Executive Certificate in Data Analysis for Occupational Health and Safety Professionals

Statistical Methods for Occupational Health

Absolute Risk: Absolute risk refers to the probability of an event occurring, and it is often used to describe the risk of developing a disease or condition. In the context of occupational health, absolute risk is used to quantify the likelihood of a worker developing a work-related illness or injury. For example, the absolute risk of developing lung cancer from exposure to asbestos in the workplace can be estimated based on the level and duration of exposure.

Acceptable Risk: Acceptable risk refers to the level of risk that is deemed tolerable by an individual or organization. In occupational health, acceptable risk is often determined by regulatory agencies and is used to establish permissible exposure limits for hazardous substances. For instance, the Occupational Safety and Health Administration (OSHA) sets acceptable risk levels for workers exposed to toxic chemicals in the workplace.

Accident: An accident refers to an unexpected event that results in injury or damage. In occupational health, accidents can be caused by a variety of factors, including equipment malfunction, human error, and environmental hazards. For example, a worker may be involved in an accident while operating a machine, resulting in physical injury.

Accident Investigation: Accident investigation involves examining the circumstances surrounding an accident to determine its causes and identify measures to prevent similar accidents from occurring in the future. In occupational health, accident investigation is a critical process for reducing workplace injuries and illnesses.

Accuracy: Accuracy refers to the degree of closeness between a measurement or estimate and the true value. In occupational health, accuracy is important for ensuring that measurements of worker exposure to hazardous substances are reliable and valid. For instance, accurate measurements of air quality are necessary to determine whether workers are exposed to hazardous levels of pollutants.

Action Level: An action level refers to the concentration of a substance at which action is required to reduce worker exposure. In occupational health, action levels are often established by regulatory agencies to protect workers from hazardous substances. For example, OSHA sets action levels for lead exposure in the workplace, requiring employers to take steps to reduce exposure when levels exceed a certain threshold.

Additive Model: An additive model is a statistical model that assumes the effect of two or more variables is the sum of their individual effects. In occupational health, additive models are used to examine the combined effects of multiple risk factors on worker health. For instance, an additive model might be used to estimate the combined effects of smoking and asbestos exposure on lung cancer risk.

Adjusted Rate: An adjusted rate refers to a rate that has been modified to account for the effects of one or more variables. In occupational health, adjusted rates are used to compare rates of injury or illness across

different groups of workers while controlling for factors such as age, sex, and occupation. For example, an adjusted rate of injury might be calculated to compare the injury rates of male and female workers while controlling for differences in job duties and work experience.

Age-Standardized Rate: An age-standardized rate refers to a rate that has been adjusted to account for the effects of age. In occupational health, age-standardized rates are used to compare rates of injury or illness across different populations while controlling for differences in age distribution. For instance, an age-standardized rate of mortality might be calculated to compare the mortality rates of workers in different industries while controlling for differences in age.

Algorithm: An algorithm refers to a set of rules or procedures used to solve a problem or make a decision. In occupational health, algorithms are used to analyze data and make predictions about worker health risks. For example, an algorithm might be used to identify workers at high risk of developing a work-related illness based on their medical history and job duties.

Alternative Hypothesis: An alternative hypothesis refers to a statement that is contrary to the null hypothesis. In occupational health, alternative hypotheses are used to test the effectiveness of interventions or to examine the relationship between risk factors and health outcomes. For instance, an alternative hypothesis might be used to test the effectiveness of a new safety training program in reducing workplace injuries.

Analysis of Variance: Analysis of variance (ANOVA) refers to a statistical technique used to compare means across two or more groups. In occupational health, ANOVA is used to examine the effects of different risk factors on worker health. For example, ANOVA might be used to compare the mean levels of noise exposure among workers in different industries.

Annual Rate: An annual rate refers to a rate that is calculated over a one-year period. In occupational health, annual rates are used to describe the frequency of injuries or illnesses over time. For instance, an annual rate of injury might be calculated to describe the number of injuries that occur per year in a particular industry.

Antecedent: An antecedent refers to a variable that precedes another variable in time. In occupational health, antecedents are used to examine the causes of workplace injuries or illnesses. For example, an antecedent might be used to examine the relationship between worker fatigue and workplace accidents.

Association: Association refers to a relationship between two or more variables. In occupational health, associations are used to examine the relationship between risk factors and health outcomes. For instance, an association might be observed between worker exposure to asbestos and the development of lung cancer.

