0363-5465/97/2525-0414$02.00/0
THE AMERICAN JOURNAL OF SPORTS MEDICINE, Vol. 25, No. 3

Current Concepts

A Statistics Primer

Prevalence, Incidence, Relative Risks, and Odds Ratios: Some Epidemiologic Concepts in the Sports Medicine Literature

John E. Kuhn, MS, MD, Mary Lou V. H. Greenfield,* MPH, MS, and Edward M. Wojtys, MD

From MedSport and the Section of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan

Epidemiology is the study of disease frequencies and the causes of diseases in human populations. Terminology from this field of study is frequently encountered in the sports medicine literature, particularly with regard to causes of injuries to athletes and the prevention of such injuries. Suppose you are a team physician for a professional football team and you noted that a relatively high number of athletes had tears of the PCL. How common is this problem? Two epidemiologic measures may be helpful in understanding this problem.

MEASURES OF DISEASE FREQUENCY

Prevalence is the number of *existing* cases of a disease or injury divided by the total population at risk at a *given point in time.* For example, suppose an NFL team physician conducting preseason examinations of a team found 2 PCL tears in 100 athletes. In this example, the population is established at a specific point in time, and the prevalence of PCL tears for this team before the season starts is 2%.

Incidence, on the other hand, refers to the number of *new* cases that develop over a specified *period of time.* There are two ways to look at incidence: the cumulative incidence and the incidence rate. The *cumulative incidence* is the number of new cases of a disease during a given duration of time divided by the total cases without the disease at the start of the time period, but who are at risk. In the example of a football team with 100 players, 2 had PCL tears before the season started. At the end of the season the physician reexamined the team and found a total of 6 players with PCL tears. Among the 98 players who were healthy at the beginning of the season, 4 new PCL tears occurred during the season, and thus the cumulative incidence for the team during the season was 4 of 98 or 4.1%, while the prevalence at the beginning of the season was 2 of 100 or 2%, and at the end of the season the prevalence was 6 of 100 or 6% (Table 1).

Unfortunately, the cumulative incidence does not account for varying time periods that players may be exposed to injury and, therefore, may not be a precise method for estimating the injury rate. Continuing with the PCL example, suppose most injuries to the PCL occur during games. During the season, football players may be traded or activated from the injured reserve list, or they may be reserves and not play in games. These players would not have had the exposure to as many injury-producing events (games) as players who had played all season. The *incidence rate* provides a more precise estimate of the impact of exposure; it is the number of new cases that develop during a given time period divided by the total person-time of observation. The total person-time of observation is calculated by adding together all of the time of exposure that each player has endured; in this example, it is the amount of playing time for each player. The incidence rate not only includes the population at risk, but also accounts for each participant's time at risk.

For example, for a football team of 100 players, the risk of PCL injuries during game playing in the 98 players with intact ligaments at the start of the season is determined by documenting the total minutes of time played in a game for each player and whether a PCL injury occurred. The game time played for each player is added to obtain the total time at risk. To determine the incidence

* Address correspondence and reprint requests to Mary Lou V. H. Greenfield, MPH, MS, University of Michigan, Orthopaedic Surgery, TC2914G-0328, 1500 East Medical Center Drive, Ann Arbor, MI 48109.

No author or related institution has received any financial benefit from research in this study.

TABLE 1
Example of Prevalence, Incidence, and Relative Risk

Hypothetical team of 100 players	PCL injuries/ starters	PCL injuries/ reserves	Totals
Before season	2/50	0/50	2/100
After season	5/50	1/50	6/100

Before season prevalence = 2/100 or 2%
After season prevalence = 6/100 or 6%
Cumulative incidence for the entire team during the season = 4 new cases/98 healthy players, or 4.1%
Incidence for starters = 3 new cases/48 healthy players, or 6.25%
Incidence for reserves = 1 new case/50 healthy players, or 2%
Relative risk = Incidence of injury in exposed group/Incidence of injury in unexposed group = Incidence for starters/Incidence for reserves = 6.25%/2%
Relative risk = 3.125

rate, the number of new cases of PCL injury is divided by the total person-time of playing (Table 2).

MEASURES OF DISEASE ASSOCIATION

A relationship often exists between the exposure to a particular risk factor and the development of a disease or injury. Football players who play frequently are more likely to get injured. One way of expressing this relationship is the relative risk. The *relative risk* is an estimate of the magnitude of the association between the exposure to a risk factor and an injury, and indicates the likelihood of developing the injury in the exposed group relative to the group that has not been exposed. It is the incidence of the injury in the exposed group divided by the incidence in the unexposed group. In the football team example, an incidence of 4.1% for PCL tears was found during the last season (Table 1). To assess the risk that playing in football games has on the development of PCL tears, the incidence of PCL tears in an unexposed group (the reserve players who practice but do not play in games) could be compared with the incidence of PCL tears in the exposed group (those players who start or play in games). For example, suppose only 1 PCL tear was noted in 50 reserves during the course of the season, producing an incidence of 2% (Table 1). On the other hand, among the 48 healthy starting players, 3 new PCL tears developed during the season, producing an incidence for starters of 6.25%. Dividing the incidence of those exposed to the risk of playing games (6.25%) by the incidence of those not exposed to games (2%) gives us the relative risk of a football player injuring his PCL by playing a season of football games, which is 3.125.

