NBE-4070 : Basics of Biomedical Data Analysis

Stéphane Deny
Prof. in Neuroscience and Biomedial Engineering and Computer Science
Aalto University

Lecture 1: Mean, Standard Deviation, Standard Error, Confidence Intervals, T-tests

What is this course about?

- It’s about learning how to analyze biomedical data.

- More precisely: it covers some of the essential notions needed to analyse time-series and other multi-dimensional data, including

  - basic statistical concepts (e.g., mean, standard error, t-tests),

  - spectral methods (e.g., Fourier transform, high-pass/low-pass filters),

  - linear methods (e.g., PCA, linear regression),

  - non-linear methods (e.g., clustering, ICA, t-SNE, deep networks)

What are biomedical signals? Examples below:

- **Bioelectric signals** : EEG (brain), ECG (heart), EMG (muscles) electrical activity,..., using electrodes
- **Biomagnetic signals** : brain, heart, lungs magnetic activity
- **Bioimpedance signals** : tissue composition, blood distribution using electrodes
- **Bioacoustic signals** : sounds from blood flow, lungs, digestive tract, joints, muscles using microphones, also – speech sound
- **Biomechanical signals** : blood pressure, limb acceleration, flow tension using diversity of sensors
- **Biochemical signals** : laboratory analysis of blood/tissue samples
- **Biooptical signals** : blood oxygenation, heart output using fiberoptic technology, heart rate from wrist device
- **'others'** : observations made by nurse, annotations alarms, (self)assessments like levels of ‘pain’, ‘tiredness’ etc.

Learning Objectives

After completing this course, you will:

- understand the fundamental linear and non-linear methods used for biomedical data analysis, and know how to explain them to a general audience.

- Given a question and biomedical measurement, you will know how to select and apply the suitable data analysis methods for this problem.

- have all the keys to avoid overinterpreting or misinterpreting the results of different analyses.

Teaching team

- **Lecturer:**
  Stéphane Deny: stephane.deny@aalto.fi
  Brief bio: - PhD in Neuroscience at UPMC (France),
  - Postdoc Neuro & ML at Stanford and Facebook AI
  - Asst. Prof. in Neuro and CS at Aalto since 2021
  Second time teaching this course:
  => feedback welcome about the lectures and exercise sessions!
  https://presemo.aalto.fi/bda2023

- **Teaching Assistants**
  Hyunkyung Choo: hyunkyung.choo@aalto.fi
  Carlos Sevilla Salcedo: carlos.sevillasalcedo@aalto.fi

- **Ask us your questions by e-mail (but not at the last minute)!**

Course Material

- **Slides**
  - red mark on the slide : things to know
  - orange mark on the slide : things to understand
  - blue mark on the slide : examples, food for thought

- **Exercises:**
  - Python notebooks containing exercises

Calendar

- 11 lectures from September to November

- 11 exercise sessions, in 2 groups:
  - both groups on Mondays 10.15 – 12.00
  - Sign up on OpenCourse to one of these groups
  - You can work together or alone during exercise sessions, but everyone should upload the completed exercise.

- Oral Exam: Dec 12 - 13

Monday	*Tuesday*	*Wednesday*	*Thursday*	*Friday*
11 H01 H02	12 L01	13	14	15

Grading – 5 credits

- Quizz at the beginning of most lectures (20% of the grade). Study the slides of the previous lecture before coming to class.

- Exercise sessions every week (50% of the grade), on the computer. Exercise notebooks will be posted online. You need to upload them completed according to the deadlines.

