CE ARTICLE

SECTION EDITOR: PAMELA V. MOORE, EdD, MPH, RN

Physical Fitness and Employee Absenteeism

MEASUREMENT CONSIDERATIONS FOR PROGRAMS

by Karen M. Conrad, PhD, MPH, RN and Carolyn L. Blue, MS, MSN, RN

The purpose of this article is to discuss the measurement issues that need to be considered by the occupational health nurse when designing and evaluating worksite physical fitness programs that have as an intended program outcome the reduction of employee absenteeism. We will also offer suggestions on the design and analysis methods that may be used in evaluating such programs.

ABOUT THE AUTHORS:

Dr. Conrad is Associate Professor, Department of Public Health, Mental Health, and Administrative Nursing, and Program Director, Occupational Health Nursing Program, University of Illinois at Chicago. Ms. Blue is a doctoral candidate, University of Illinois at Chicago, and Associate Professor, Community Health Division, School of Nursing, Purdue University, West Lafayette, IN.

In actuality, when a physical fitness program is ongoing, the occupational health nurse can use periodic outcome information about the program to tailor and adjust it so that program goals, e.g., improved physical fitness and reduced absenteeism, will more likely be met. When program measures are good—accurate, precise, and specific—there is an increased likelihood of observing the truth; that is, to validly conclude that a program is working when it really is and that a program is not working when it really is not (Conrad, 1994). Even the seemingly unwanted news that a program is not working can be used to turn the program around so that it becomes successful. Having available valid program outcome information based on sound measures is increasingly critical in today's work environment, where the competition for scarce resources demands that the occupational health nurse make informed program decisions.

BACKGROUND

For young children, physical activity is typically considered a form of play; however, by adulthood, most individuals give up a physically active lifestyle for a more sedentary one. For instance, only about 22% of adults engage in at least 30 minutes of activity five or more times per week (U.S. Department of Health and Human Services, 1991). However, staying physically active is good for one's health.

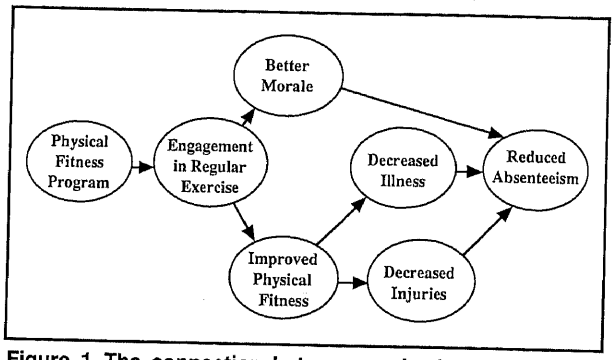

Figure 1. The connection between a physical fitness program and employee absenteeism.

Considerable evidence exists to suggest the positive effects of engagement in physical activity. Physical activity is associated with many health benefits, including improved cardiovascular and physical fitness, weight control, normalization of carbohydrate metabolism, psychological benefits, and improved quality of life in persons with chronic diseases (Bouchard, 1990; Fletcher, 1992; Gebhardt, 1990; International Society of Sport Psychology, 1992; Powell, 1989). Worksite research reveals that physically fit workers incur fewer injuries, have better morale and productivity, and report fewer absences than less physically fit workers (Hilyer, 1991; Leutzinger, 1991; Lynch, 1990; Rosenfeld, 1992).

Worksite physical fitness programs offer the opportunity to increase the level of physical fitness of workers. In fact, in listing health objectives for the nation, the document *Healthy People 2000* (U.S. Department of Health and Human Services, 1991) singles out, as one worksite objective, the increase in the proportion of worksites offering employer sponsored physical activity and fitness programs. In response to this national objective, the percent of worksites offering activities to promote exercise and physical fitness has increased from 22% in 1985 to 42% in 1992 (U.S. Department of Health and Human Services, 1993).

THE LINKAGE BETWEEN PHYSICAL FITNESS AND ABSENTEEISM

Reducing employee absenteeism is among the reasons worksites initiate health promotion programs (Shephard, 1992). Of those worksites that conduct formal evaluations of their programs, 36% of them collect data on absenteeism (United States Department of Health and Human Services, 1993). The occupational health nurse, as manager of a physical fitness program, may be responsible for such an evaluation.

More specifically, a number of researchers have examined the linkage between physical fitness and absenteeism. Shephard's (1992) review identified 39 published and non-published studies appearing between 1965 and 1990. A supplemental literature search by these authors located an additional seven studies from published sources. The above studies support the conclusion that the average reported impact of fitness programs on absenteeism is between 0.5 and 2.0 days' improvement in attendance per participant per year. This improvement translates to an estimated dollar saving of 0.35% to 1.4% of payroll (Shephard, 1992). Although the studies generally support a program effect, selection bias is a concern because of the voluntary nature of program participation and the tendency for those who participate to have a healthier lifestyle than those workers not engaged in physical fitness activities (Conrad, 1987).

Whether stated explicitly or not in the program plan, every physical fitness program is based on hypothesized linkages between the program and the intended outcome(s). The hypotheses serve to guide the selection of measures to be collected in the program. For example, it is logical to hypothesize the following about a physical fitness program: a) the program will increase engagement in regular exercise among the workers; b) participation in the program and engagement in exercise will result in improved physical fitness and job morale; c) increased physical fitness will decrease illness and work related injuries; and d) absenteeism will be reduced due to better morale, improved physical fitness, decreased illness, and decreased injury. Figure 1 is a graphic depiction of these hypothesized linkages.

MEASURING ABSTRACT CONCEPTS IN THE FIELD SETTING

It is generally recognized that biometric or physiologic measures need to be calibrated carefully and standard protocols followed. For example, when conducting blood pressure screenings, manometers are calibrated and readings are taken following a prescribed protocol. It is often *not* so obvious that serious attention also needs to be given to measuring more abstract concepts such as physical fitness and absenteeism.