A relative risk of 1.0 means that the incidence rate of a disease or injury in the exposed group is identical to the incidence rate of the unexposed group, indicating there is no association observed between the exposure to a risk factor and the development of a disease. When the relative risk is greater than 1.0, a positive association exists between the exposure to a risk factor and a disease. In this example, the football players who play in games have 3.125 times the risk of developing a PCL tear compared with reserves. If the relative risk is less than 1.0, an inverse association exists, such that exposure to a particular variable has a protective effect and decreases the risk of disease or injury to those exposed.

Cohort studies are particularly useful for estimating relative risk. In a cohort study, a group of participants who are disease- or injury-free is observed over time *prospectively* with exposures and the onset of the disease or injury documented as they occur. Many studies, however, are *case-control studies* in which the participants are selected on the basis of their disease status, and their histories are evaluated *retrospectively* to identify exposures and risks preceding the disease or injury. The case-control study, which by definition begins with a population of diseased or injured participants, cannot be evaluated prospectively to determine exposure risks, and as such the relative risk calculation cannot be used. Instead, the relative risk can be approximated by determining the odds ratio. The *odds ratio* is the odds of exposure in those with the disease or injury compared with the odds of exposure in those without the disease or injury. These probabilities can be reduced to a simple formula useful for calculating the odds ratio (Table 3).

For example, suppose the team physician is asked to evaluate 100 potential new players for the team. The top 100 potential new players fall into two major groups: there are 20 players with PCL injuries and 80 with intact ligaments (Table 3). Of the 20 players with PCL injuries, 15 were starters and played in games during their college seasons and 5 were reserves and did not play. Of the 80 players with intact PCLs, 50 were reserves and 30 were starters who played in games. The odds ratio is calculated and is found to be 5.0 (Table 3). Thus, the odds ratio takes retrospective data and determines the odds of the exposure, given that the player had the injury. In this case, the football player has a 5-fold greater chance of having an injured PCL if he played in games as opposed to being a reserve player.

TABLE 2
Example of Incidence Rate[a]

Player	Minutes played in games during current season	Did PCL tear develop during current season?
1	1200	No
2	800	Yes
3	1100	No
4	900	No
5	1400	No
6	600	No
7	1600	No
8	400	Yes
9	1000	No
10	1000	No
Totals	10,000	2 PCL injuries

Incidence rate = two new PCL injuries/10,000 player-minutes

[a] In this hypothetical example 10 new players are followed over a season and the number of minutes of game time played as well as the development of a PCL tear is assessed. The incidence rate of 2 PCL injuries per 10,000 minutes of playing time assesses the risk of sustaining a PCL injury during play.

TABLE 3
Example of Determining Odds Ratio

	Disease			PCL Tears	
	Yes	No		Yes	No
Exposure			College history		
Yes	a	b	Played in games	15	30
No	c	d	Reserve player	5	50

a = numbers of players who are exposed and have the condition or disease
b = number of players who are exposed and do not have the condition or disease
c = number of players who are not exposed and have the condition or disease
d = number of players who are not exposed and do not have the condition or disease

$$\text{Odds ratio} = \frac{\text{Probability of exposure among those with the disease}}{\text{Probability of exposure among those without the disease}}$$

$$= \frac{\text{P(exposed/diseased)/P(not exposed/diseased)}}{\text{P(exposed/not diseased)/P(not exposed/not diseased)}}$$

$$= \frac{(a/a+c)/(c/a+c)}{(b/b+d)/(d/b+d)}$$

$$\text{In the example: } \frac{(15/15+5)/(5/15+5)}{(30/30+50)/(50/30+50)} = 5.0$$

It is important to understand that relative risks and odds ratios should be accompanied by confidence intervals. The confidence interval provides us with an estimate of the degree of assurance as to the range within which the true relative risk or odds ratio lies. The confidence interval guides us in our interpretation of the study findings. Confidence intervals will be covered in more detail in a future "Current Concepts" article.

It also is important to understand that because there is an association between the risk factor being studied and the injury or disease that occurs, it is not necessarily true that the risk factor *causes* the injury or disease. It is possible that the association demonstrated by the relative risk or odds ratio may be due, totally or in part, to bias, confounding, or chance.

Measures of disease frequency such as prevalence and incidence and measures of association such as relative risk and odds ratios are basic tools that help us to quantify the relationship between exposures to risk factors and diseases or injuries. The data provided by these tools provide the foundation from which researchers can formulate testable hypotheses for further study.

ACKNOWLEDGMENT

The authors thank Dr. M. Anthony Schork from the University of Michigan, School of Public Health, Department of Biostatistics, for his review of this article.