Asymptotic: Asymptotic refers to a value that is approached as the sample size increases. In occupational health, asymptotic values are used to describe the behavior of statistical estimates as the sample size becomes very large. For example, an asymptotic value might be used to describe the behavior of a confidence interval as the sample size increases.

Attack Rate: An attack rate refers to the proportion of individuals who develop a disease or condition over a specified period. In occupational health, attack rates are used to describe the incidence of work-related illnesses. For instance, an attack rate might be calculated to describe the proportion of workers who develop a respiratory illness after exposure to a hazardous substance.

Attributable Fraction: An attributable fraction refers to the proportion of cases that can be attributed to a specific risk factor. In occupational health, attributable fractions are used to examine the contribution of different risk factors to worker health outcomes. For example, an attributable fraction might be calculated to describe the proportion of lung cancer cases that can be attributed to worker exposure to asbestos.

Average: An average refers to a measure of central tendency, such as the mean or median. In occupational health, averages are used to describe the typical values of worker exposure to hazardous substances or the frequency of workplace injuries. For instance, an average might be calculated to describe the mean level of noise exposure among workers in a particular industry.

Bayesian Method: A Bayesian method refers to a statistical approach that uses Bayes' theorem to update probabilities based on new data. In occupational health, Bayesian methods are used to analyze data and make predictions about worker health risks. For example, a Bayesian method might be used to estimate the probability of a worker developing a work-related illness based on their medical history and job duties.

Benchmark: A benchmark refers to a standard or reference point against which performance is measured. In occupational health, benchmarks are used to evaluate the effectiveness of safety programs or to compare the performance of different industries. For instance, a benchmark might be established for the rate of workplace injuries in a particular industry, and employers might be encouraged to strive to meet or exceed this standard.

Bias: Bias refers to a systematic error or distortion in a measurement or estimate. In occupational health, bias can occur in the collection or analysis of data, and it can affect the validity of conclusions drawn from the data. For example, bias might occur in the selection of participants for a study, resulting in a sample that is not representative of the larger population.

Binary Variable: A binary variable refers to a variable that takes on only two possible values. In occupational health, binary variables are used to describe the presence or absence of a risk factor or the occurrence or non-occurrence of an event. For instance, a binary variable might be used to describe whether a worker is exposed to a hazardous substance (yes/no).

Bivariate Analysis: Bivariate analysis refers to the examination of the relationship between two variables. In occupational health, bivariate analysis is used to examine the relationship between risk factors and health outcomes. For example, bivariate analysis might be used to examine the relationship between worker exposure to noise and the development of hearing loss.

Case-Control Study: A case-control study refers to a research design in which cases (individuals with a disease or condition) are compared to controls (individuals without the disease or condition). In occupational health, case-control studies are used to examine the relationship between risk factors and work-related illnesses. For instance, a case-control study might be used to examine the relationship

between worker exposure to asbestos and the development of lung cancer.

Causal Model: A causal model refers to a theoretical framework that describes the causal relationships between variables. In occupational health, causal models are used to examine the causes of workplace injuries or illnesses. For example, a causal model might be used to describe the relationship between worker fatigue, equipment malfunction, and workplace accidents.

Censoring: Censoring refers to the loss of follow-up or incomplete data. In occupational health, censoring can occur when workers leave their jobs or are lost to follow-up, resulting in incomplete data on their health outcomes. For instance, censoring might occur in a study of worker exposure to a hazardous substance, resulting in incomplete data on the long-term health effects of exposure.

Coefficient of Variation: The coefficient of variation refers to a measure of variability that is expressed as a percentage. In occupational health, the coefficient of variation is used to describe the variability of worker exposure to hazardous substances. For example, the coefficient of variation might be calculated to describe the variability of noise exposure among workers in a particular industry.

Confidence Interval: A confidence interval refers to a range of values within which a parameter is likely to lie. In occupational health, confidence intervals are used to describe the uncertainty associated with estimates of worker exposure to hazardous substances or the frequency of workplace injuries. For instance, a confidence interval might be calculated to describe the uncertainty associated with an estimate of the mean level of noise exposure among workers in a particular industry.

Confounding Variable: A confounding variable refers to a variable that is related to both the exposure and the outcome. In occupational health, confounding variables can affect the validity of conclusions drawn from data and must be controlled for in the analysis. For example, a confounding variable might be worker age, which is related to both exposure to a hazardous substance and the development of a work-related illness.

