Building your Biostatistics Toolbox for Precepting Journal Clubs

Michael C. Brown, PharmD
Associate Dean for Academic Affairs and Professor
Concordia University School of Pharmacy

Audrey B. Kostrzewa, PharmD, MPH, BCPS
Assistant Professor
Concordia University School of Pharmacy

Kajua B. Lor, PharmD, BCACP
Chair, Department of Clinical Sciences & Practice Advancement and Associate Professor
Medical College of Wisconsin School of Pharmacy

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Conflict of Interest Statement

- None of the presenters have any conflicts of interest to disclose
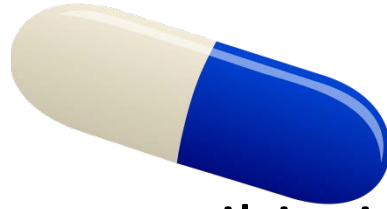
Objectives

- Describe the statistical concepts of Type I Error, Type II Error, p-values, confidence intervals, and risk ratios.

- Apply statistical concepts in sample literature case studies.

- Explain when to use statistical analyses such as chi square, survival analysis, t-test, ANOVA, and regression/correlation.

- Assess if the statistical analyses used match the study design and outcomes in a study and discuss their implications.

- Interpret the clinical and statistical significance of statistics in sample literature case studies.

- Describe teaching pearls and best practices for precepting journal clubs and explain where to find resources for precepting journal clubs.

Michael C. Brown, PharmD

TYPE I ERROR, TYPE II ERROR, P-VALUE, CONFIDENCE INTERVALS

Type I & Type II Error

Example

- *GMCCP527* is a new antibiotic being studied for treatment for sinusitis.

- To test whether GMCCP527 is effective or not, a clinical trial was performed as follows: a sample of 160 subjects with sinus infections were randomized to receive either GMCCP527 or placebo twice daily for 10 days.

Errors in Hypothesis Testing

"Truth about the population"

What we are trying to find out by doing one or more studies BUT WE NEVER GET TO KNOW

Outcome same	Outcome different
GMCCP527 does not work	GMCCP527 works

Adapted from Moore, DS and McCabe, GP. Introduction to the Practice of Statistics, 5th Edition. W.H. Freeman and Co., NY:2006, p. 436.

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Errors in Hypothesis Testing

What we see when we do one or more studies

Outcomes Different
GMCCP527 works

Decision
based
on
sample

Outcomes Same
GMCCP527 does
not work

Adapted from Moore, DS and McCabe, GP. <u>Introduction to the Practice of Statistics</u>, 5th Edition. W.H. Freeman and Co., NY:2006, p. 436.

Summary Table

"Truth" about the population

	Outcome same GMCCP527 does not work	Outcome different GMCCP527 works
Decision based on sample — Outcomes Different GMCCP527 works	Study shows GMCCP527 works, but it doesn't in population	Study shows GMCCP527 works and in population it does
Outcomes Same GMCCP527 does not work	Study shows GMCCP527 does not work and in population it would not	Study shows GMCCP527 does not work, but in population it would have worked

Adapted from Moore, DS and McCabe, GP. <u>Introduction to the Practice of Statistics</u>, 5th Edition. W.H. Freeman and Co., NY:2006, p. 436.

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Type I Error & Alpha

- **<u>Type I Error</u>**: the probability that a study concludes there IS a difference between two treatments when there "actually" is not (wrong decision)

- **<u>Alpha</u>**: the <u>*acceptable threshold*</u> of probability that a study would conclude that there IS a difference between two treatments when there "actually" is not (it was CHANCE)
  - Usually set at 0.05 or 5%

Type II Error & Beta

- **Type II Error**: the possibility that a study concludes there is NOT a difference between two treatments when there "actually" is one (wrong decision)

- **Beta**: the *acceptable* probability that a study would conclude that there is NOT a difference between two treatments when there "actually" is
  - Usually set at 0.2 or 20%

Summary Table

"Truth" about the population

	Outcome same GMCCP527 does not work	Outcome different GMCCP527 works
Decision based on sample — Outcomes Different GMCCP527 works	**Type I error** (threshold α)	**Correct decision**
Outcomes Same GMCCP527 does not work	**Correct decision**	**Type II error** (threshold β)

Adapted from Moore, DS and McCabe, GP. Introduction to the Practice of Statistics, 5th Edition. W.H. Freeman and Co., NY:2006, p. 436.

p-value

p-value

- **p-value** is the probability of an outcome this or more extreme occurring due to only chance (that is, assuming the null hypothesis is "true")

- **p-value** is the probability that the study found a difference by CHANCE (the difference doesn't actually exist)

- "**p-value** is the probability that the null hypothesis is true"
  - This is technically not statistically correct, although some find it helpful to "think" about p-values this way

- p-value can range from nearly zero to exactly 1

p-value and Statistical Significance

- What does a p-value < 0.05 (alpha) mean?

  – Statistically significantly different

- What does a p-value > 0.05 (alpha) mean?

  – Not statistically significantly different

More about p-value

– Some p-values are reported as exact values, but they are often reported as "less than (<)" a certain value. This may be due to the sensitivity of the tool used to determine p-value. Interpretation should not change.

- Can there be a gradient of statistical significance? (ex. p <0.5 vs. p<0.01 vs. p<0.001)
    - **NO.** Statistical significance is a yes or no question. There is no such thing as *very* significant, *almost* significant, etc.

