



TRAINING COURSE ON DATA MANAGEMENT AND DATA ANALYSIS IN HEALTH SCIENCE RESEARCH

1. Objectives

After the training course, participants should be able to:

- Describe the basic study designs in health science research
- Describe the basic concepts in statistics, including hypothesis testing, p-value, and confidence interval
- Create the data entry form using Epidata. Conduct data processing and cleaning.
- Use STATA software in processing, analyzing data, and interpreting the results.

2. **Format:** 6 sessions (March 13-15th, 2020)

3. **Location:** Hanoi University of Public Health, 1A Duc Thang Road, Duc Thang Ward, North Tu Liem District, Hanoi

4. Syllabus

Section 1: <i>(March 13rd)</i>	Basics of Health Science Research <ul style="list-style-type: none">- Basic quantitative study designs in health science research- Basic concepts in probability and statistics- Hypothesis testing, p-value, and confidence interval
Section 2: <i>(March 13rd)</i>	Data management <ul style="list-style-type: none">- Creating the data entry form using Epidata- Introduction to Stata software- Some commonly used Stata commands in data management
Section 3: <i>(March 14th)</i>	Data cleaning and descriptive statistics <ul style="list-style-type: none">- Data cleaning with Stata<ul style="list-style-type: none">+ Cleaning data+ Generating, editing variables- Descriptive statistics with Stata<ul style="list-style-type: none">+ Descriptive statistics for quantitative variables+ Descriptive statistics for qualitative variables
Section 4: <i>(March 14th)</i>	Hypothesis testing <ul style="list-style-type: none">- Hypothesis testing for continuous variables<ul style="list-style-type: none">+ t-test+ ANOVA+ Non-parametric tests+ Correlation tests (Pearson r, Spearman...)- Hypothesis testing for binary and categorical variables<ul style="list-style-type: none">+ One proportion Z test+ Two proportion Z test+ Chi-squared test+ Fisher exact test+ Relative risk and odds ratio
Section 5: <i>(March 15th)</i>	Regression <ul style="list-style-type: none">- Linear regression- Data transformation methods- Generalized linear models<ul style="list-style-type: none">+ Logistic regression+ Poisson regression- Using OR/PR in cross-sectional study

Section 6: Methods for building multivariable regression models and create result tables for publication with Stata

(March 15th)

- Methods for building multivariable regression models
- Using Directed Acyclic Graphs (DAGs) in causal inference and multivariable model construction
- Creating result tables for publication with Stata
- Discussion and questions