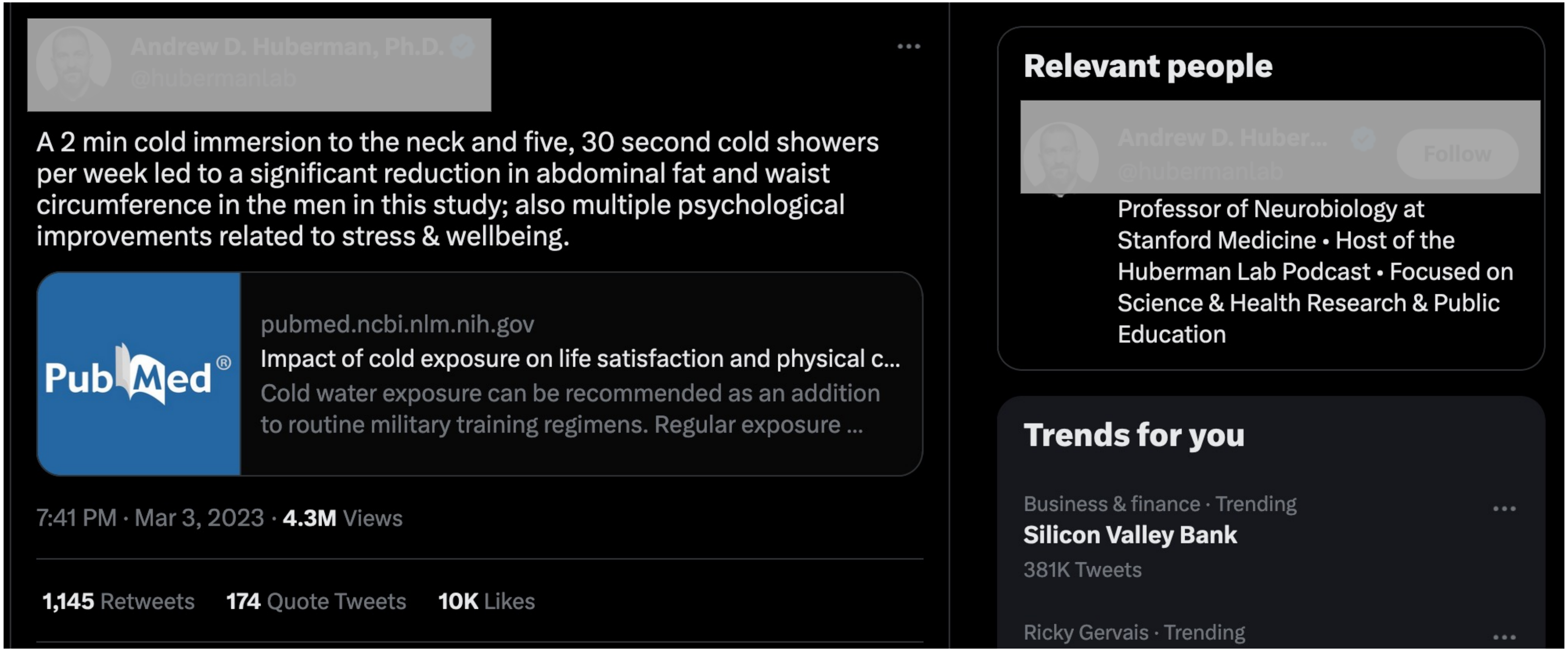
- Oral examination (30%): During a 20-min examination, you will be asked to reason about a data analysis problem, select and apply the correct data analysis technique, and know how to explain this technique.