Clinical vs Statistical Significance

- Frequently CLINICAL and STATISTICAL SIGNIFICANCE agree but they are NOT equivalent

  - Consider if…

    GMCCP527 cost $10,000 per treatment AND

    Cure rate 85% for GMCCP527 vs 83% for sulfa at 2 weeks AND

    $p=0.03$

  - Consider if…

    GMCCP527 comparably priced to sulfa

    Infection being treated led to hospitalization 50% of time if not cured in two weeks AND

    Cure rate 97% for GMCCP527 vs 92% for sulfa at 2 weeks AND

    $p=0.073$

Confidence Intervals

Confidence Interval

- Definition
  - A 95% CI contains 95% of the point estimates (means, ratios, etc) that would be found for multiple samples drawn from the population in the same way as the study's sample
    - 95% certain point estimate falls in this range

  - It does NOT say 95% CI contains 95% of individuals' values
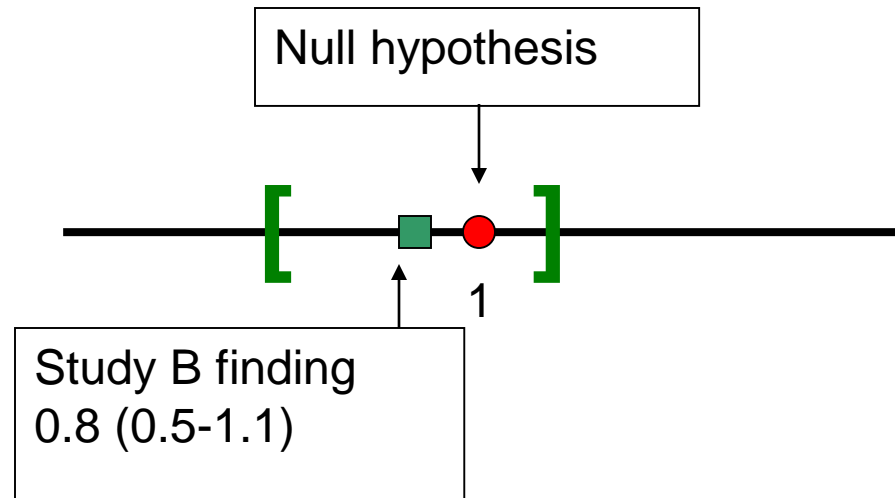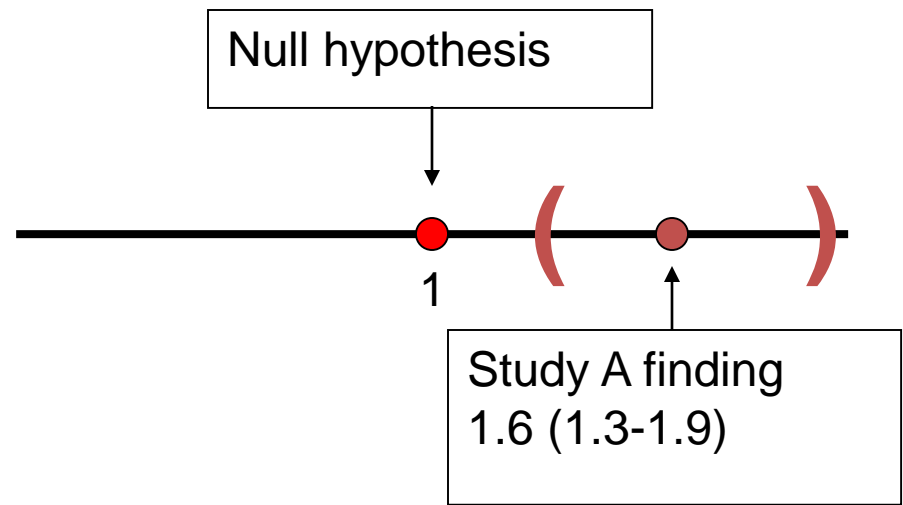
Why are confidence intervals useful?

- Clinical significance – range of values helps determine clinical usefulness/ applicability

- Ex: Change in blood pressure
  - 24.4 mm Hg (95% 11.4 – 37.4)

- Ex: Change in blood pressure
  - 1.4 mm Hg (95% 0.2 – 2.6)

Using Confidence Intervals to Determine Statistical Significance

1. Risk ratios

2. Change/difference in point estimate (change/difference in mean, change/difference in proportion, etc)

3. Comparing two independent point estimates (two independent means, two independent proportions, etc)

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Risk RATIO and CI

- Key question – Does CI include 1?
- Ex A: No: Statistically significant
- Ex B: Yes: Not statistically significant

Null hypothesis

1

Study A finding
1.6 (1.3-1.9)

Null hypothesis
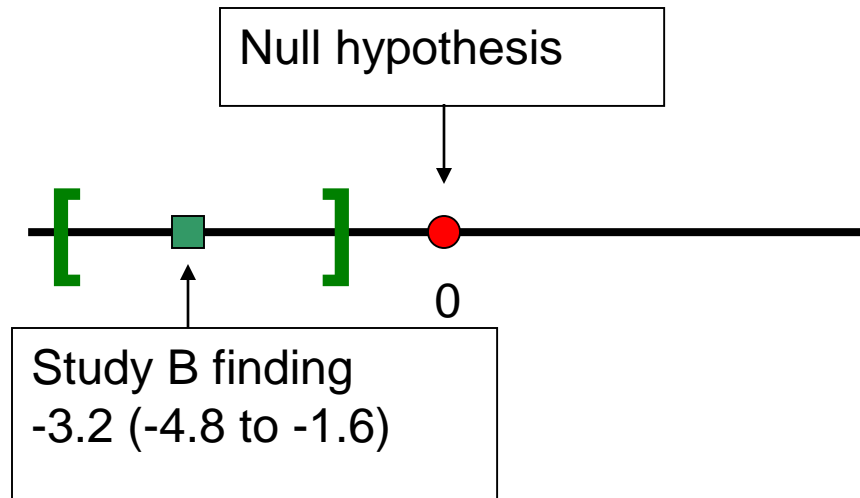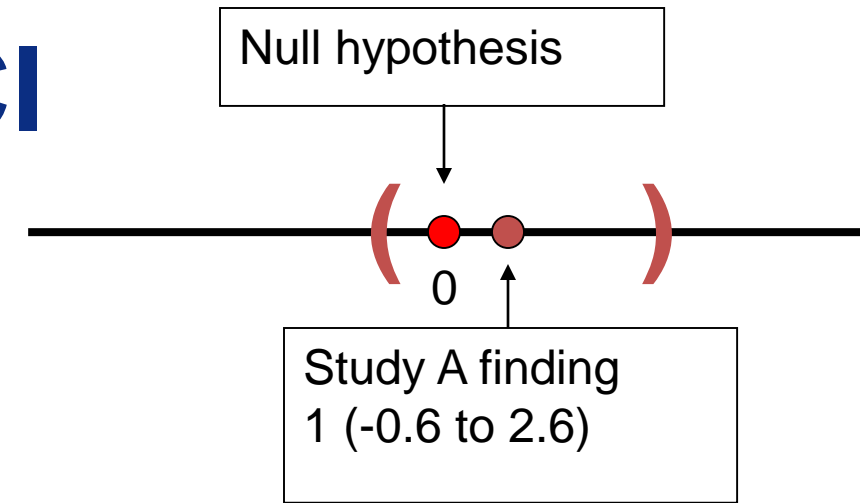
1