In the field setting, finding out the truth—that is, whether or not a physical fitness program reduces employee absenteeism—is not easy. In observing relationships, there is not the same control as in the laboratory. There are many factors operating that make it difficult to know what causes what. Thus, in the world of physical fitness and absenteeism it is necessary to talk about *probability,* acknowledging that absolute certainty of results cannot be attained, no matter what the statistical values.

An important goal is to maximize the probability that results are accurate or valid.

Thus, dealing with error is inherent in evaluating worksite programs. For example, when conducting a physical fitness program in the workplace and examining its effect on employee absenteeism, there are a number of different scenarios that may occur: a) the program really worked and positive results were observed in the analysis (true positive result); b) the program really worked but the statistical results did not show this (false negative result). (This latter case happens more than one would hope); c) the program did not work and results show it did not work (true negative result); d) the program did not work, but because of chance factors a statistically significant result was obtained (false positive). Having available reliable and valid measures of physical activity, physical fitness, and absenteeism increases the likelihood of finding true results (true positive and true negative results).

RELIABILITY AND VALIDITY OF PHYSICAL FITNESS AND ABSENTEEISM

It is critical to preface any discussion of measurement issues with a few words about reliability and validity. Reliability is concerned with consistency or stability of the measuring technique, whether it be an instrument or a rater. Reliability refers to the extent to which a measure has few random error components and assesses a content domain consistently (Carmines, 1979; Nunnally, 1994; Waltz, 1991).

Interrater and intrarater reliability are especially important when physical activity or physical fitness is being observed or recorded. Likewise, if absenteeism data are being abstracted from the personnel record, it is important to know that the information is being recorded in a consistent manner over time and between raters. The occupational health nurse will want to know of any changes in company policy on how absenteeism is defined. To assess interrater reliability, two observers would record measurements on a small sample of individuals or records, as the case may be. Their results would then be inspected to determine the agreement between the two raters. Intrarater reliability is the consistency with which one rater or observer classifies individuals into levels or codes responses from a record or interview on two different occasions (Waltz, 1991).

An example of intrarater reliability is for one rater to observe an individual's routine physical activity at work and classify the activity using some predetermined classification scheme. At a later time, the same rater would again observe and classify the physical activity of this worker to determine the consistency of agreement in the two measures. With intrarater reliability, it is assumed that the physical activity of individuals is stable on the two occasions.

Validity of a measure is the extent to which it measures the intended construct (Nunnally, 1994). To determine validity, measures of physical activity are usually compared to measures of physical fitness, as the two concepts are closely related (for reviews, see Blair, 1984; LaPorte, 1985; Montoye, 1984). It is difficult to assess physical activity questionnaires for validity because of the lack of a "gold standard" measurement (Blair, 1984). Physical fitness, however, has valid and reliable measures (Blair, 1984; Heyward, 1991).

The validity of absenteeism measures has been debated for years. Of concern here is the need to be sure that the absenteeism measure selected is measuring the kind of absenteeism that is likely to be affected by a physical fitness program. (This issue will be discussed in more detail later.) The occupational health nurse will also want to ensure that the absenteeism records available are valid for all workers. For example, the accuracy of the absenteeism record for hourly workers may be greater than that for management level employees, where more flexibility may exist in what is viewed as an absence.

MEASUREMENT OF PHYSICAL FITNESS

Before beginning a discussion on the measurement issues involved with physical fitness, it is important to clarify three related terms: physical activity, exercise, and physical fitness (see Table 1). Physical fitness is the end product, and has been defined as the ability to perform and maintain occupational, recreational, and daily activities throughout life without becoming unduly fatigued (Heyward, 1991). Physical activity and exercise can have as their desired outcome improved physical fitness. Physical activity is broadly defined as "any bodily movement produced by skeletal muscles and resulting in energy expenditure" (Bouchard, 1990, Howley, 1992). The domains of physical activity include leisure, work, household chores and other activities, and exercise—the latter of which is defined as planned activity for the purpose of improving physical fitness (Blair, 1984; Cowles, 1992). Thus, exercise is a type of physical activity with physical fitness being influenced by exercise habits. Figure 2 illustrates the relationship among these concepts.

A worksite program may have as its goal improved physical fitness of workers. To meet this goal, the program may offer exercise classes as well as other forms of physical activity. Components of individual physical fitness include: aerobic fitness, including cardiovascular and pulmonary fitness; anaerobic fitness, including

TABLE 1

Measurement of Physical Activity and Physical Fitness*

Measures of Physical Activity	Measures of Physical Fitness
Type of Activity	Cardiorespiratory fitness
Leisure	Treadmill performance
Occupational	Cycle ergometer performance
Exercise	Bench step
Household and other	Jog/run test
Intensity of activity	Walk test
Low intensity	Body composition
Moderate intensity	Underwater weighing
High intensity	Skinfold measures
Frequency of activity	Body mass index
Duration of activity	Relative weight
	Energy costs
	VO_2 direct
	VO_2 indirect
	Muscular strength and endurance
	Bench press
	Grip dynamometers
	Flexibility
	Back strength
	Hamstring stretch
	Sit and reach test

*Examples

strength, power, and muscle endurance; flexibility; body composition; and psychological well being such as tension control (Bouchard, 1990; Dunn, 1987; International Society of Sport Psychology, 1992).

Planners of worksite exercise programs need to be selective as to which component to target. It is useful in planning a worksite physical fitness program to consider that worker health is affected not only by individual health behavior practices, but also by factors within the workplace, such as the job hazards, job tasks, and equipment. The job tasks of workers, including those engaged in physically demanding jobs, need to be quantified and considered.

For example, if the desired program outcome is increased flexibility of the back for a group of workers whose job tasks involve bending, lifting, and twisting, then an appropriate physical fitness measure is back flexibility. Other jobs that are sedentary at times and require bursts of energy and physical exertion at other times (fire fighting, for example) may focus on both aerobic and anaerobic fitness and flexibility. A job task analysis is therefore necessary to determine which component of fitness is to be addressed and measured. For this reason, the objectives of the program will determine the physical fitness measures used.

What Should Be Measured?