Contingency Table: A contingency table refers to a table used to display the relationship between two or more categorical variables. In occupational health, contingency tables are used to examine the relationship between risk factors and health outcomes. For instance, a contingency table might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Continuous Variable: A continuous variable refers to a variable that can take on any value within a given range. In occupational health, continuous variables are used to describe the level of worker exposure to hazardous substances or the frequency of workplace injuries. For example, a continuous variable might be used to describe the level of noise exposure among workers in a particular industry.

Control Group: A control group refers to a group of individuals who do not receive the intervention or treatment being studied. In occupational health, control groups are used to compare the outcomes of workers who receive a safety training program to those who do not. For instance, a control group might be used to evaluate the effectiveness of a new safety training program in reducing workplace injuries.

Control Measure: A control measure refers to a method or procedure used to reduce or eliminate a hazard.

In occupational health, control measures are used to reduce worker exposure to hazardous substances or to prevent workplace injuries. For example, a control measure might be the use of personal protective equipment (PPE) to reduce worker exposure to hazardous chemicals.

Correlation Coefficient: A correlation coefficient refers to a measure of the strength and direction of the relationship between two variables. In occupational health, correlation coefficients are used to examine the relationship between risk factors and health outcomes. For instance, a correlation coefficient might be calculated to examine the relationship between worker exposure to noise and the development of hearing loss.

Cox Proportional Hazards Model: A Cox proportional hazards model refers to a statistical model used to examine the relationship between one or more variables and the risk of an event occurring. In occupational health, Cox proportional hazards models are used to examine the relationship between risk factors and worker health outcomes. For example, a Cox proportional hazards model might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Cross-Sectional Study: A cross-sectional study refers to a research design in which data are collected at a single point in time. In occupational health, cross-sectional studies are used to examine the relationship between risk factors and health outcomes at a particular point in time. For instance, a cross-sectional study might be used to examine the relationship between worker exposure to hazardous substances and the prevalence of work-related illnesses.

Crude Rate: A crude rate refers to a rate that has not been adjusted for the effects of one or more variables. In occupational health, crude rates are used to describe the frequency of workplace injuries or illnesses without controlling for factors such as age, sex, or occupation. For example, a crude rate might be calculated to describe the number of injuries that occur per year in a particular industry.

Data Quality: Data quality refers to the accuracy, completeness, and reliability of data. In occupational health, data quality is critical for ensuring that conclusions drawn from data are valid and reliable. For instance, data quality might be affected by errors in data collection or incomplete data on worker exposure to hazardous substances.

Decision Tree: A decision tree refers to a graphical representation of a decision-making process. In occupational health, decision trees are used to identify the most effective interventions or control measures for reducing worker exposure to hazardous substances or preventing workplace injuries. For example, a decision tree might be used to identify the most effective control measures for reducing worker exposure to noise in a particular industry.

Dose-Response Relationship: A dose-response relationship refers to the relationship between the level of exposure to a substance and the resulting health effect. In occupational health, dose-response relationships are used to examine the relationship between worker exposure to hazardous substances and the development of work-related illnesses. For instance, a dose-response relationship might be observed between worker exposure to asbestos and the development of lung cancer.

Ecological Study: An ecological study refers to a research design in which the unit of analysis is a group or

population rather than an individual. In occupational health, ecological studies are used to examine the relationship between risk factors and health outcomes at the population level. For example, an ecological study might be used to examine the relationship between the prevalence of work-related illnesses and the level of worker exposure to hazardous substances in different industries.

Effect Modifier: An effect modifier refers to a variable that modifies the relationship between a risk factor and a health outcome. In occupational health, effect modifiers can affect the validity of conclusions drawn from data and must be controlled for in the analysis. For instance, an effect modifier might be worker age, which can modify the relationship between exposure to a hazardous substance and the development of a work-related illness.

Empirical Bayes Method: An empirical Bayes method refers to a statistical approach that uses Bayes' theorem to update probabilities based on new data. In occupational health, empirical Bayes methods are used to analyze data and make predictions about worker health risks. For example, an empirical Bayes method might be used to estimate the probability of a worker developing a work-related illness based on their medical history and job duties.