Study B finding
0.8 (0.5-1.1)

Change in Point Estimate

- Example: Improvement in LDL following TLC
  - Measure LDL at baseline
  - Measure LDL at 6 months
  - Outcome: change in LDL baseline to 6 months

- What value would represent no change (or no difference) between baseline and 6 months?

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Change in Point Estimate and CI

- Key question – Does CI include 0?
- Ex A: Yes: Not statistically significant
- Ex B: No: Statistically significant

Null hypothesis

Study A finding
1 (-0.6 to 2.6)

Null hypothesis

Study B finding
-3.2 (-4.8 to -1.6)

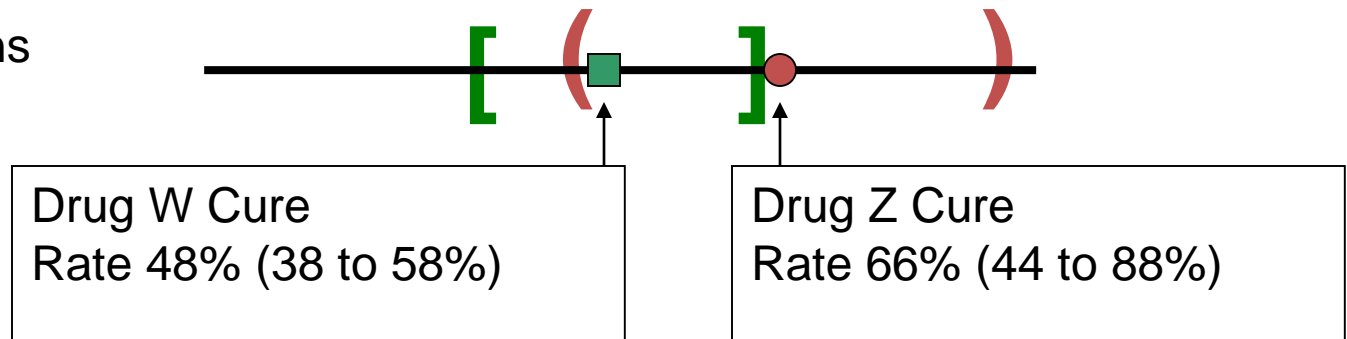# Comparing two independent point estimates

- Example: cure rates from two different antibiotic treatments (W and Z)
  - One group with infection receives drug W
    - Some percent of them are cured
  - Another group with infection receives drug Z
    - Some percent of them are cured

- What would represent no change (or no difference) in cure rates?

MCCP
Greater Milwaukee College of Clinical Pharmacy

Comparing two separate point estimates with CIs

Case 1: CI don't overlap
Statistically
Significant
Difference

Drug W Cure
Rate 43% (34 to 52%)

Drug Z Cure
Rate 66% (57 to 75%)
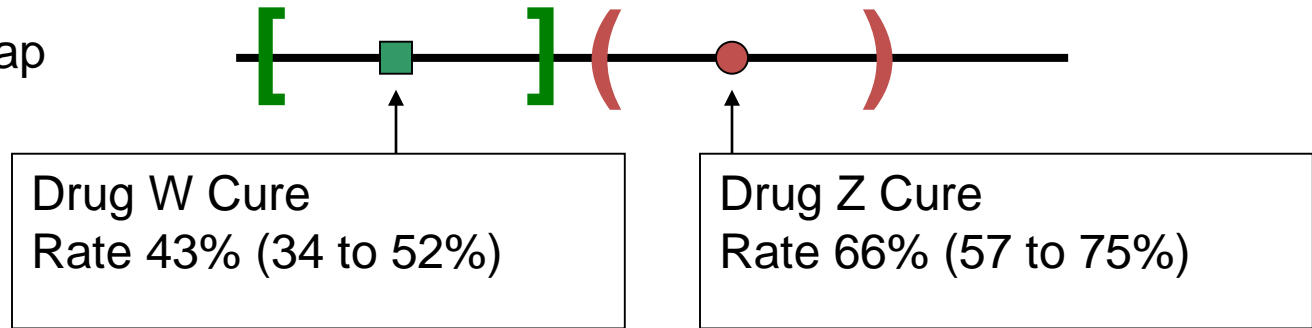
Case 2: one CI contains other point estimate
No Statistically
Significant
Difference

Drug W Cure
Rate 48% (38 to 58%)

Drug Z Cure
Rate 66% (44 to 88%)

Case 3: CIs overlap BUT point estimates don't fall in other CI: ????
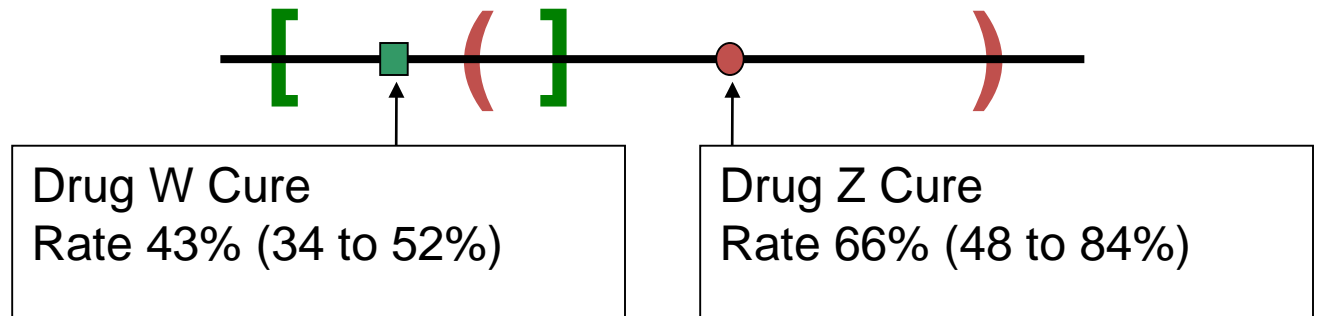
Drug W Cure
Rate 43% (34 to 52%)

Drug Z Cure
Rate 66% (48 to 84%)

GMCCP
Greater Milwaukee College of Clinical Pharmacy

References

1. Dawson B, Trapp RG. Basic & Clinical Biostatistics. 4th edition. New York, New York: Lange Medical Books/Mc-Graw-Hill; 2004.