Generally, the goal of a worksite exercise program is to increase an individual's physical activity pattern (i.e., leisure, occupational, exercise, and other activities) which should result in improved physical fitness. It depends on the desired program outcomes as to whether all estimates of activity or whether only high intensity aerobic activities are to be measured (Blair, 1984; Washburn, 1986). A physical activity pattern is a comprehensive description of type, frequency, duration, and intensity over a specified period of time (Bouchard, 1990). The American College of Sports Medicine (ACSM) (1990, 1991) recommends 20 to 60 minutes of physical activity at an intensity of 60% to 90% of maximal heart rate, 3 to 5 days a week, to improve car-

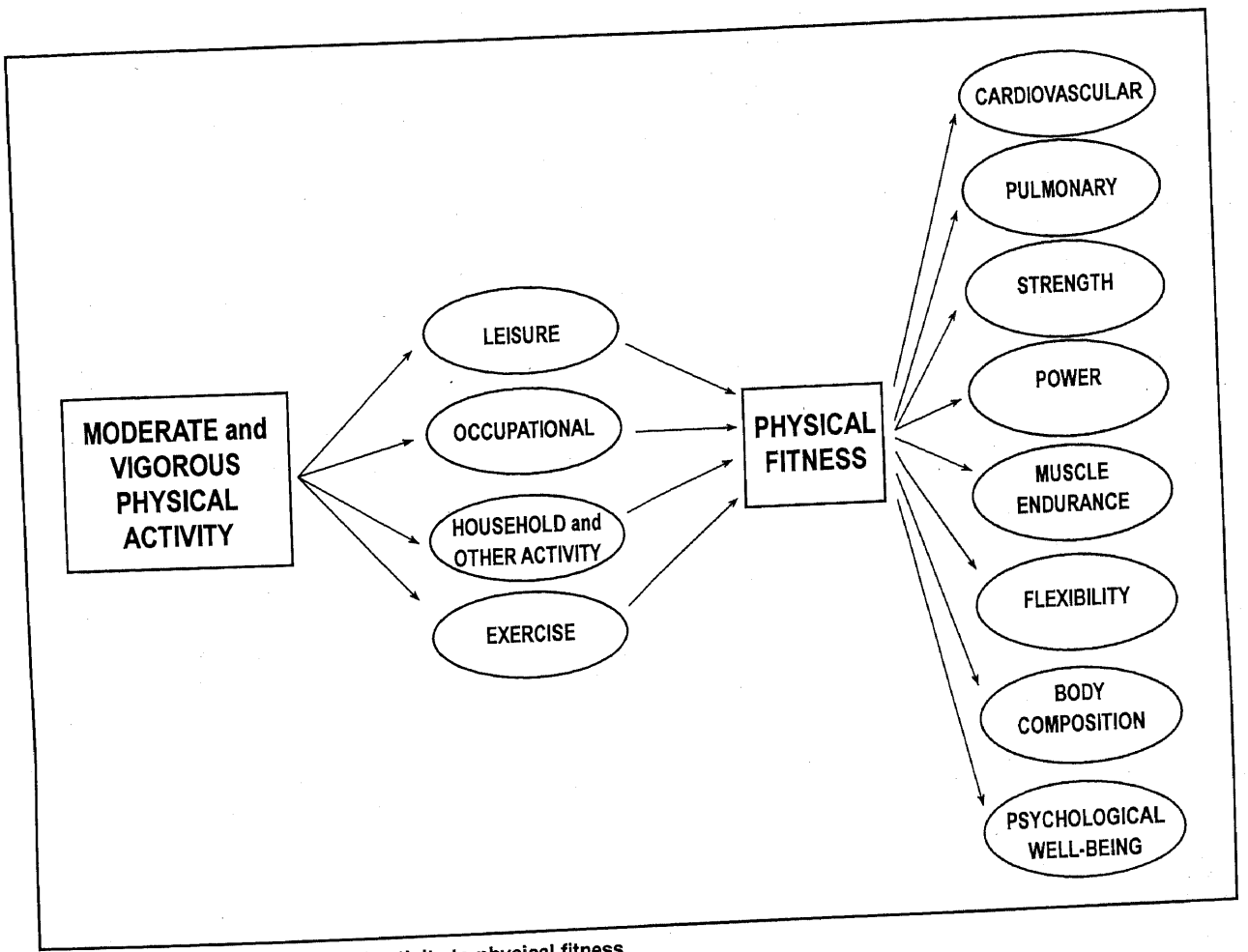

Figure 2. Relationship of physical activity to physical fitness.

diorespiratory fitness. Lower intensity activities should be performed for longer periods of time, while the duration of vigorous (i.e., moderate and high intensity) activities may be performed for shorter periods of time (ACSM, 1990).

Exercise can be performed at various intensities, depending on the current health of the worker and fitness goals—desired performance, health, or physical fitness (see Table 1). Howley (1992) described the various exercise intensities and benefits. Low intensity exercise is physical activity that produces little increase in respiration and no discomfort and is adequate for minimal health benefits (walking is an example). Moderate intensity causes mild breathlessness, some perspiration, and perhaps initial soreness. High-intensity exercise is activity at 80% to 120% of a person's maximal heart rate and is recommended for those interested in high level performance. Usually low or moderate intensity is adequate for worksite programs, but jobs requiring bursts of strength and power might benefit from high-intensity exercise.

Physical fitness (e.g., cardiovascular health or increased flexibility) is influenced by exercise patterns or habits of physical activity (Blair, 1984). Different patterns of physical activity are needed for different physical fitness goals. Thus, the measurement of physical activity should include type of activity, intensity, duration, and frequency (Haskell, 1985).

Measures of Physical Fitness

Physical fitness measures are frequently used to evaluate the impact of worksite exercise programs. Before any physical fitness measure is taken, it is important to screen persons who have potential risks and negative effects from exercise (ACSM, 1991). Protocol for pre-testing fitness testing for moderate and high intensity

exercise is outlined in *Guidelines for Exercise Testing and Prescription* (ACSM, 1991).