Environmental Monitoring: Environmental monitoring refers to the process of measuring and assessing the level of exposure to hazardous substances in the workplace. In occupational health, environmental monitoring is used to identify and control hazards in the workplace. For instance, environmental monitoring might involve measuring the level of noise exposure among workers in a particular industry.

Epidemiology: Epidemiology refers to the study of the distribution and determinants of health-related events, diseases, or health-related characteristics among populations. In occupational health, epidemiology is used to examine the relationship between risk factors and worker health outcomes. For example, epidemiology might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Error: Error refers to a deviation from the true value. In occupational health, error can occur in the collection or analysis of data and can affect the validity of conclusions drawn from the data. For instance, error might occur in the measurement of worker exposure to hazardous substances, resulting in inaccurate estimates of the level of exposure.

Exposure Assessment: Exposure assessment refers to the process of measuring and characterizing the level of exposure to hazardous substances in the workplace. In occupational health, exposure assessment is used to identify and control hazards in the workplace. For example, exposure assessment might involve measuring the level of noise exposure among workers in a particular industry.

Exposure Limit: An exposure limit refers to the maximum allowable level of exposure to a hazardous substance. In occupational health, exposure limits are used to protect workers from hazardous substances and are often established by regulatory agencies. For instance, an exposure limit might be established for worker exposure to lead in the workplace.

False Negative: A false negative refers to a test result that incorrectly indicates the absence of a condition or disease. In occupational health, false negatives can occur in the diagnosis of work-related illnesses and can

result in delayed or inadequate treatment. For example, a false negative might occur in the diagnosis of lung cancer, resulting in delayed treatment and a poorer prognosis.

False Positive: A false positive refers to a test result that incorrectly indicates the presence of a condition or disease. In occupational health, false positives can occur in the diagnosis of work-related illnesses and can result in unnecessary treatment or interventions. For instance, a false positive might occur in the diagnosis of a work-related illness, resulting in unnecessary medical treatment and lost productivity.

Frequency: Frequency refers to the number of times an event occurs over a specified period. In occupational health, frequency is used to describe the number of workplace injuries or illnesses that occur over a given period. For example, frequency might be used to describe the number of injuries that occur per year in a particular industry.

Generalized Linear Model: A generalized linear model refers to a statistical model that is used to examine the relationship between one or more variables and a continuous outcome variable. In occupational health, generalized linear models are used to examine the relationship between risk factors and worker health outcomes. For instance, a generalized linear model might be used to examine the relationship between worker exposure to noise and the development of hearing loss.

Hazard: A hazard refers to a condition or situation that has the potential to cause harm or injury. In occupational health, hazards can include physical, chemical, or biological agents that can cause harm to workers. For example, a hazard might be the presence of asbestos in the workplace, which can cause lung cancer and other respiratory diseases.

Hazard Ratio: A hazard ratio refers to a measure of the relative risk of an event occurring. In occupational health, hazard ratios are used to examine the relationship between risk factors and worker health outcomes. For instance, a hazard ratio might be calculated to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Health Effect: A health effect refers to a change in a worker's health status that is caused by exposure to a hazardous substance or other occupational hazard. In occupational health, health effects can range from mild and temporary to severe and permanent. For example, a health effect might be the development of lung cancer in a worker who is exposed to asbestos in the workplace.

Incidence: Incidence refers to the number of new cases of a disease or condition that occur over a specified period. In occupational health, incidence is used to describe the number of new cases of work-related illnesses that occur over a given period. For example, incidence might be used to describe the number of new cases of lung cancer that occur per year in a particular industry.

Incidence Rate: An incidence rate refers to the rate at which new cases of a disease or condition occur over a specified period. In occupational health, incidence rates are used to describe the rate at which new cases of work-related illnesses occur over a given period. For instance, an incidence rate might be calculated to describe the rate at which new cases of lung cancer occur per year in a particular industry.

Independent Variable: An independent variable refers to a variable that is being manipulated or changed in

an experiment or study. In occupational health, independent variables can include risk factors such as worker exposure to hazardous substances or the presence of workplace hazards. For example, an independent variable might be the level of noise exposure among workers in a particular industry.

Inferential Statistics: Inferential statistics refers to the process of using sample data to make inferences about a larger population. In occupational health, inferential statistics are used to examine the relationship between risk factors and worker health outcomes and to make predictions about the health effects of exposure to hazardous substances. For instance, inferential statistics might be used to estimate the probability of a worker developing a work-related illness based on their medical history and job duties.

