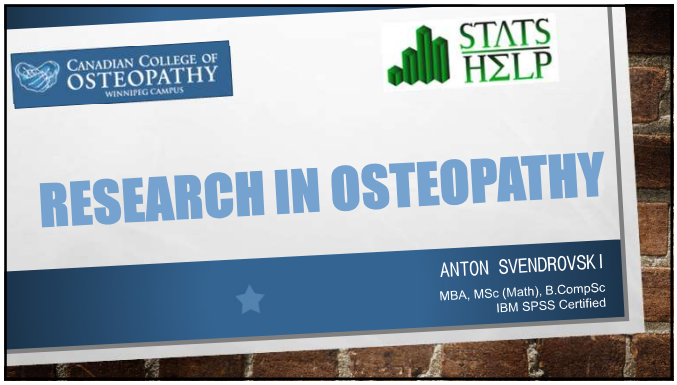


AGENDA

- REVIEW OF MATERIAL
  - HYPOTHESIS/RESEARCH QUESTION
  - P-VALUE
  - STUDY DESIGNS
  - VARIABLES AND MEASUREMENTS
- DESCRIPTIVE STATISTICS
- INFERENTIAL STATISTICS
- SAMPLE SIZE DETERMINATION
- CHOOSING MEASUREMENT INSTRUMENT/TOOL
- QUASI-EXPERIMENTAL DESIGNS
- RELIABILITY STUDIES

REVIEW: RESEARCH QUESTION

1. WHY DO PATIENTS SEEK OSTEOPATHIC TREATMENT?
2. DOES OSTEOPATHIC INTERVENTION X EFFECTIVELY REDUCE PATIENTS' PAIN AFTER 5 SESSIONS?
3. IS THERE AN ASSOCIATION BETWEEN THE AGE OF PARTICIPANTS AND THE NUMBER OF OSTEOPATHIC SESSIONS ATTENDED?
4. IS THERE A DIFFERENCE BETWEEN OSTEOPATIC INTERVENTION X AND INTERVENTION Y IN INCREASING THE PARTICIPANTS' QUALITY OF LIFE?
5. HOW RELIABLE IS A PARTICULAR TECHNIQUE IN DIFFERENTIATING EMPTY VS FILLED BLADDER?
6. IS THERE A CONSENSUS IN PUBLISHED STUDIES REGARDING THE EFFECTIVENESS OF INTERVENTION X?

REVIEW: HYPOTHESIS

Hypothesis = Research Question + Measurement Tool + "$p \leq 0.05$"

Examples of Hypothesis formulation:

1. Osteopathic treatment will significantly reduce the redness associated with acne as measured by infra-red photography, $p \leq 0.05$.
2. Five sessions of osteopathic intervention X will result in significant reduction in patients' pain as measured by Visual Analog Scale, $p \leq 0.05$.
3. Three trained osteopathy students at the end of their curriculum could achieve at least moderate agreement on osteopathic sacral palpatory diagnostic tests, evaluated using Fleiss K (Kappa) statistics, $p \leq 0.05$.
4. Osteopathic treatment X is more effective than osteopathic intervention Y in increasing the participants' quality of life as measured by WHOQOL questionnaire, $p \leq 0.05$.

REVIEW: HYPOTHESES

Null Hypothesis (Ho):

Osteopathic treatment **will NOT** significantly reduce the redness associated with acne as measured by infra-red photography, $p > 0.05$.

Alternative Hypothesis (Ha):

Osteopathic treatment **will** significantly reduce the redness associated with acne as measured by infra-red photography, $p \leq 0.05$.

$p < 0.05$	0.05	$p > 0.05$
Reject null and accept an alternative hypothesis. There is statistically significant reduction of acne skin redness as a result of osteopathic treatment.		Failed to reject the null hypothesis. There is insufficient evidence to conclude that osteopathic treatment is effective.

p-value

UNDERSTANDING RESEARCH ARTICLES

Table 2 Comparison of the VAS, MOV and ROM values between OMT and CCT groups ($n = 25$) at T0, T1 and T2.