2. DiCenzo R, ed. *Clinical Pharmacist's Guide to Biostatistics and Literature Evaluation*. Lenexa, KS: American College of Clinical Pharmacy; 2011.

Risk Ratios

When are risk ratios used?

- **Nominal data**
  - Categorical data with no inherent order
  - Eg, Y/N, proportions, rates

- **Contingency table ("2x2 table")**
  - Often how nominal data can be displayed
  - Helpful to visualize data and estimate association (calculate risk ratios) between variables

Risk Ratios

- Strength of association between risk factor and outcome
- Resulting from various regression models

- **<u>Odds Ratio (OR)</u>**
  - Used in case-control studies (retrospective)
  - The odds that a case was exposed/the odds that a control was exposed

- **<u>Relative Risk (RR)</u>**
  - Used in cohort and clinical trials (prospective)
  - Incidence in exposed subjects/incidence of unexposed subjects

- **<u>Hazard Ratio (HR)</u>**
  - Used in survival analysis (prospective)
  - Similar to OR/RR with time aspect

Determining "Risk"

	Event	No Event
Risk Factor Present (Exposed/Experimental)	A	B
Risk Factor Absent (Unexposed/Control)	C	D

- Number of patients with risk factor = A + B
- Number of patients without risk factor = C + D
- Number of patients with event = A + C
- Number of patients without event = B + D

Absolute & Relative Risk

- **Absolute Risk (Event Rates)**
  - Experimental Event Rate (EER) = A/(A+B)
  - Control Event Rate (CER) = C/(C+D)
  - Absolute Risk Reduction (ARR) = | EER – CER |
  - Number Needed to Treat (NNT) = 1/ARR

- **Relative Risk**
  - Change from baseline risk
  - Relative Risk Reduction (RRR) = $\dfrac{|\ EER - CER\ |}{CER}$

What sounds more impressive?

- 17.9% or 63%?

- Extreme example of RRR misuse:
    - Risk of disease X in general population = 0.2% (2 in every 1,000 people will get disease X)
    - Drug A vs. placebo tested in 20,000 people. 5/10,000 people treated with drug A diagnosed with disease X. 20/10,000 people treated with placebo diagnosed with disease X.
    - ARR = |0.0005 – 0.002| = 0.0015 x 100% = 0.15%
    - RRR = 0.0015/0.002 = 0.75 x 100% = 75%
    - What sounds more impressive? Reducing risk of disease X by 0.15% or by 75%?
    - If baseline risk or ARR not reported or obvious, reader may not know the clinical impact of a 75% RRR
        - 75% RRR could mean 0.2%→0.05% or potentially 92%→23%

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Calculating Risk Ratios

	Event	No Event
Risk Factor Present (Exposed)	A	B
Risk Factor Absent (Unexposed)	C	D

Odds Ratio (OR) $= \dfrac{[A/(A+C)]/[C/(A+C)]}{[B/(B+D)]/[D/(B+D)]} = (A/C)/(B/D) =$ **AD/BC**

Relative Risk (RR) $= \dfrac{A/(A+B)}{C/(C+D)} =$ **EER/CER**

Hazard Ratio (HR) cannot be manually calculated

MCCP
Greater Milwaukee College of Clinical Pharmacy

Adapted from Dawson and Trapp. Basic & Clinical Biostatistics. 4th ed. Lange Medical Books. 2004.

Interpreting Risk Ratios

- All essentially interpreted the same way:
  - If risk ratio = 1, risk factor is not associated with event
  - If risk ratio > 1, risk factor is positively associated with event
  - If risk ratio < 1, risk factor is negatively associated with event

- Interpretation of risk ratio (OR/RR) provides clinical significance

- Still need statistical test to provide measure of statistical significance
  - P-value and/or confidence interval (if crosses 1, not statistically significant)

Additional Note: Unadjusted vs. Adjusted

- Relationship of two factors being studied may be influenced by other factors (confounding variables)

- Unadjusted or crude = only two factors being studied are taken into account; likely able to calculate by hand

- Adjusted = incorporates potential confounders into statistical model; cannot calculate by hand

- Which one is better or more accurate?

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Chi Squared

Chi Squared (X2)

- Think of 2x2 (contingency) table

- Non-parametric statistical test that compares whether table values are statistically different than random distribution
  - **<u>Nominal data</u>**
  - **<u>Independent data</u>**
  - **<u>2 groups or 3+ groups of data</u>**
    - If used for 3+ groups, must use post hoc correction (Bonferroni) to correct for multiple comparisons

- Results in a p-value

Fisher Exact

- Similar to chi squared, used when expected frequencies in the 2x2 table are small
  - "small" typically means 5

Survival Analysis

"Classic" Statistics to Illustrate Survival

- Mean time-to-event

- Event rates at pre-defined times
  - 5-year survival

- Events per 100 patient-years

Survival Analysis

- Time to event (eg, disease progression, death)

- Type of regression model (ie, association)
  - Often presented as a Kaplan Meier curve

- Advantages:
  - Accounts for incomplete observation (censoring)
  - Less biased by time length of study

Kaplan-Meier Curve: Key Features

- % free (or with) event

- Number of events calculated each time an event occurs

- Takes into account censored patients ("number at risk")