Information Bias: Information bias refers to a systematic error or distortion in the collection or interpretation of data. In occupational health, information bias can occur in the collection or analysis of data and can affect the validity of conclusions drawn from the data. For example, information bias might occur in the measurement of worker exposure to hazardous substances, resulting in inaccurate estimates of the level of exposure.

Intervention Study: An intervention study refers to a research design in which a treatment or intervention is applied to a group of individuals to examine its effect on a health outcome. In occupational health, intervention studies are used to examine the effectiveness of safety programs or control measures in reducing worker exposure to hazardous substances or preventing workplace injuries. For instance, an intervention study might be used to evaluate the effectiveness of a new safety training program in reducing workplace injuries.

Kaplan-Meier Estimate: A Kaplan-Meier estimate refers to a statistical method used to estimate the probability of survival or the occurrence of an event over time. In occupational health, Kaplan-Meier estimates are used to examine the relationship between risk factors and worker health outcomes. For example, a Kaplan-Meier estimate might be used to estimate the probability of a worker developing a work-related illness based on their medical history and job duties.

Latin Square Design: A Latin square design refers to a research design in which the treatment or intervention is applied to a group of individuals in a specific order. In occupational health, Latin square designs are used to examine the effectiveness of safety programs or control measures in reducing worker exposure to hazardous substances or preventing workplace injuries. For instance, a Latin square design might be used to evaluate the effectiveness of a new safety training program in reducing workplace injuries.

Least Squares Method: A least squares method refers to a statistical method used to estimate the parameters of a model by minimizing the sum of the squared errors. In occupational health, least squares methods are used to examine the relationship between risk factors and worker health outcomes. For example, a least squares method might be used to estimate the relationship between worker exposure to noise and the development of hearing loss.

Life Table: A life table refers to a table that displays the probability of survival or the occurrence of an event over time. In occupational health, life tables are used to examine the relationship between risk factors and worker health outcomes. For instance, a life table might be used to estimate the probability of a worker

developing a work-related illness based on their medical history and job duties.

Logistic Regression: Logistic regression refers to a statistical model used to examine the relationship between one or more variables and a binary outcome variable. In occupational health, logistic regression is used to examine the relationship between risk factors and worker health outcomes. For example, logistic regression might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Longitudinal Study: A longitudinal study refers to a research design in which data are collected over a long period. In occupational health, longitudinal studies are used to examine the relationship between risk factors and worker health outcomes over time. For instance, a longitudinal study might be used to examine the relationship between worker exposure to hazardous substances and the development of work-related illnesses over a period of many years.

Lost Time Injury: A lost time injury refers to an injury that results in a worker being unable to work for a period of time. In occupational health, lost time injuries are used to describe the severity of workplace injuries and to evaluate the effectiveness of safety programs or control measures. For example, a lost time injury might be used to describe the number of days that a worker is unable to work due to a work-related injury.

Matched Case-Control Study: A matched case-control study refers to a research design in which cases (individuals with a disease or condition) are matched to controls (individuals without the disease or condition) based on one or more characteristics. In occupational health, matched case-control studies are used to examine the relationship between risk factors and work-related illnesses. For instance, a matched case-control study might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Maximum Likelihood Estimate: A maximum likelihood estimate refers to a statistical method used to estimate the parameters of a model by maximizing the likelihood of the observed data. In occupational health, maximum likelihood estimates are used to examine the relationship between risk factors and worker health outcomes. For example, a maximum likelihood estimate might be used to estimate the relationship between worker exposure to noise and the development of hearing loss.

Mean: The mean refers to a measure of central tendency that is calculated by summing all the values and dividing by the number of values. In occupational health, the mean is used to describe the typical values of worker exposure to hazardous substances or the frequency of workplace injuries. For instance, the mean might be calculated to describe the average level of noise exposure among workers in a particular industry.

Measurement Error: Measurement error refers to a deviation from the true value that occurs during the measurement process. In occupational health, measurement error can occur in the collection or analysis of data and can affect the validity of conclusions drawn from the data. For example, measurement error might occur in the measurement of worker exposure to hazardous substances, resulting in inaccurate estimates of the level of exposure.

Median: The median refers to a measure of central tendency that is the middle value in a dataset. In

occupational health, the median is used to describe the typical values of worker exposure to hazardous substances or the frequency of workplace injuries. For instance, the median might be calculated to describe the median level of noise exposure among workers in a particular industry.