Components of physical fitness include cardiorespiratory fitness, body composition, muscular strength and endurance, and flexibility (Caspersen, 1985; Heyward, 1991) (see Table 1). Although all of the components can be improved through physical activity, cardiorespiratory endurance is improved most by aerobic exercise (Heyward, 1991).

Cardiorespiratory endurance (CRE) is defined as "the ability to perform large muscle, dynamic, moderate to high intensity exercise for prolonged periods" (ACSM, 1991). Although the direct measure of CRE is maximal oxygen consumption (VO_{2max}), in community studies, other measurements have been developed that correlate well with the direct measure (ACSM, 1991). Tests appropriate for use in the worksite are submaximal cycle ergometer and treadmill tests, step tests, distance runs, and walk tests (ACSM, 1991; Heyward, 1991). The object for all of these tests is to calculate the heart rate response to exercise.

Body composition is another way of measuring fitness. Body composition is the amount of lean body mass (LBM) in proportion to body fat (ACSM, 1991). Fat weight decreases and LBM increases in individuals who exercise due to increases in muscle tissue and blood volume (Astrand, 1986). Hydrostatic (underwater) weighing is the "gold standard" for measuring LBM. However, the technique requires special equipment and trained technicians, and is expensive, time consuming, and complicated (ACSM, 1991). Reasonably reliable and valid measures of body composition used more frequently in community and worksite studies include skinfold thickness, circumferences, and body mass index (BMI) (ACSM, 1991; Heyward, 1991; Howley, 1992).

Muscular strength and endurance are related in that improvement in one leads to improvement in the other. Strength of a muscle is the "maximum force that can be generated by a specific muscle or muscle group," while muscular endurance is the "ability of a muscle group to execute repeated contractions over a period of sufficient time to cause muscular fatigue" (ACSM, 1991). Muscle strength and endurance are improved when the force and repetition are increased beyond what the muscle normally encounters (Howley, 1992). Some common strength tests include single push-up, pull-up, sit-up, abdominal curl or extension, and grip test using a spring loaded instrument (dynamometer) (ACSM, 1991; Heyward, 1991). Muscular endurance is measured by having an individual do the maximum number of sit-ups, pull-ups, push-ups, or lift a predetermined weight for as many repetitions as possible (Heyward, 1991).

There are two important concerns in the measurement of muscle strength and endurance. The first is to identify which variable is to be measured (i.e., strength and/or endurance), and the second is to identify the muscle groups important to the program evaluation. For example, grip strength measures will not demonstrate a person's overall strength. For some forms of work, such as fire fighting and nursing, power or the rate of performing work (ACSM, 1991) also should be included in the measurement criteria. Power can be assessed by including a time factor or limitation to strength measures (ACSM, 1991).

Flexibility is the ability to move a joint or series of joints smoothly through the complete range of motion (ACSM, 1991). The direct measures of flexibility, expressed in degrees, have been more reliable and valid than indirect measures (ACSM, 1991; Heyward, 1991). Common devices for measuring joint range of motion direction include the goniometer, electrogoniometer, and flexometer (ACSM, 1991). Indirect measures of shoulder elevation, ankle flexibility, sit and reach and trunk extension tests are frequently employed in fitness evaluation (ACSM, 1991). However, the sit and reach test measure of flexibility was not found to be a valid measure of back flexibility (Jackson, 1986, 1989). Flexibility is joint specific; therefore, the flexibility test (i.e., measure) chosen for program evaluation should accurately reflect the desired outcome. For example, if the desired program outcome is increased hip flexibility, then the sit and reach test may be done using established test protocols (ACSM, 1991).

Reducing Physical Fitness Measurement Error

Measurement error can be reduced by following specific protocol and calibrating equipment used. Error in measures can result from inter-individual variability (age, gender, heredity, and race), situational contaminants, transitory personal factors, test administration variations, and precision of the measuring tool (Polit, 1995; Waltz, 1991). Situational factors such as room temperature, water vapor pressure gradient, and humidity affect work performance (Alpaugh, 1988; Howley, 1992). Transitory personal factors include pretest fatigue, hunger, mood, and illness. It is important that: a) the environment, the measurement tools, and the person be in optimal state prior to any physical fitness evaluation; b) the administration of the test be done using controls and standardization of the testing protocol; and c) measurement instruments be calibrated and in good working order to assure both reliability and validity.

Reliability depends on the stability, consistency, and dependability of equipment from one measurement time to another. When pulse, blood pressure readings, weight, or skinfolds are not taken with the same method each time, the results may not be reliable and the evaluation of the fitness program will be difficult to interpret. When instruments are not calibrated correctly, measures will be biased consistently in one direction (either too high or too low). For example, the step height, duration, and timing of taking the pulse when doing a step test are critical to the validity of the test. If step height is not adjusted for size or the pulse taken too soon or too long after the stepping is finished, the results will be biased in one direction, and the test may not demonstrate a positive change when there is one. Finally, when the test is an inappropriate measure (such as grip strength measures for lifting strength), results will not be valid. Therefore, it is incorrect to make generalizations about a person's general strength and/or endurance based on only one measurement (hand grip dynamometer, for example).

Selection of physical fitness measures requires careful consideration of desired program results, instrument reliability, validity, convenience, and cost. Because the desired result of an exercise program is improved physical fitness, fitness measures are appropriate for evaluation purposes. To achieve the best possible evaluation requires efforts to reduce measurement error. Otherwise, the program may be effecting change, but the evaluation results will not demonstrate change.

MEASUREMENT OF EMPLOYEE ABSENTEEISM

The measurement of absenteeism is not straightforward (Baun, 1995; Markham, 1983). When measuring absenteeism for research purposes, three different aspects of absenteeism need to be considered: the conceptual definition, the type of absenteeism, and the measure of absenteeism (see Table 2). A simple conceptual definition of absenteeism offered by Brooke (1986) stated that absenteeism is non-attendance for scheduled work. A slightly more specific definition used by the National Center for Health Statistics stated that absenteeism is work loss days on which a person did not work at his or her job or business for at least half of the normal workday because of specific illness or injury (Porwoll, 1980). In these definitions, absenteeism is represented by a binary variable (absent/not absent). Absenteeism may be conceptually defined in more elaborate ways. The reader is referred to Hulin (1994) for further discussion.