		OMT		CCT		t	P
T0	VAS[a]	6.9	±0.88	6.40	±1.42		NS
	MOV[b]	35.1	±4.36	34.9	±34.5		NS
	ROM[c]	62.4	±10.67	64.5	±9.55		NS
T1	VAS[a]	1.5	±0.85	2.6	±0.7	−4.995	0.000
	MOV[b]	46.0	±4.78	41.3	±4.52	3.572	0.000
	ROM[c]	81.9	±10.31	71.9	±9.05	3.654	0.000
T2	VAS[a]	3.8	±1.26	4.4	±1.75		NS
	MOV[b]	42.9	±2.69	40.4	±2.41	3.461	0.001
	ROM[c]	80.5	±5.44	72.4	±2.95	6.545	0.000

[a] The visual analogue pain scale was scored from 0 to 10.
[b] Measure in millimeters.
[c] Measure in degrees.

Source: A.M. Cuccia et al. Osteopathic manual therapy versus conventional conservative therapy in the treatment of temporomandibular disorders: A randomized controlled trial. Journal of Bodywork & Movement Therapies (2010) 14, 179-184
https://pdfs.semanticscholar.org/849d/3c122af15e27b3dc59de93a76dde196e52a4.pdf

UNDERSTANDING RESEARCH ARTICLES

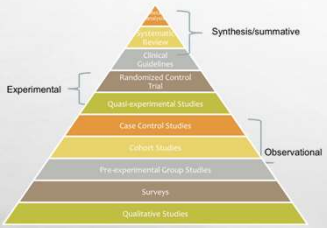
Table 1.—Headache Measures-Comparison of Means and *t*-Tests

	Group	N	Std. Mean	Range Deviation	Range Minimum	Maximum	*t*-Value	Sig. df	2-tailed
HA free days per week									
	C	12	.21 days	1.685 days	−3	4	−2.589	24	.016
	E	14	1.79 days	1.424 days	−1	4			
	Total	26	1.06 days	1.717 days	−3	4			
HA degree of improvement									
	C	12	.656	1.953	−2.58	2.36	−1.860	24	.075
	E	14	1.881	1.394	.00	4.46			
	Total	26			−2.58	4.46			

C = control group; E = experimental group.

Source: Rosemary Anderson & Caryn Seniscal (2006). A comparison of selected osteopathic treatment and relaxation for tension-type headaches. American Headache Society, doi: 10.1111/j.1526-4610.2006.00535.x

REVIEW: STUDY DESIGNS



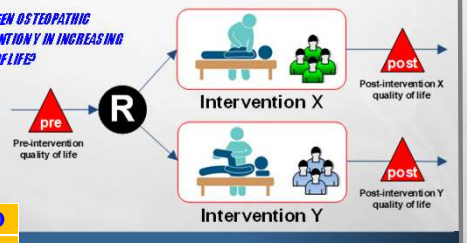
Acceptable Forms of Student Research:

1. Experimental and quasi-experimental research
2. Reliability, validity, and palpation studies
3. Technique studies
4. Case series studies
5. Fundamental studies - investigative but non-experimental
6. Qualitative studies

REVIEW: EXPERIMENTAL (RCT)

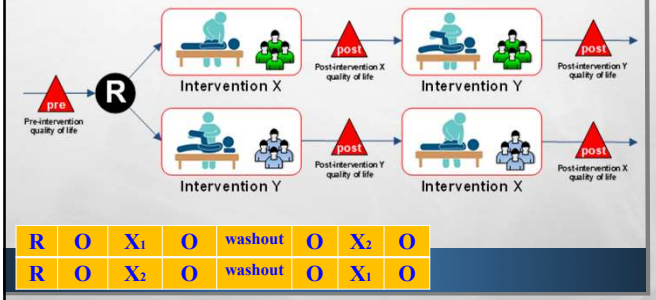
RESEARCH QUESTION:

IS THERE A DIFFERENCE BETWEEN OSTEOPATHIC INTERVENTION X AND INTERVENTION Y IN INCREASING THE PARTICIPANTS' QUALITY OF LIFE?



R	O	X_1	O
R	O	X_2	O

REVIEW: QUASI-EXPERIMENTAL (CROSSOVER)



R	O	X_1	O	washout	O	X_2	O
R	O	X_2	O	washout	O	X_1	O

REVIEW: QUASI-EXPERIMENTAL (WITHIN SUBJECT)

RESEARCH QUESTION:

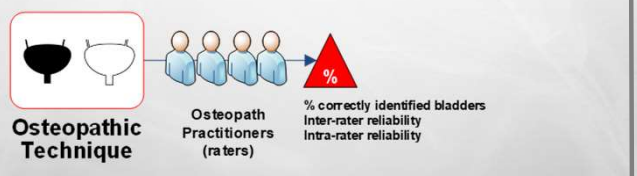
DOES OSTEOPATHIC INTERVENTION X EFFECTIVELY REDUCE PATIENTS' PAIN AFTER 5 SESSIONS?