Meta-Analysis: A meta-analysis refers to a statistical method used to combine the results of multiple studies to draw more general conclusions. In occupational health, meta-analyses are used to examine the relationship between risk factors and worker health outcomes. For example, a meta-analysis might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Misclassification: Misclassification refers to the incorrect classification of a worker or a disease or condition. In occupational health, misclassification can occur in the collection or analysis of data and can affect the validity of conclusions drawn from the data. For instance, misclassification might occur in the diagnosis of a work-related illness, resulting in delayed or inadequate treatment.

Mode: The mode refers to a measure of central tendency that is the most frequently occurring value in a dataset. In occupational health, the mode is used to describe the typical values of worker exposure to hazardous substances or the frequency of workplace injuries. For example, the mode might be calculated to describe the most common level of noise exposure among workers in a particular industry.

Morbidity: Morbidity refers to the state of being diseased or ill. In occupational health, morbidity is used to describe the incidence or prevalence of work-related illnesses. For instance, morbidity might be used to describe the number of workers who develop a work-related illness over a given period.

Mortality: Mortality refers to the state of being dead. In occupational health, mortality is used to describe the number of workers who die from work-related illnesses or injuries. For example, mortality might be used to describe the number of workers who die from lung cancer caused by exposure to asbestos in the workplace.

Multivariate Analysis: Multivariate analysis refers to the examination of the relationship between two or more variables and a continuous outcome variable. In occupational health, multivariate analysis is used to examine the relationship between risk factors and worker health outcomes. For instance, multivariate analysis might be used to examine the relationship between worker exposure to noise, asbestos, and the development of lung cancer.

Natural Experiment: A natural experiment refers to a research design in which the treatment or intervention is applied by nature or by circumstances beyond the control of the researcher. In occupational health, natural experiments are used to examine the relationship between risk factors and worker health outcomes. For example, a natural experiment might be used to examine the relationship between worker exposure to hazardous substances and the development of work-related illnesses.

Noise-Induced Hearing Loss: Noise-induced hearing loss refers to a type of hearing loss that is caused by exposure to loud noise. In occupational health, noise-induced hearing loss is a common work-related illness that can be caused by exposure to loud noise in the workplace. For instance, noise-induced hearing loss might occur in workers who are exposed to loud noise in a manufacturing plant.

Null Hypothesis: A null hypothesis refers to a statement that there is no effect or relationship between variables. In occupational health, null hypotheses are used to test the effectiveness of interventions or to examine the relationship between risk factors and health outcomes. For example, a null hypothesis might be used to test the effectiveness of a new safety training program in reducing workplace injuries.

Observational Study: An observational study refers to a research design in which the researcher observes the subjects without intervening or manipulating the treatment. In occupational health, observational studies are used to examine the relationship between risk factors and worker health outcomes. For instance, an observational study might be used to examine the relationship between worker exposure to hazardous substances and the development of work-related illnesses.

Occupational Disease: An occupational disease refers to a disease or condition that is caused by exposure to a hazardous substance or other occupational hazard. In occupational health, occupational diseases can range from mild and temporary to severe and permanent. For example, an occupational disease might be lung cancer caused by exposure to asbestos in the workplace.

Odds Ratio: An odds ratio refers to a measure of the association between a risk factor and a health outcome. In occupational health, odds ratios are used to examine the relationship between risk factors and worker health outcomes. For instance, an odds ratio might be calculated to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Outlier: An outlier refers to a value that is significantly different from the other values in a dataset. In occupational health, outliers can occur in the collection or analysis of data and can affect the validity of conclusions drawn from the data. For example, an outlier might occur in the measurement of worker exposure to hazardous substances, resulting in inaccurate estimates of the level of exposure.

Overmatching: Overmatching refers to the process of matching cases and controls too closely, resulting in a loss of information or power. In occupational health, overmatching can occur in the design of studies and can affect the validity of conclusions drawn from the data. For instance, overmatching might occur in a study of the relationship between worker exposure to hazardous substances and the development of work-related illnesses.

Parametric Test: A parametric test refers to a statistical test that assumes a specific distribution of the data. In occupational health, parametric tests are used to examine the relationship between risk factors and worker health outcomes. For example, a parametric test might be used to examine the relationship between worker exposure to noise and the development of hearing loss.