- P-value or HR with confidence interval to determine significance

Comparing Two Kaplan-Meier Curves

- Eyeball test is good for describing trends
  - Caution with authors making data look more spread out than it really is

- Must also use statistical test to determine statistical significance
  - Eg, Log-rank, Mantel-Haenszel Chi-Square Statistic, Cox Proportional Hazards Model
  - Results in either a p-value or HR with confidence interval
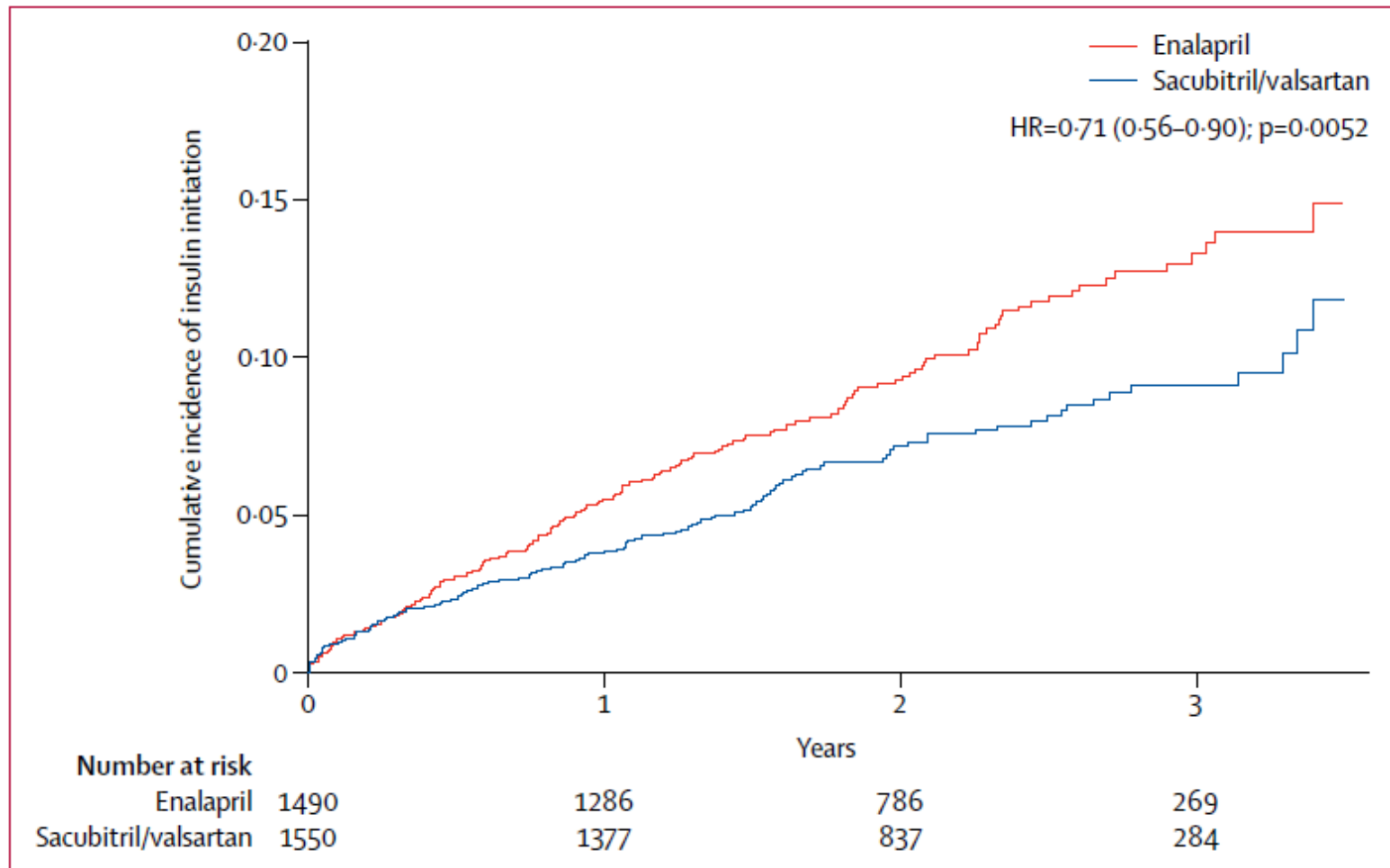
Comparing Two Kaplan-Meier Curves: Example

Figure 2: Kaplan-Meier curve showing time to insulin initiation in the sacubitril/valsartan and enalapril groups, in patients previously not treated with insulin

Seferovic JP, et al. *Lancet Diabetes Endocrinol*. 2017.

