AT SEVEN YEARS mean = SD	CASES WITH MENTAL RETARDATION	
			no.	%
Total				
Cases	98	91.5±20.9	14	14.3
Sibling controls	98	95.3±15.1	5	5.1
Early normal evaluation only				
Cases	59	99.8±14.9	1	1.7
Sibling controls	59	98.0±13.0	1	1.7

\*IQ<70 at seven years old

†P<0.01 for the difference between SDs for the cases and the siblings(Priman test for comparison of correlated variances).

P<0.05 for the difference in the rates of mental retardation (McNemar test for paired samples.)

The Title of the article asks a specific question.

- a. Does the last sentence of the abstract answer the question to your satisfaction ie. does it provide a definitive answer? (concentrate on numerical differences, not on biases, causation, effect of medication, etc.)
  
- b. If not, rewrite the last sentence of the abstract your way. (you might want to use the 1.8 IQ point difference from Table 2 but don't have to use exact numbers.)