Parsimony: Parsimony refers to the principle of simplicity or economy in the design of studies or the analysis of data. In occupational health, parsimony is used to guide the selection of variables and the analysis of data. For instance, parsimony might be used to guide the selection of variables to include in a model of the relationship between worker exposure to hazardous substances and the development of work-related illnesses.

Pearson Correlation Coefficient: A Pearson correlation coefficient refers to a measure of the strength and direction of the relationship between two variables. In occupational health, Pearson correlation coefficients

are used to examine the relationship between risk factors and worker health outcomes. For example, a Pearson correlation coefficient might be calculated to examine the relationship between worker exposure to noise and the development of hearing loss.

Percentile: A percentile refers to a measure of the percentage of values that are below a certain value. In occupational health, percentiles are used to describe the distribution of worker exposure to hazardous substances or the frequency of workplace injuries. For instance, a percentile might be calculated to describe the percentage of workers who are exposed to hazardous substances at or below a certain level.

Personal Protective Equipment: Personal protective equipment (PPE) refers to equipment or devices that are worn by workers to protect them from hazardous substances or other occupational hazards. In occupational health, PPE is used to reduce worker exposure to hazardous substances and to prevent workplace injuries. For example, PPE might include respirators, gloves, or safety glasses.

Poisson Distribution: A Poisson distribution refers to a statistical distribution that is used to model the number of events that occur over a fixed period. In occupational health, Poisson distributions are used to examine the relationship between risk factors and worker health outcomes. For instance, a Poisson distribution might be used to model the number of workplace injuries that occur over a given period.

Population: A population refers to the group of individuals or workers that are being studied. In occupational health, populations can range from a small group of workers in a single industry to a large group of workers across multiple industries. For example, a population might be all workers in a particular industry or all workers who are exposed to a hazardous substance.

Positive Predictive Value: The positive predictive value refers to the probability that a worker who tests positive for a disease or condition actually has the disease or condition. In occupational health, the positive predictive value is used to evaluate the accuracy of diagnostic tests or screening programs. For instance, the positive predictive value might be used to evaluate the accuracy of a test for lung cancer in workers who are exposed to asbestos.

Power: Power refers to the ability of a study to detect a statistically significant effect or relationship. In occupational health, power is used to determine the sample size required to detect a certain effect or relationship. For example, power might be used to determine the sample size required to detect a significant relationship between worker exposure to hazardous substances and the development of work-related illnesses.

Precision: Precision refers to the degree of closeness between repeated measurements or estimates. In occupational health, precision is used to evaluate the reliability of data or the accuracy of estimates. For instance, precision might be used to evaluate the reliability of measurements of worker exposure to hazardous substances.

Prevalence: Prevalence refers to the number of cases of a disease or condition that exist in a population at a given time. In occupational health, prevalence is used to describe the number of workers who have a work-related illness or condition. For example, prevalence might be used to describe the number of workers who have lung cancer caused by exposure to asbestos in the workplace.

Preventive Medicine: Preventive medicine refers to the practice of preventing or reducing the risk of disease or injury. In occupational health, preventive medicine is used to reduce the risk of work-related illnesses or injuries. For instance, preventive medicine might include programs to reduce worker exposure to hazardous substances or to promote healthy behaviors among workers.

Primary Prevention: Primary prevention refers to the intervention or action taken to prevent a disease or condition from occurring. In occupational health, primary prevention is used to reduce the risk of work-related illnesses or injuries. For example, primary prevention might include programs to reduce worker exposure to hazardous substances or to promote healthy behaviors among workers.

Probability: Probability refers to the chance or likelihood that an event will occur. In occupational health, probability is used to describe the likelihood of a worker developing a work-related illness or condition. For instance, probability might be used to describe the likelihood of a worker developing lung cancer caused by exposure to asbestos in the workplace.

Proportional Hazards Model: A proportional hazards model refers to a statistical model that is used to examine the relationship between one or more variables and the risk of an event occurring. In occupational health, proportional hazards models are used to examine the relationship between risk factors and worker health outcomes. For example, a proportional hazards model might be used to examine the relationship between worker exposure to asbestos and the development of lung cancer.

Prospective Study: A prospective study refers to a research design in which the subjects are followed over time to examine the relationship between risk factors and health outcomes. In occupational health, prospective studies are used to examine the relationship between worker exposure to hazardous substances and the development of work-related illnesses. For instance, a prospective study might be used to examine the