O	X	O

REVIEW: RELIABILITY STUDY

RESEARCH QUESTION:

HOW RELIABLE IS A PARTICULAR TECHNIQUE IN DIFFERENTIATING EMPTY VS FILLED BLADDER?



REVIEW: VARIABLES

Variable is a thing that changes in experiment. A variable is any factor, trait, or condition that can exist in differing amounts or types.

Independent Variable – the variable that is changed or controlled in a scientific experiment. Usually the Treatment: technique, global or regional osteopathic intervention vs control.

Dependent Variable – the outcome of interest, what we are hoping to change or alter.

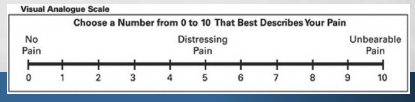
Variable type: **Numerical** (Age) or **Categorical** (Gender, Group)

REVIEW: MEASUREMENTS

Measurement is a variable that is being assessed (quantified / measured) using a particular technique, tool or instrument.

$$BMI = \frac{\text{weight (kg)}}{\text{height (m}^2\text{)}}$$

In the last 4 weeks have you	No Problem	Slight Problem	Moderate Problem	Marked Problem	Extreme Problem
1. Had difficulty moving?	☐	☐	☐	☐	☐
2. Had difficulty walking?	☐	☐	☐	☐	☐
3. Had problems with your balance?	☐	☐	☐	☐	☐
4. Had difficulty standing up without support?	☐	☐	☐	☐	☐
5. Had difficulty speaking?	☐	☐	☐	☐	☐



TWO AREAS OF STATISTICS

DESCRIPTIVE statistics

- SUMMARIZE SAMPLE DATA
  - MEAN, MEDIAN, MODE
  - STANDARD DEVIATION, RANGE
  - FREQUENCY, PROPORTIONS (%)
- VISUALIZE DATA IN A SAMPLE
  - HISTOGRAM
  - BAR GRAPH
  - BOX-WHISKER PLOT

INFERENTIAL statistics

- INFER/GENERALIZE RESULTS TO THE TARGET POPULATION
- CONFIDENCE INTERVALS (95% CI)
- STATISTICAL TESTS (P-VALUE)
  - PARAMETRIC VS NON-PARAMETRIC
- TYPE I AND TYPE II ERRORS

DESCRIPTIVE STATISTICS

Table 1
Demographic Characteristics of Subjects (N = 14)*

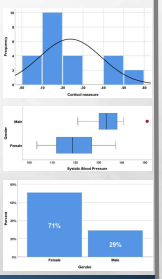
Variable	Treatment Group (n = 7)	Control Group (n = 7)
Age, y	82.3 ± 4.7	84.3 ± 5.6
Gender, No. (%)		
Female	6 (86)	6 (86)
Male	1 (14)	1 (14)
Race/ethnicity, No. (%)		
White	7 (100)	7 (100)
Received flu shot in past year, No. (%)		
Yes	6 (86)	7 (100)
Not known	1 (14)	0 (0)
History of dementia, No. (%)		
Yes	1 (14)	1 (17)
No	6 (86)	5 (83)

*Data are presented as mean ± SD for continuous variables and as number (percent of group) for categorical variables. Total counts may not add up to n due to missing values.

MEASURES OF CENTRAL TENDENCY
- MEAN = AVERAGE
- MEDIAN = 50/50 CUT-OFF
- MODE = MOST FREQUENT

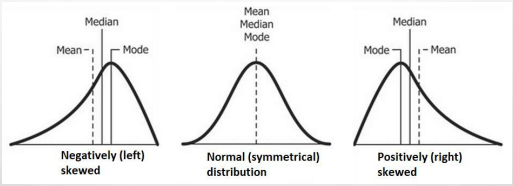
MEASURES OF VARIABILITY
- STANDARD DEVIATION
- IQR OR RANGE

CATEGORICAL (QUALITATIVE) DATA
- FREQUENCY
- PROPORTIONS (%)



Reference: Donald R. Noll, Brian F. Degenhardt, Melissa Stuart, Rene McGovern & Michelle Matteson (2004). Effectiveness of a Sham Protocol and Adverse Effects in a Clinical Trial of Osteopathic Manipulative Treatment in Nursing Home Patients. JAOA vol 104 (3).