50% is a pass, 80% is an excellent grade

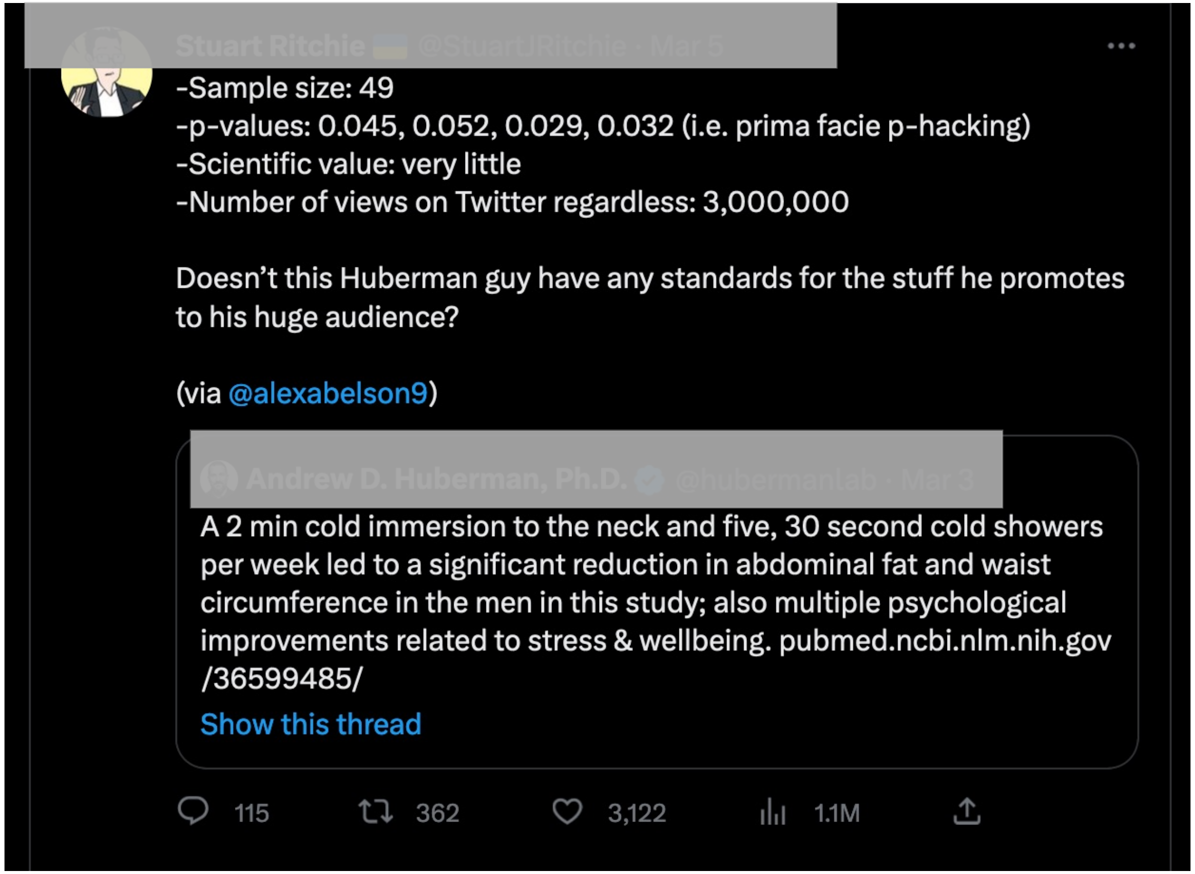
Outline of the course

1. Mean, Standard Deviation, Standard Error, Confidence Intervals, T-test
2. Fourier Transform, Wavelet Transforms, Spectrograms, High-pass, Low-pass filters
3. Principal Component Analysis / Singular Value Decomposition
4. Linear Regression / Logistic Regression
5. Clustering Methods
6. Non-linear Methods: Independant Component Analysis, t-Stochastic Neighbour Embedding, Random Forests, Deep Networks
7. Invited lecture from the biomedical industry

- In what way(s) do you think these claims could be wrong or exaggerated?
- What steps would you take to verify these claims?



Checking the comments



[redacted]

-Sample size: 49
-p-values: 0.045, 0.052, 0.029, 0.032 (i.e. prima facie p-hacking)
-Scientific value: very little
-Number of views on Twitter regardless: 3,000,000

Doesn't this Huberman guy have any standards for the stuff he promotes to his huge audience?

(via @alexabelson9)

[redacted]

A 2 min cold immersion to the neck and five, 30 second cold showers per week led to a significant reduction in abdominal fat and waist circumference in the men in this study; also multiple psychological improvements related to stress & wellbeing. pubmed.ncbi.nlm.nih.gov/36599485/

Show this thread

115 362 3,122 1.1M

Reading the abstract

The claims in the tweet are consistent with the claims of the abtract.

Andrew D. Huberman, Ph.D. @hubermanlab · Mar 3

A 2 min cold immersion to the neck and five, 30 second cold showers per week led to a significant reduction in abdominal fat and waist circumference in the men in this study; also multiple psychological improvements related to stress & wellbeing. pubmed.ncbi.nlm.nih.gov/36599485/

Show this thread

ABSTRACT

Introduction Optimal mental state and physical fitness are crucial factors affecting training of military personnel. Incorporating components aimed at improving mental status and physical composition of soldiers into training programmes can lead to better outcomes. Previously, cold exposure has been used to promote human health in multiple ways, with a plethora of reported benefits. Thus, the aim of this study was to examine the effect of regular cold exposure on the psychological status and physical composition of healthy young soldiers in the Czech Army.