In addition to deciding on the conceptual definition of absenteeism, the occupational health nurse needs to consider the absenteeism types (or reasons for absence) to be recorded. Absenteeism types may be categorized as voluntary or involuntary. Voluntary absenteeism, for example, may include maternity leave, family leave, or even vacation time. Involuntary absenteeism includes absences due to illness or injury and may also include absences due to jury duty or funeral leave. It is involuntary absenteeism that is of concern as an outcome measure in worksite physical fitness programs.

TABLE 2
Measures of Absenteeism*

- Frequency—Number of absence episodes in a given time period
- Time Lost—Total time lost in a given time period
- Duration—Average length of time of absences in a given time period

*Examples

Obviously, it is important in the planning of the program to decide which type(s) of absenteeism to collect by considering the logical connection between the program and outcomes. If the program is hypothesized to affect only illness and work related injury, then data on these two kinds of absences need to be collected. Including data on maternity leave absences not under the control of the fitness program may lead to misinformed results about program effectiveness. Akin (1984) logically posited that each type of absenteeism may be the outcome of a different set of underlying causes, and each type may be represented by a different model.

There is no short supply of absenteeism measures. Muchinsky's (1977) classic work noted over 41 different measures. The absenteeism measure selected will be affected by available company data. Three of the more frequently used measures of absenteeism include: time lost, frequency, and average duration (Conrad, 1990). Time lost is the most frequently used measure in studies of physical fitness programs because of its logical linkage with health and its availability. Time lost refers to the total time absent (hours or days) within a specified period of time. Frequency refers to the number of episodes of absence during a specified period of time. The frequency measure has received research attention because it is considered to indicate volition and is considered as a measure

TABLE 3
Types of Research Designs*

One Group Pretest Posttest Design

O×O

Randomized True Experiment

$$\frac{O \times O}{O \quad O}$$

Quasi-Experimental Design

$$\frac{O \times O}{O \quad O}$$

*Examples

O=observation, X=intervention

of morale (Farrell, 1988). Duration refers to the average length of each absence during a specified period of time and is computed by dividing total time lost by frequency.

Once the occupational health nurse makes the forgoing decisions, it is time to consider the desired time period for which absenteeism data are to be collected. Ideally, absenteeism data should be collected prior to the start of the program as well as during and after the program. Collecting pretest absenteeism data provides an opportunity to observe absenteeism trends over time, to compare absenteeism rates before and after program inception, and to compare pretest absenteeism differences between participants and non-participants.

For how long should absenteeism data be collected? Although a few guidelines exist, there does not appear to be an aggregation period that is best for all situations. One factor to consider is that absenteeism has a relatively low base rate. Some workers will be absent less, and a few will be absent a great deal more than the average. In general, a longer period of observation time provides a more reliable or stable picture of absenteeism patterns not affected by short, cyclical variations such as seasons. If cyclical variations exist at a worksite, then it is also important to consider that the pretest and posttest absenteeism periods are parallel points in time. For example, if absenteeism was measured for a short period of time, such as 3 months of pretest in the fall followed by 3 months of posttest in the winter, then a rash of absenteeism in the winter due to the flu could make the program appear ineffective. In actuality, the flu season could have obscured a true program effect. Fitness program research often reports using 1 or more years of pretest and posttest absenteeism data (Blair, 1986; Lynch, 1990; Shephard, 1992).

A few employees may have extensive absenteeism within the data reporting period. Such cases are termed "outliers." A few extreme cases of extensive absenteeism can greatly affect the group mean. In the following hypothetical example, the number of days of posttest absenteeism for a group was generally in the range of 5 days per year. However, someone in the group was hospitalized for a catastrophic illness and was absent for 200 days in that same year. Especially for a small size group, this one high absenteeism figure will shift the group mean up considerably. If this case is in the program group at the posttest, the program may look as if it had no effect. On the other hand, if this case ends up in the control/comparison group at posttest, then the program may look like a success, not because it really was, but because the outlier increased the control/comparison group absenteeism mean considerably. Texts on statistical methods are available on how to assess and treat outliers.

In summary, with regard to the measurement of absenteeism, the occupational health nurse needs to be clear about how absenteeism is to be conceptually defined, what types of absenteeism should be recorded, and what measure(s) need to be recorded. Also, a decision needs to be made about how absenteeism will be recorded—in hours or days and for what period of time. Finally, outliers need to be inspected and a decision made on how to handle them in a meaningful way.

PROGRAM EVALUATION DESIGNS

The research literature reports using several different research designs for examining the effect of a worksite physical fitness program on employee absenteeism. Some designs are more elaborate than others and some support causal inferences better than others. This section presents three of the more frequently used designs for examining the effect of a fitness program on employee absenteeism (see Table 3 for graphic depictions of research design examples).

One such design is called the one group pretest-posttest design. With this design, one group of employees is studied at two points in time: before the program is initiated and again later on in the program or at the end of the program. Using this approach, the occupational health nurse may record a group mean pretest absenteeism score. In addition, a posttest mean score of total days absent for the year after the program would also be recorded. With this design, as well as with those

described below, the nurse investigator observes for changes in absenteeism that may be due to the program. Ideally, the rate of absenteeism will decrease between Time 1 and Time 2. The problem with this design is that even if absenteeism does decrease, the investigator can rarely conclude that the decreased absenteeism was due to the program, because many other factors may account for the observed change. For example, if a new, stricter absenteeism policy was instituted at the same time as the program, it is possible that absenteeism reduction noted may be due to the new policy, rather than the program.

A commonly used quasi-experimental design is pretest/posttest nonequivalent control group design (Cook, 1979). In this design, like the one described above, there is no random assignment of employees to groups. Instead, the pretest (i.e., baseline absenteeism rate) is examined to identify initial differences between participants and non-participants. Pretest differences are assumed to exist. For example, compared to non-participants, those employees who join fitness programs tend to be younger, more educated, and have higher status jobs. Attempts must be made to control for these baseline differences either statistically or logically. The non-equivalence between participants and non-participants is referred to as selection bias and is one of the major reasons why it is difficult to attribute desired outcomes, such as aerobic fitness, to a particular worksite program (Conrad, 1991).