NORMAL DISTRIBUTION

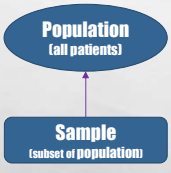


ASSESSED BY HISTOGRAMS AND COMPARING MEAN AND MEDIAN

NORMAL DISTRIBUTION IS DESIRED FOR (PARAMETRIC) STATISTICAL ANALYSIS

INFERENTIAL STATISTICS

HELPS US TO INFER AND GENERALIZE THE FINDINGS IN A SAMPLE (INDIVIDUAL STUDY) TO THE ENTIRE POPULATION

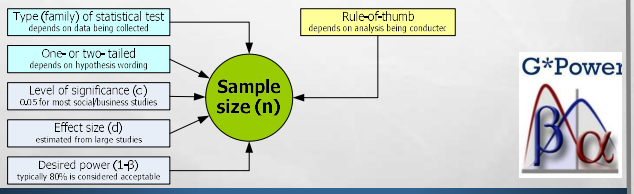


1) CONFIDENCE INTERVALS (CI)
- ESTIMATE POPULATION PROPORTION
- ESTIMATE POPULATION MEAN

2) STATISTICAL HYPOTHESIS TESTS
- EVALUATE (SAMPLE) EVIDENCE TO MAKE CONCLUSION ABOUT UNKNOWN POPULATION CHARACTERISTIC
- COURTROOM EXAMPLE: NULL HYPOTHESIS = NOT GUILTY, ALTERNATIVE HYPOTHESIS = GUILTY

SAMPLE SIZE DETERMINATION

Level of significance (Type I error) – chance of finding effect if it does not exist
Effect size – expected amount of change in dependent variable (treatment effect)
Statistical power – credibility of the test, chance of finding effect if it does exist



http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register

HOW DO I KNOW EFFECT SIZE?

Approaches to determine effect size:

- Previous (published) studies with similar research question
  - similar Population, Intervention, Outcome
  - look for numbers to quantify effect size (mean, standard deviation, %)
- Pilot study conducted with small group of participants (*n* = 12)
- Based on practical significance
  - Clinically important change, Minimal Important Difference (MID)
- Assume to be medium effect (*Cohen's d* = 0.5)

FINDING PUBLISHED STUDIES

- **GOOGLE SEARCH (START WITH GOOGLE SCHOLAR)**
- **PREVIOUS YEARS CCO STUDENTS' THESIS**
- **THE JOURNAL OF THE AMERICAN OSTEOPATHIC ASSOCIATION**
  HTTP://JAOA.ORG/
- **INTERNATIONAL JOURNAL OF OSTEOPATHIC MEDICINE**
  HTTP://WWW.JOURNALOFOSTEOPATHICMEDICINE.COM/
- **THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE**
  HTTPS://WWW.LIEBERTPUB.COM/LOI/ACM
- **INTERNATIONAL JOURNAL OF OCCUPATIONAL MEDICINE AND ENVIRONMENTAL HEALTH**
  HTTP://IJOMEH.EU/
- **INTERNATIONAL JOURNAL OF PHYSIOTHERAPY**
  HTTPS://WWW.IJPHY.ORG/

Google Scholar

CANADIAN COLLEGE OF OSTEOPATHY

SAMPLE SIZE – RULES-OF-THUMB

Experimental:	**Minimum 12**
Quasi-Experimental:	**Minimum 16**
Reliability Studies:	**Minimum 40**
Technique Studies:	**Minimum 24**
Case Series:	**Minimum 6**

Final notes on sample size:

- For multiple groups, aim for **balanced** design (equal number of participants in each group).
- Account for **non-response rate** during recruitment.
- Account for **attrition/drop-out rate** during the study.

PILOT STUDIES / PRE-STUDIES

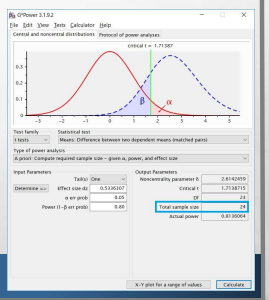
Pre-study is a *small (preliminary)* study undertaken before *large* one.