Methods A total of 49 (male and female) soldiers aged 19–30 years were randomly assigned to one of the two groups (intervention and control). The participants regularly underwent cold exposure for 8 weeks, in outdoor and indoor environments. Life Satisfaction Questionnaire and InBody 770 device were used to evaluate life satisfaction and body composition, respectively. Zung Self-Rating Anxiety Scale was used to assess anxiety produced by cold exposure.

Results Theoretical and practical training in cold immersion in the winter did not induce anxiety. Regular cold exposure led to a significant ($p=0.045$) increase of 6.2% in self-perceived sexual satisfaction compared with the pre-exposure measurements. Furthermore, considerable increase (6.3% compared with the pre-exposure period) was observed in self-perceived health satisfaction; the change was borderline significant ($p=0.052$). In men, there was a reduction in waist circumference (1.3%, $p=0.029$) and abdominal fat (5.5%, $p=0.042$). Systematic exposure to cold significantly lowered perceived anxiety in the entire test group ($p=0.032$).

Conclusions Cold water exposure can be recommended as an addition to routine military training regimens. Regular exposure positively impacts mental status and physical composition, which may contribute to the higher psychological resilience. Additionally, cold exposure as a part of military training is most likely to reduce anxiety among soldiers.

Accessing the study.

BMJ Military Health

Original research

Impact of cold exposure on life satisfaction and physical composition of soldiers

Jiří Néma,[1] J Zdara,[1] P Lašák,[2] J Bavlovič,[3] M Bureš,[4] J Pejchal,[5] H Schvach[1]

[1]Department of Military Medical Services Organisation and Management, University of Defence in Brno Faculty of Military Health Sciences, Hradec Kralove, Czech Republic
[2]Department of Military Internal Medicine and Military Hygiene, University of Defence in Brno Faculty of Military Health Sciences, Hradec Kralove, Czech Republic
[3]Department of Molecular Pathology and Biology, University of Defence in Brno Faculty of Military Health Sciences, Hradec Kralove, Czech Republic
[4]Department of Computer Science, Czech Technical University in Prague, Prague, Czech Republic
[5]Department of Toxicology and Military Pharmacy, University of Defence in Brno Faculty of Military Health Sciences, Hradec Kralove, Czech Republic

Correspondence to
J Bavlovič, Department of Molecular Pathology and Biology, University of Defence in Brno Faculty of Military Health Sciences, Hradec Kralove 50001, Czech Republic; jan.bavlovic@unob.cz

Received 30 August 2022
Accepted 7 December 2022

ABSTRACT

Introduction Optimal mental state and physical fitness are crucial factors affecting training of military personnel. Incorporating components aimed at improving mental status and physical composition of soldiers into training programmes can lead to better outcomes. Previously, cold exposure has been used to promote human health in multiple ways, with a plethora of reported benefits. Thus, the aim of this study was to examine the effect of regular cold exposure on the psychological status and physical composition of healthy young soldiers in the Czech Army.

Methods A total of 49 (male and female) soldiers aged 19–30 years were randomly assigned to one of the two groups (intervention and control). The participants regularly underwent cold exposure for 8 weeks, in outdoor and indoor environments. Life Satisfaction Questionnaire and InBody 770 device were used to evaluate life satisfaction and body composition, respectively. Zung Self-Rating Anxiety Scale was used to assess anxiety produced by cold exposure.

Results Theoretical and practical training in cold immersion in the winter did not induce anxiety. Regular cold exposure led to a significant (p=0.045) increase of 6.2% in self-perceived sexual satisfaction compared with the pre-exposure measurements. Furthermore, considerable increase (6.3% compared with the pre-exposure period) was observed in self-perceived health satisfaction; the change was borderline significant (p=0.052). In men, there was a reduction in waist circumference (1.3%, p=0.029) and abdominal fat (5.5%, p=0.042). Systematic exposure to cold significantly lowered perceived anxiety in the entire test group (p=0.032).

Conclusions Cold water exposure can be recommended as an addition to routine military training regimens. Regular exposure positively impacts mental status and physical composition, which may contribute to the higher psychological resilience. Additionally, cold exposure as a part of military training is most likely to reduce anxiety among soldiers.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Regular cold exposure increases resilience to stress factors.