Comparing Two Kaplan-Meier Curves: Examples
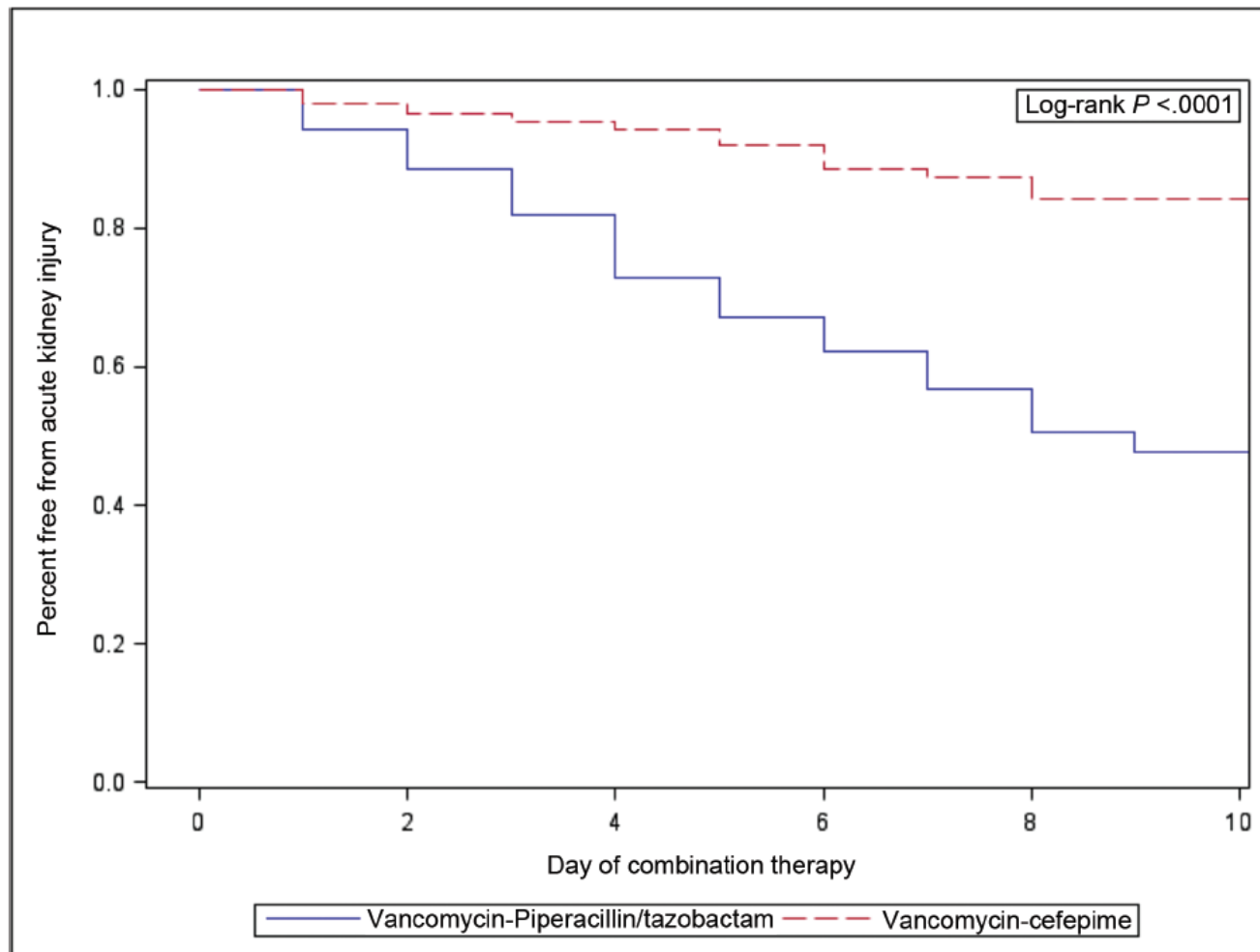
Figure 2. Kaplan–Meier survival analysis for acute kidney injury as a function of treatment group.

Navalkele B, et al. *Clin Infect Dis*. 2017;64:116-23.

References

1. Dawson B, Trapp RG. Basic & Clinical Biostatistics. 4th edition. New York, New York: Lange Medical Books/Mc-Graw-Hill; 2004.
2. DiCenzo R, ed. *Clinical Pharmacist's Guide to Biostatistics and Literature Evaluation*. Lenexa, KS: American College of Clinical Pharmacy; 2011.

Kajua B. Lor, PharmD, BCACP
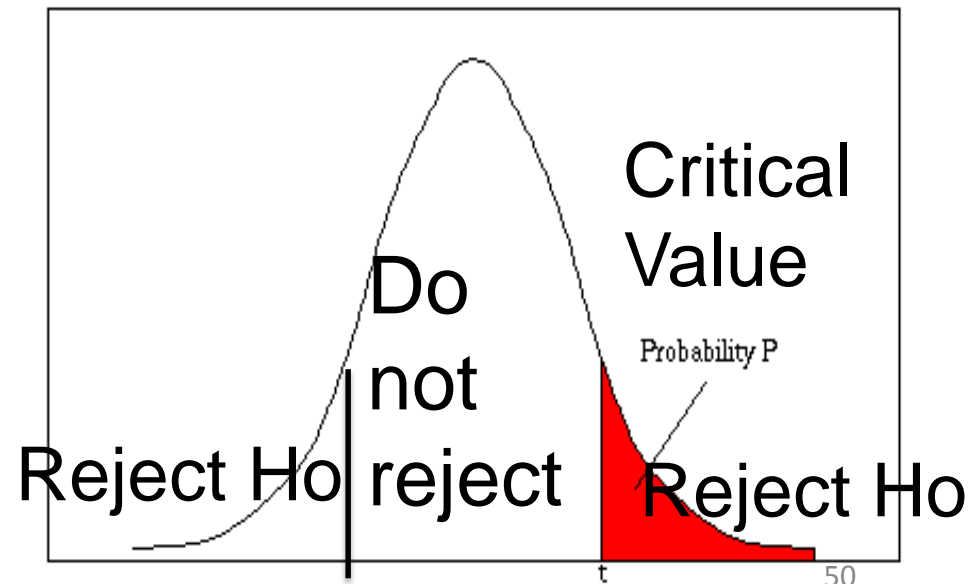
T-TEST, ANOVA, REGRESSION/CORRELATION

MCCP
Greater Milwaukee College of Clinical Pharmacy

2-level Discrete → Continuous

Note: Chi-squared not the best test to use as the critical value gets larger

	Treatment	Control
Glucose readings (mg/dL)	81	86
	81	89
	79	81
	80	87
	84	82
	87	87
	75	85
	83	89
	88	83

$H_O = \mu_1 = \mu_2$
$H_A = \mu_1 \neq \mu_2$

1) Convert
2) Compare   **t-test**

Critical Value

Probability P

Do not reject

Reject Ho

Reject Ho

t

GMCCP
Greater Milwaukee College of Clinical Pharmacy

3 types of t-tests

2-Sample t-test

- Experimental vs control (independent)

Paired t-test

- Pre- vs Post for same person

Single t-test

- Is the sample mean different from the population mean?