- Applicable when no previous studies are available on the research topic
- **Feasibility** assessment to validate
  - study design and research protocol
  - subjects recruitment strategy, consent rate, dropout rate
  - treatment, intervention
  - outcome measures, instruments, measurement/assessment tools
- Helpful to explore the effect size and determine sample size needed for a large study
- Recommendations for future large-scale study

SAMPLE SIZE DETERMINATION EXAMPLE

Research Question:
A global osteopathic treatment will increase urinary pH levels, as measured using urine test strips.

- crossover design
- "increase" → one-tail test
- literature search → Buscemi et al. (2015) study reported effect size
- G*Power calculation → 24 subjects
- 10% dropout rate → 27 subj to recruit



Reference: Buscemi, A., Carbone, J., Tacchi, M., Buttafuoco, S., Rapisarda, A., Perciavalle, V., & Coco, M. (2015). Changes of urine pH after the compression of the fourth ventricle. Medicina, Ricerche, Scienza della vita, Retrieved from http://www.scienza-ricerche.it/

STUDENTS' RESEARCH

- Proposal (PICO statement)
  - P = patient/problem (research question)
  - I = intervention (experiment design)
  - C = comparison (control)
  - O = outcome (validated instrument to measure)

MEASUREMENT INSTRUMENT/TOOL

Examples:

Strain	→ **Strain gauge**
Angle	→ **Goniometer** (manual or digital)
Acceleration (3-axis)	→ **Accelerometer** (Fitbit or less expensive alternatives)
Ground reaction force	→ **Force platform/plate**
Object thickness	→ **Caliper**
Time interval	→ **Stopwatch** (iPhone has one built-in)
Weight	→ **Scale**

Clinical measurements (pulse, blood pressure, temperature, respiratory rate)

Ensure sufficient level of **accuracy/precision** and **range**

MEASUREMENT INSTRUMENT/TOOL

Google Scholar

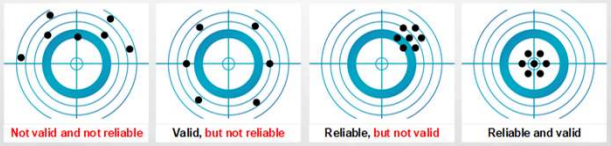
CANADIAN COLLEGE OF OSTEOPATHY TORONTO CAMPUS

Examples:

Tinnitus symptoms	→ **Tinnitus Handicap Inventory** (THI)
Quality of life	→ **Quality of Life Scale** (QOLS) questionnaire
Pain	→ **Visual Analog Scale** (VAS)
Feet functioning	→ **Foot and Ankle Survey** (FAOS) or **Foot Functioning Index** (FFI)

Good instrument is both **Reliable** and **Valid** (validated).

INSTRUMENT RELIABILITY AND VALIDITY



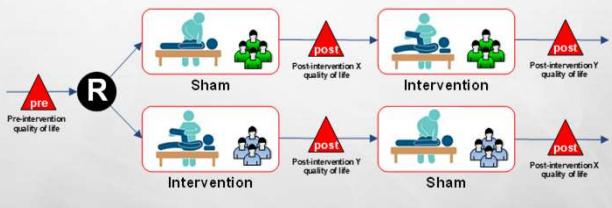
INSTRUMENT RELIABILITY AND VALIDITY

Reliability:
Internal consistency reliability (Cronbach's α > 0.8)
Test-retest reliability correlation (r > 0.7)
Inter-rater (inter-observer) reliability (Kappa > 0.4 or interclass correlation coefficient > 0.7)

Validity:
Correlation with "gold standard" instrument (r > 0.7)
Overall accuracy with respect to actual state (diagnostic accuracy, sensitivity, specificity, PPV, NPV)

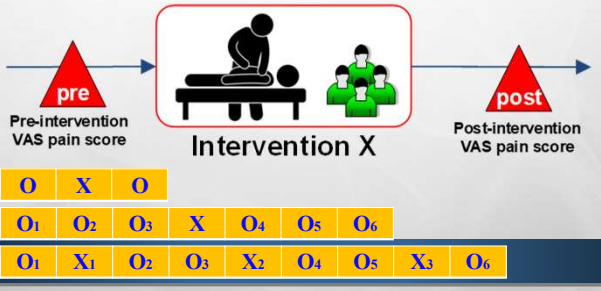
< 0	0.00-0.20	0.21-0.40	0.41-0.60	0.61-0.80	0.81-1.00
Poor	Slight	Fair	Moderate	Substantial	Almost perfect