WHAT THIS STUDY ADDS

⇒ Theoretical and practical training in cold immersion in the winter does not induce anxiety.

⇒ Regular exposure to cold increases self-perceived health and sexual satisfaction and reduces waist circumference and abdominal fat.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Cold water exposure as a supplement to routine military training may be a pleasant diversion while positively affecting the well-being of soldiers.

that cold exposure is associated with a reduction in morbidity. However, the number of days of illness in adults remained unchanged (ie, if individuals become ill, they will be treated for the same length of time as they would have been treated if not exposed to cold).[3] In the past, cold showers or cold water swimming has been used as a possible non-pharmacological supportive treatment for depression, but without satisfactory results.[3 4] Regular cold exposure has been shown to increase resilience to stress factors[5] and improve psychological well-being.[6] It is also known to exert a positive effect on accelerating recovery after high physical stress.[7] Furthermore, cold exposure has been demonstrated to lower BP, suggesting its utility as a preventive measure against cardiovascular diseases.[8–10] At the same time, short-term exposure to cold (up to 1 hour in a cold bath, under optimal cold exposure conditions as defined in the literature) appears to be

Result table

Table 2 Change in perceived anxiety before and after the theoretical training and first immersion into the cold lake water during the winter measured by Zung SAS (mean±SD)

	1st measurement	2nd measurement	P value
Intervention group (N=25)	41.1±11.0	39.3±11.2	**0.032 (0.031)**
Control group (N=24)	40.9±9.3	41.8±8.2	0.721 (0.730)

Bold significant differences between the first and second measurements: p≤0.05 (tested by Wilcoxon test; in brackets, we state the exact test as Monte Carlo).
SAS, Self-Rating Anxiety Scale.

Definitions: mean

- Given a dataset of N measurements $X = \{x_1, x_2, ..., x_N\}$

- The ***mean*** of this dataset is given by:

$$\mu = \frac{x_1 + x_2 + ... + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Definitions: variance and standard deviation

- Given a dataset of N measurements $X = \{x_1, x_2, ..., x_N\}$

- The <u>mean</u> is given by: $\mu = \frac{1}{N} * \sum_{i=1}^{N} x_i$

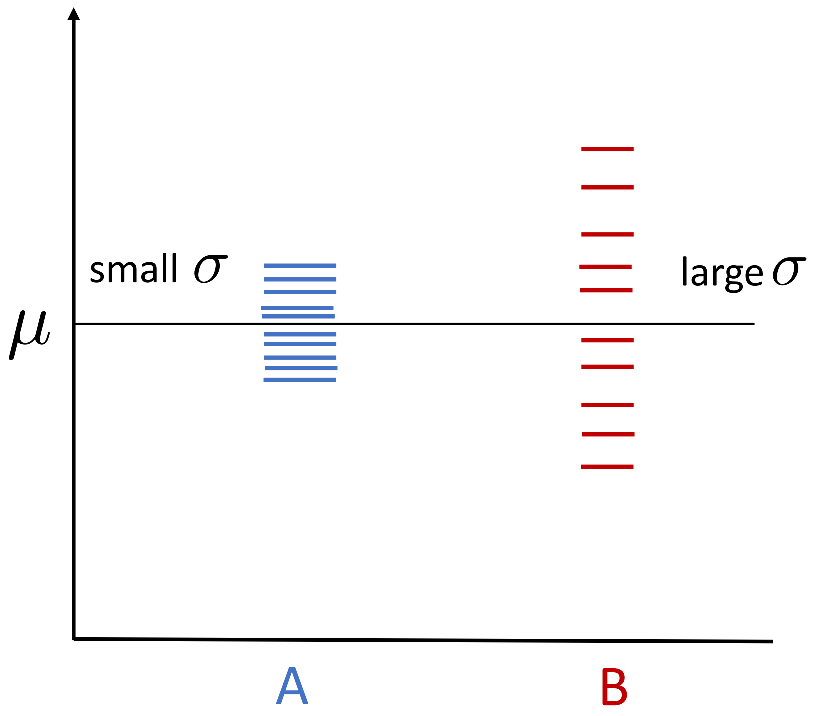
- The <u>*variance*</u> is given by: $Var(X) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2$

- The <u>*standard deviation*</u> is given by: $\sigma(X) = \sqrt{Var(X)}$

The standard deviation describes the spread of the data around the mean.