2 – Sample t-test

$H_O = \mu_1 = \mu_2$
$H_A = \mu_1 \neq \mu_2$

	Treatment	Control
Glucose readings (mg/dL)	81	86
	81	89
	79	81
	80	87
	84	82
	87	87
	75	85
	83	89
	88	83
Mean, $\bar{x}$		
S_1^2	16.2	9.1
n_1	9	9

Step 1: Convert

$$s_{pooled}^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

$$= \frac{16.2 * 8 + 9.1 * 8}{9 + 9 - 2} = \frac{202.4}{16} = 12.65$$

Effect Size

Variance

$$t_{n_1+n_2-2,\frac{\alpha}{2}} \sim \underline{\qquad\qquad}$$

$$= \frac{82 - 85.6}{\sqrt{s_{pooled} \cdot \left[\frac{1}{9} + \frac{1}{9}\right]}} = -2.15$$

GMCCP
Greater Milwaukee College of Clinical Pharmacy

2-Sample t-test

Step 2: Compare
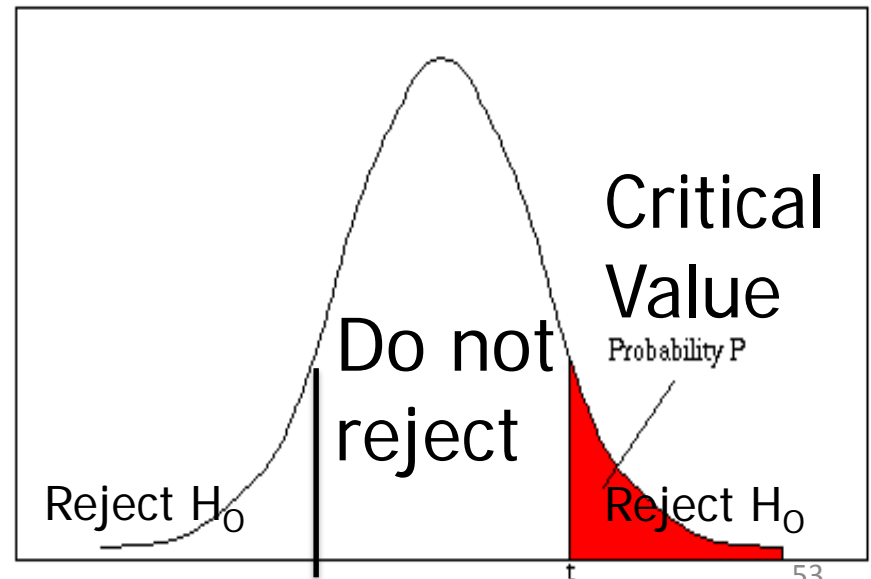
- Is this value significant?

Critical Value: 2.12

Observed Value: -2.15

- Do we care?

Is the change in blood glucose of 3.6 mg/dL important?

Therefore, we reject the null hypothesis that the mean blood glucose between experimental and control group are the same.

Critical Value

Probability P

Do not reject

Reject H_O

Reject H_O

t

ANOVA aka F Test

- Ex:

Group 1: Diet/Education
Group 2: Drug
Group 3: Placebo

Group 1	Group 2	Group 3
280	240	310
250	230	320
270	180	370
280	270	240

H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4 \ldots = \mu_k$
The mean of all groups are the same
H_A: H_o is wrong

Rej
DNR

1) Convert

$df_1 = k - 1$
$df_2 = n - k$
k: # of groups
MSB: means squared between groups, 6400
MSW: means squared within groups, 1488.9

What is n?

What is k?

$$F_{df1,df2} = $$

Ho = wrong

Ho = correct

Calculate F

$$= \frac{MSB}{MSW} = \frac{6400}{1488.9} = 4.29848$$

GMCCP
Greater Milwaukee College of Clinical Pharmacy
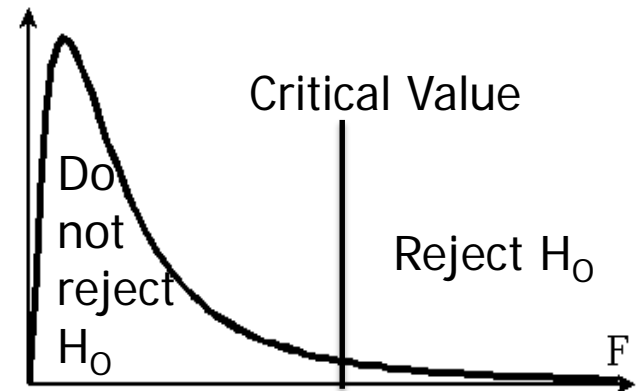
ANOVA Step 2: Compare

Observed value = 4.29848, critical value, F = 4.26
Therefore, reject H_o and one mean is different from another mean in the population.

- Do we care?
We don't know which mean is different among the 3 groups
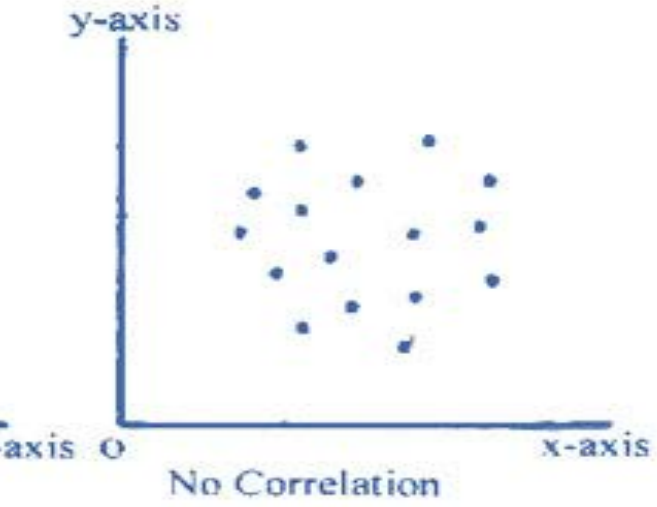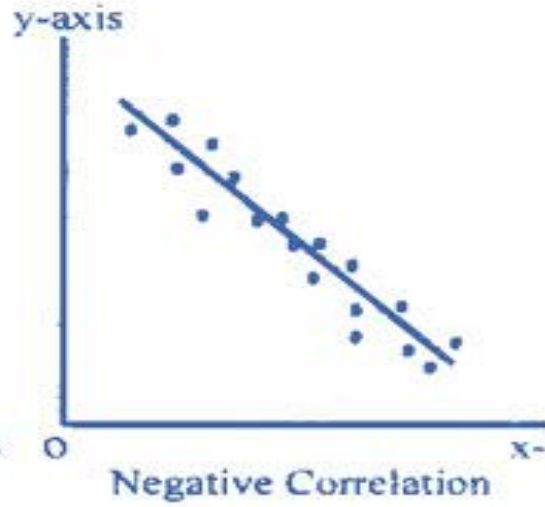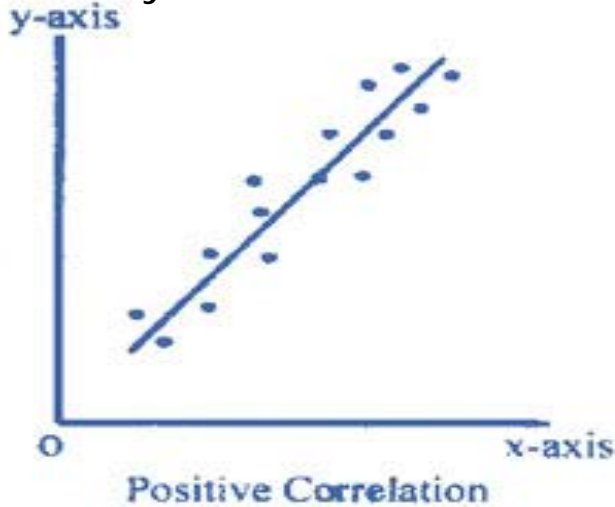1 vs 2, 2 vs 1
1 vs 3, 3 vs 1
2 vs 3, 3 vs 2