QUASI-EXPERIMENTAL (CROSSOVER)



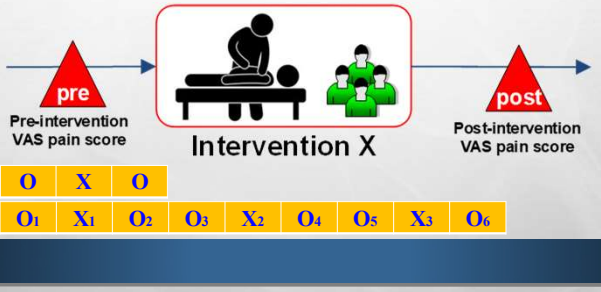
R	O		O	washout	O	X	O
R	O	X	O	washout	O		O

QUASI-EXPERIMENTAL (WITHIN SUBJECT)



O	X	O						
O_1	O_2	O_3	X	O_4	O_5	O_6		
O_1	X_1	O_2	O_3	X_2	O_4	O_5	X_3	O_6

CASE SERIES

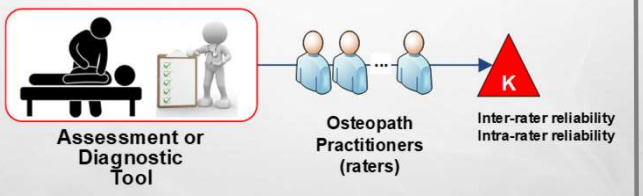


O	X	O						
O_1	X_1	O_2	O_3	X_2	O_4	O_5	X_3	O_6

RELIABILITY/VALIDITY/PALPATION STUDIES

- Practical aspects
  - Live patients or objects (models)
  - Repeated trials to make a diagnosis
- Benefits
  - Relative simplicity in design
  - Contribution to osteopathic profession
  - Improving manual skills
  - Osteopathic students as study participants

RELIABILITY STUDY EXAMPLE

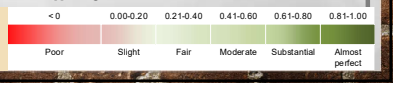


Example:

Consorti et al. (2017) study explored inter-rater reliability of Osteopathic Sacral Palpatory Diagnostic Test using 52 patients and 3 trained osteopathy students (raters). Fleiss Kappa ranges between 0.06 to 0.34 (Table 3).

Categorical outcomes:
Cohen's Kappa (2 raters), Fleiss Kappa (3+ raters)

Numerical outcomes:
Cronbach's α, Interclass Correlation Coefficient



VALIDITY STUDY EXAMPLE

Examples:

- **Assessing accuracy of palpation technique to differentiate between empty and filled bladders**
- **Using wax blocks to assess participants' skills in differentiating two heights (Christopher Reiach study)**
- **Evaluating palpation technique to determine knee problems (validate through radiographs)**
- **Palpation sensitivity study using a hydrodynamic model (Monica Noy project)**

Categorical outcomes:
Overall accuracy, sensitivity, specificity, NPV, PPV

Numerical outcomes:
Correlation coefficient, mean absolute error

PALPATION STUDY EXAMPLE

Intervention examples:

- **Feedback when using wax blocks**
- **Take home models to self-practice palpation skills**
- **Workshops with group practice sessions**

TRAINING STATION FOR SURGEONS

Presented with the permission of Dr. Ilay Habaz and Dr. Eran Shlomovitz (University Health Network)

STUDENTS' RESEARCH – PARTICIPANTS

Recruitment of study participants

- Specialized clinics
- Osteopathic practices
- Social media (Facebook, LinkedIn, Twitter)
  - Post message on your own page
  - Ask friends to re-post your message on their pages
  - Join relevant Facebook group
  - Paid advertisement
- Kijiji and other online posting sites

QUESTIONS? COMMENTS? THOUGHTS?

- **ANTON SVENDROVSKI**
- **647-833-3359**
- **WWW.STATSHELP.CA/OSTEO/**
- **INFO@STATSHELP.CA**

STATS HELP

Research Proposals | Sample Size Calculation | Methodology/Design | Statistical Data Analysis | Interpretation