The most rigorous research design is the true experiment. The pretest/posttest control group design is one of the more popular true experiments. Like the above designs, in this design there is some comparison from which change can be inferred. What is unique about this design is the random assignment of employees to groups. Employees would be assigned randomly to either participate in the fitness program or not participate. Absenteeism data would be recorded for both groups. When random assignment is possible, it is considered the preferred approach because it can control for many threats to internal validity; that is, the ability to infer that A (the program) caused B (the outcome, e.g., physical fitness and/or reduced absenteeism). However, true experiments do not automatically control such internal validity threats as local history, attrition, diffusion of treatment, or compensatory rivalry, nor do they control threats to construct validity, external validity, or statistical conclusion validity (Conrad, 1994; Cook, 1979).

True experiments are often dismissed as infeasible for worksite health promotions programs because of the employer's reluctance to refuse a program to employees who want it. Sometimes control group members can be placed on a waiting list and admitted to the program at a later date. However, with programs such as physical fitness, the nurse needs to consider that if employees are refused admission to a worksite fitness program, they may join a non-worksite program. In this case, the control group also participates in a program, and is not truly a control group. The result of such a situation is that few differences may be noted between participants and non-participants. Thus, the program may appear to be ineffective, not because it was, but because the control group participated in a program of its own. For this reason, exercise information also should be collected on the comparison group.

Regardless of what design is used, it is important to collect exercise data on the comparison group, if one is available. It may be faulty to assume that employees do not exercise because they are not in the fitness program. Too often, exercise data are not obtained for the control/comparison group. If both the treatment group and the control/comparison group exercised, concluding that a worksite fitness program did not have an effect on absenteeism can be misleading.

ANALYSIS APPROACHES

Multiple factors are involved in deciding what analysis approach to use in a particular study, such as the program objectives, the design of the study, the number of variables, the level of measurement of the variables, and statistical, computer, and financial resources. This section highlights two of the more common approaches. The reader is referred to statistic texts for further detail. Before conducting a statistical analysis, the occupational health nurse investigator needs to be aware of the statistical assumptions of the test.

If the objective of the fitness program is to examine the effect of the program on absenteeism or to examine differences between participants and non-participants on some variable, such as absenteeism, then a quantitative approach, such as *t* tests, analysis of variance (ANOVA) or covariance (ANCOVA), is appropriate.

The *t* test is used when examining pretest-posttest differences across one group (e.g., the dependent *t* test for use in the one group pretest/posttest design) or when examining differences between participant and non-participant groups at the pretest and/or posttest (i.e., the independent *t* test).

A statistically stronger test, analysis of covariance (ANCOVA), is especially useful when examining multiple groups and pretest differences are assumed to exist between the groups. For example, in a quasi-experimental design where non-equivalence between groups is assumed (e.g., the baseline absenteeism rate is different

between the program participants and non-participants), ANCOVA can help statistically control for the baseline difference. Controlling for the baseline differences strengthens one's ability to infer a program effect.

CRITICAL ISSUES TO CONSIDER

In reality, a physical fitness program may or may not show a statistically significant effect on employee absenteeism. This paper focused on one of those reasons—poor and/or inappropriate measures. There are other reasons as well for findings that do not support the study hypotheses (Cook, 1979). For example, a program may actually have reduced absenteeism, but the effect is not discernible because its impact was too small or the sample size was too small. On the other hand, in some cases a program may be judged to have an effect on absenteeism when in reality that conclusion is faulty. Why? As previously mentioned, when dealing with statistical procedures, there is always some degree of uncertainty and there will always be a margin of error in the results. While statistical results and "*p* values" can be impressive to the on-looker, statistics cannot replace the need for sound judgment. Whether or not a statistically significant test result is obtained, it is important for the occupational health nurse to consider the clinical or substantive significance of the results. For example, how much difference in absenteeism rates between the intervention and control groups is needed before deciding that the program is effective in reducing absenteeism? Do improvements in physical fitness test scores translate into meaningful improvements in worker functional capacity?

In evaluating the goal that a physical fitness program reduces absenteeism, careful selection of measures is imperative. In addition, the occupational health nurse must assure that the measures are, in fact, capturing the aspects of physical fitness and absenteeism of interest.

REFERENCES

Akin, R.S., & Goodman, P.S. (1984). Methods of defining and measuring absenteeism. In P.S. Goodman and R.S. Atkin (Eds.). *Absenteeism* (pp. 47-109). San Francisco: Jossey Bass.

Alpaugh, E.L. (1988). Temperature extremes. In B.A. Plog (Ed.). *Fundamentals of industrial hygiene* (3rd Ed.). Chicago: National Safety Council.

American College of Sports Medicine. (1990). Position stand: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness in healthy adults. *Medicine and Science in Sports and Exercise, 22,* 265-274.

American College of Sports Medicine (1991). *Guidelines for exercise testing and prescription* (4th ed.). Philadelphia: Lea & Febiger.

Astrand, P., & Rodahl, K. (1986). *Textbook of work physiology* (3rd ed.). New York: McGraw-Hill.

Baun, W.B. (1995). The impact of worksite health promotion programs on absenteeism. In R.L. Kaman (Ed.). *Worksite health promotion economics: Consensus and analysis* (pp. 131-145). Champaign, IL: Human Kinetics.

Blair, S.N. (1984). How to assess exercise habits and physical fitness. In J.D. Matarazzo, S.M. Weiss, J.A. Herd, & N.E. Miller (Eds.). *Behavioral health: A handbook of health enhancement and disease prevention.* New York: John Wiley & Sons.

Blair, S.N., Smith, M., Collingwood, T.R., Reynolds, R., Prentice, M.C., & Sterling, C.L. (1986). Health promotion of educators: Impact on absenteeism. *Preventive Medicine, 15,* 166-175.