Critical Value

Do not reject H_O

Reject H_O

F

Multiple Comparisons Procedure aka post hoc procedure
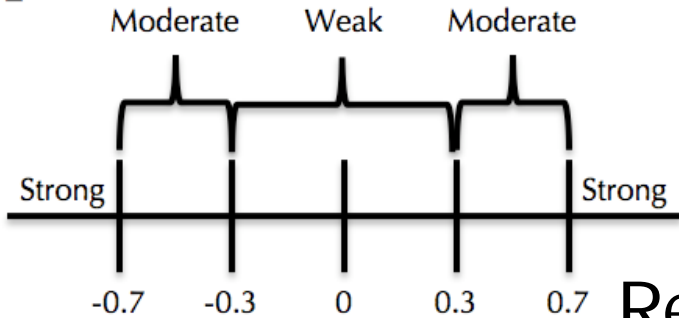
- Bonferonni
- Tukey
- Dunnett

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Correlation

Recall: y = mx + b

Positive Correlation   Negative Correlation   No Correlation

$$r = \frac{\text{covariance}(y, x)}{\sqrt{\text{VAR}_x \cdot \text{VAR}_y}} = \frac{\sum(x_i - \bar{x}) \sum(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

$-1 \leq r \leq 1$

Moderate   Weak   Moderate

Strong   Strong

-0.7   -0.3   0   0.3   0.7

H_O: ρ ≠ 0 (there is no correlation)

Rej   Convert

DNR Compare

$$t_{n-2} \sim \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{\text{effect}}{\text{variance}}$$

GMCCP
Greater Milwaukee College of Clinical Pharmacy

Correlation

- Correlation does not predict causality; it only confirms that a relationship exists.

- Using a scatter plot, it is possible to visualize this definite positive correlation

Regression (Basics)

Analysis Type	Purpose	Analysis Methods
Univariate Analysis	Methods for analyzing data on a single variable	Frequency distribution
Bivariate Analysis	Assess relationship of two variables	Correlation analysis Linear regression
Multivariable Analysis	Assess relationship of multiple variables to a single outcome	Multiple regression Proportional hazards

- Regression: perform typically when baseline demographics are different from each other, generally, includes more than 1 variable

- Linear Regression example:
  - $y = b_o + b_1 x_1 + \ldots + b_k x_k$

Biostats 101 Summary

X: Intervention/ Antecedent (Independent Variable)	Y: Outcome (Dependent Variable)	Statistical Test Used
None	Discrete	Chi Squared
2-level Discrete	Continuous	T-test
K-level Discrete	Continuous	ANOVA Multiple Partial F Test
None	Continuous	Correlation Test
Continuous	Continuous	Multivariate: Linear Regression
Discrete	Discrete	Univariate: Odds Ratio Multivariate: Logistic Regression
Discrete	Time to Event Time is Continuous Event is Binary	Univariate: Kaplan-Meier Multivariate: Cox Proportional Hazards Model

Tip!

Table 2.

Parametric Tests and Some of the More Common Nonparametric Alternatives

Parametric Test	Corresponding Nonparametric Test(s)
Two-sample t test	Mann-Whitney U, median test
Paired t test	Wilcoxon signed rank test, sign test
One-way analysis of variance	Kruskal-Wallis test
Two-way analysis of variance	Friedman two-way analysis of variance
Correlation	Spearman ρ, Kendall τ
Linear regression	Theil's incomplete method

References

- De Muth JE. Overview of biostatistics used in clinical research. *Am J Health-Syst Pharm.* 2009; 66:70 – 81.

- De Muth JE. Preparing for the first meeting with a statistician. *Am J Health-Syst Pharm.* 2008; 65: 2358- 2366.

- Pagano M, Gauvreau K. Principles of Biostatistics (2nd edition). Belmont, CA: Brooks/Cole, Cengage Learning; 2000.

Michael C. Brown, PharmD

Audrey B. Kostrzewa, PharmD, MPH, BCPS

Kajua B. Lor, PharmD, BCACP

PRECEPTING JOURNAL CLUB

Building Your Biostatistics Toolbox for Precepting Journal Clubs

CE must be claimed by 1/27/18 at accp.com

Claiming CE
- Log in to ACCP (https://www.accp.com/) – you will have to register as a guest and create a free account if not an ACCP member
- Click on Education & Meetings
- Click on CPE Center
- Click on Claim CPE for a Chapter Activity
- Enter CE code provided at meeting

- Contact us:
  – Email: secretary.gmccp@gmail.com
  – Web: http://gmccp.weebly.com