Bouchard, C., Shephard, R.J., Stephens, T., Sutton, J.R., & McPherson, B.D. (Eds.). (1990). *Exercise, fitness, and health: A consensus of current knowledge.* Champaign, IL: Human Kinetics Books.

Brooke, P.P. (1986). Beyond the Steers and Rhodes model of employee absenteeism. *Academy of Management Review, 11,* 345-361.

Carmines, E.G., & Zeller, R.A. (1979). *Reliability and validity assessment.* Newbury Park, CA: Sage Publications.

Caspersen, C.J., Powell, K.E., & Christenson, G.M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports, 100,* 126-131.

Conrad, K.J., & Conrad, K.M. (1994). Reassessing validity threats in experiments: Focus on construct validity. In K.J. Conrad (Ed.). *Critically evaluating the role of experiments. New Directions for Program Evaluation Series, 63,* 5-25.

Conrad, K.M., Conrad, K.J., & Walcott-McQuigg, J. (1991). Threats to internal validity in worksite health promotion program research: Common problems and possible solutions. *American Journal of Health Promotion, 6,* 112-122.

Conrad, K.M., Riedel, J.E., & Gibbs, J.O. (1990).The effect of worksite health promotion programs on employee absenteeism: A comparative analysis. *AAOHN Journal, 38,(12)* 573-580.

Conrad, P. (1987). Who comes to work-site wellness programs? A preliminary review. *Journal of Occupational Medicine, 29,* 317-320.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin Co.

Cowles, M. (1992). *The assessment of physical activity by self report—A review and critical appraisal.* (Report to the Ministry of Tourism and Recreation Sports and Fitness Branch). Ontario, Canada.

Dunn, M.M. (1987). Guidelines for an effective personal fitness prescription. *Nurse Practitioner, 12,* 9-26.

Farrell, D., & Stamm, C.L. (1988). Meta-analysis of the correlates of employee absences. *Human Relations, 41,* 211-227.

Fletcher, G.F., Blair, S.N., Blumenthal, J., Caspersen, C., Chaitman, B., Epstein, S., Falls, H., Froelicher, E.S.S., Froelicher, V.F., & Pina, I.L. (1992). Benefits and recommendations for physical activity programs for all Americans: A statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. *Circulation, 86,* 340-344.

Gebhardt, D.L., & Crump, C.E. (1990). Employee fitness and wellness programs in the workplace. *American Psychologist, 45,* 262-272.

Haskell, W.L., Montoye, H.J., & Orenstein, D. (1985). Physical activity and exercise to achieve health-related physical fitness components. *Public Health Reports, 100,* 202-211.

Heyward, V.H. (1991). *Advanced fitness assessment & exercise prescription* (2nd ed.). Champaign, IL: Human Kinetics Books.

Hilyer, J.C., Brown, K.C., Sirles, A.T., & Peoples, L. (1991). A flexibility intervention to reduce the incidence and severity of joint injuries among municipal firefighters. *Journal of Occupational Medicine, 32,* 631-637.

Howley, E.T., & Franks, B.D. (1992). *Health fitness: Instructor's handbook* (2nd ed.). Champaign, IL: Human Kinetics Books.

Hulin, C.L. (1994). Suggested directions for defining, measuring, and controlling absenteeism. In P.S. Goodman and R.S. Atkin (Eds.).

Absenteeism (pp. 391-421). San Francisco: Jossey Bass.

International Society of Sports Psychology. (1992). Physical activity and psychological benefits: A position statement. *International Journal of Sports Psychology, 23,* 86-91.

Jackson, A.W., & Baker, A.A. (1986). The relationship of the sit-and-reach test to criterion measures of hamstring and back flexibility in young males. *Research Quarterly for Exercise and Sport, 57,* 183-186.

Jackson, A.W., & Langford, N.J. (1989). The criterion-related validity of the sit-and-reach test: Replication and extension of previous findings. *Research Quarterly for Exercise and Sport, 60,* 384-387.

LaPorte, R.E., Montoye, H.J., & Caspersen, C.J. (1985). Assessment of physical activity in epidemiologic research: Problems and prospects. *Public Health Reports, 100,* 131-146.

Leutzinger, J., & Blanke, D. (1991). The effect of a corporate fitness program on perceived worker productivity. *Health Values, 15,* 20-29.

Lynch, W.D., Golaszewski, T.J., Clearie, A.F., Snow, D., & Vickery, D.M. (1990). Impact of a facility-based corporate fitness program on the number of absences from work due to illness. *Journal of Occupational Medicine, 32,* 9-12.

Markham, S.E., Dansereau, F., & Alutto, J.A. (1983). Absenteeism rates as measures in organizational research: Hidden cyclical and structural variation. *Review of Business and Economic Research, 18,* 21-31.

Montoye, H.J., & Taylor, H.L. (1984). Measurement of physical activity in population studies: A review. *Human Biology, 56,* 195-216.

Muchinsky, P.M. (1977). Employee absenteeism: A review of recent literature. *Journal of Vocational Research, 10,* 316-340.

Nunnally, J.C., & Bernstein I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Polit, D., & Hungler, B. (1995). *Nursing research: Principles and methods* (5th ed.). Philadelphia: J.B. Lippincott Company.

Porwoll, P.J. (1980). *Employee absenteeism: A summary of research.* Arlington, VA: Educational Research Services.

Powell, K.E., Caspersen, C.J., Koplan, J.P., & Ford, E.S. (1989). Physical activity and chronic disease. *American Journal of Clinical Nutrition, 49,* 999-1006.

Rosenfeld, O., & Tenenbaum, G. (1992). Physical fitness in the industrial environment: Perceived physical well-being benefits. *International Journal of Sport Psychology, 23,* 227-242.

Shephard, R.J. (1992). A critical analysis of work-site fitness programs and their postulated economic benefits. *Medicine and Science in Sports and Exercise, 24,* 354-370.

U.S. Department of Health and Human Services. (1991). *Healthy people 2000: National health promotion and disease prevention objectives* (DHHS Publication No. 91-50212). Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Health and Human Services (1993). 1992 national survey of worksite health promotion activities: Summary. *American Journal of Health Promotion, 7,* 452-464.

Waltz, C.F., Strickland, O.L., & Lenz, E.R. (1991). *Measurement in nursing research* (2nd ed.). Philadelphia: F.A. Davis.

Washburn, R.A., & Montoye, H.J. (1986). The assessment of physical activity by questionnaire. *American Journal of Epidemiology, 123,* 563-576.

Physical Fitness and Employee Absenteeism

This issue of the AAOHN JOURNAL contains a Continuing Education Module on "Physical Fitness and Employee Absenteeism," and has been approved by AAOHN for 1 contact hour of continuing education credit upon successful completion of the posttest and evaluation.

A certificate will be awarded and the scored test will be returned when the following requirements are met by the participant: The completed answer sheet is received at AAOHN on or before October 31, 1996; (2) A score of 70% (7 correct answers) is achieved by the participant; (3) The answer sheet is accompanied by a $10.00 processing fee. Expect up to 6 weeks for delivery of the certificate.

Upon completion of this lesson, the occupational health nurse will be able to:

1. Describe the relationship between physical fitness and absenteeism.
2. List valid and reliable outcome measures for physical fitness and absenteeism.
3. Identify appropriate research designs and analytic techniques to determine the effect of physical fitness programs on absenteeism.

AAOHN is accredited as a provider of continuing education in nursing by the American Nurses Credentialing Center's Commission on Accreditation. Alabama provider number ABNP0063. California provider number CEP9283. Louisiana provider number LSBN3.

Contact hour credits received for successful completion of the posttest and evaluation may be used for relicensure, certification, or re-certification.

Directions: Circle the letter of the best answer on the answer sheet provided. (Note: you may submit a photocopy for processing.)

1. According to the U.S. Department of Health and Human Services (1993), the percentage of worksites that offer physical activity and fitness programs is:
A. 28%.
B. 36%.
C. 42%.
D. 51%.

2. The research literature suggests that the average reported effect of fitness programs on absenteeism ranges from _____ days' improvement in attendance per participant.
A. 0.1 to 0.4.
B. 0.5 to 2.0.
C. 2.3 to 3.1.
D. 3.5 to 4.1.

3. In conducting a study on the effects of a physical fitness program on absenteeism, two occupational health nurses will record absenteeism data from personnel records. Determining the degree to which the two nurses agree when recording measurements is known as:
A. Interrater reliability.
B. Intrarater reliability.
C. Internal validity.
D. External validity.

4. The definition of "any bodily movement produced by skeletal muscles and resulting in energy expenditure" refers to:
A. Physical fitness.
B. Exercise.
C. Aerobic fitness.
D. Physical activity.

5. To improve cardiorespiratory fitness, the American College of Sports Medicine (ACSM, 1990, 1991) recommends how much physical activity 3 to 5 days per week?
A. 10 to 20 minutes at 95% of maximal heart rate.
B. 20 to 30 minutes at 50% of maximal heart rate.
C. 20 to 60 minutes at 60% to 90% of maximal heart rate.
D. 65 to 80 minutes at 40% to 50% of maximal heart rate.

6. Which of the following components of physical fitness is improved most by aerobic exercise?
A. Body composition.
B. Cardiorespiratory endurance.
C. Muscular strength and endurance.
D. Flexibility.

7. In planning a fitness study, the occupational health nurse recognizes that the most valid measurement for lean body mass is:
A. Hydrostatic weighing.
B. Skinfold thickness.
C. Circumferences.
D. Body mass index.

8. The most frequently used absenteeism measure in studies of physical fitness programs is:
A. Time lost.
B. Frequency.
C. Average duration.
D. Voluntary absence rate.

9. Which of the following research designs is the most rigorous in examining the effect of a worksite physical fitness program on employee absenteeism?
A. Descriptive correlational.
B. One group pretest/posttest design.
C. Pretest/posttest control group design.
D. Pretest/posttest non-equivalent control group design.

10. A study of the effect of a worksite fitness program on absenteeism included program participants and non-participants. Preliminary analysis shows that the baseline absenteeism rates for these groups are different. Which of the following statistical tests will control for this baseline difference in absenteeism?
A. Dependent *t* test.
B. Independent *t* test.
C. Analysis of variance (ANOVA).
D. Analysis of covariance (ANCOVA).

ANSWER SHEET

Continuing Education Module

Physical Fitness and Employee Absenteeism

November 1995

Mark one answer only!
(You may submit a photocopy of the answer sheet for processing.)

1.	A	B	C	D	**6.**	A	B	C	D
2.	A	B	C	D	**7.**	A	B	C	D
3.	A	B	C	D	**8.**	A	B	C	D
4.	A	B	C	D	**9.**	A	B	C	D
5.	A	B	C	D	**10.**	A	B	C	D

EVALUATION (must be completed to obtain credit)

Please use the scale below to evaluate this continuing education module.

	4 - To a great extent	3 - To some extent	2 - To little extent	1 - To no extent
1. As a result of completing this offering, I am able to:				
A. Describe the relationship between physical fitness and absenteeism.	4	3	2	1
B. List valid and reliable outcome measures for physical fitness and absenteeism.	4	3	2	1
C. Identify appropriate research designs and analytic techniques to determine the effect of physical fitness programs on absenteeism.	4	3	2	1
2. The offering content was relevant to the objectives.	4	3	2	1
3. Independent study was an effective teaching method for the content.				
4. How much time (in minutes) was required to read this offering and take the test?	50	60	70	80

Please print or type: (this information will be used to prepare your certificate of completion for the module).
DEADLINE: October 31, 1996

NAME _____
ADDRESS _____ *STATE* _____ *ZIP* _____
CITY _____
PHONE _____
*LICENSE NUMBER** _____

Enclose check or money order for $10.00 payable to AAOHN in U.S. Funds and mail to:

Professional Affairs—CE Module
AAOHN
50 Lenox Pointe
Atlanta, Georgia
30324-3176