Vizualisation of standard deviation

A and B have same mean
but different standard deviations.



Definitions: standard error

- Given a dataset of N measurements $X = \{x_1, x_2, ..., x_N\}$

- The standard deviation is given by: $$\sigma(X) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

It describes the spread of the data around the mean.

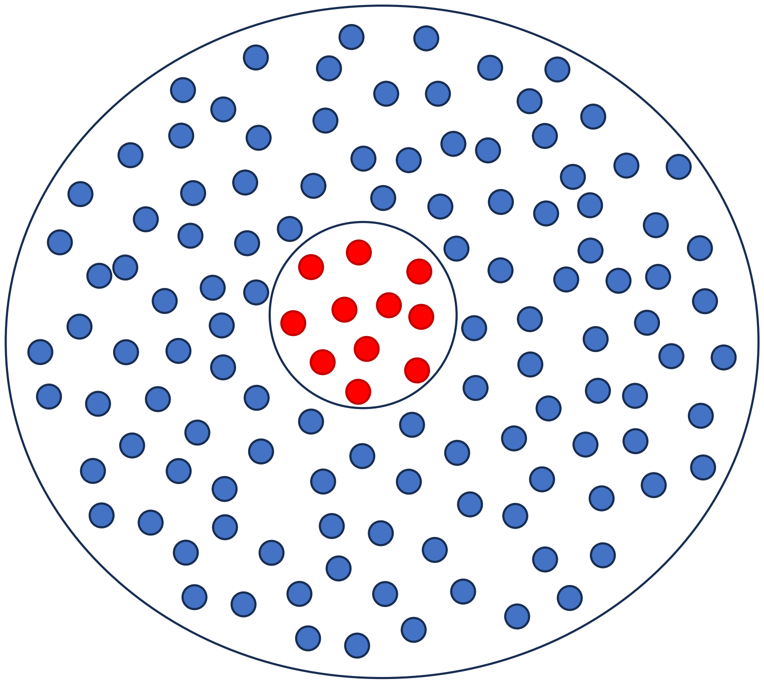
- The ***standard error (of the mean)*** is given by: $$SE = \frac{\sigma(X)}{\sqrt{N}}$$

It describes how variable the mean would be, if the measurement was repeated many times.

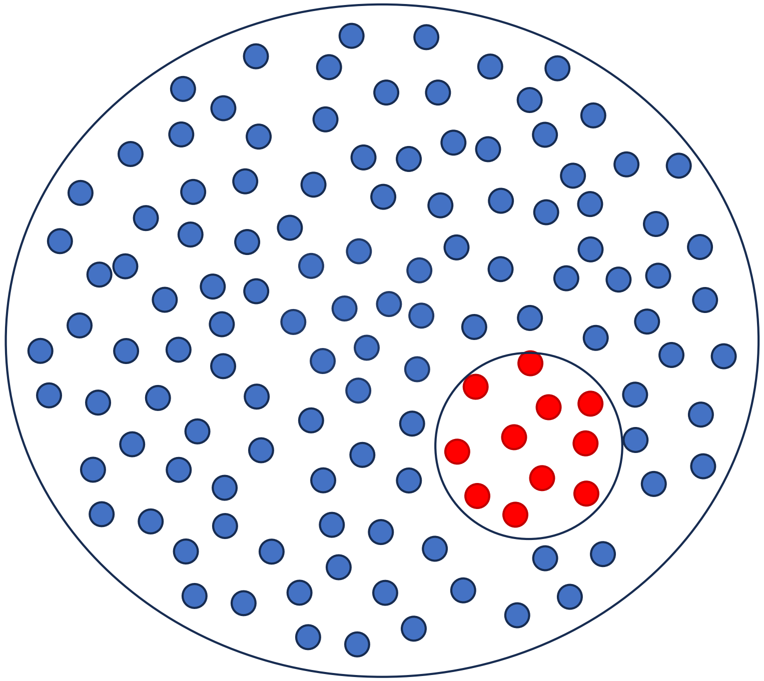
Explanation of the standard error of the mean

The whole population has true mean μ

The sampled population has mean $\hat{\mu}$



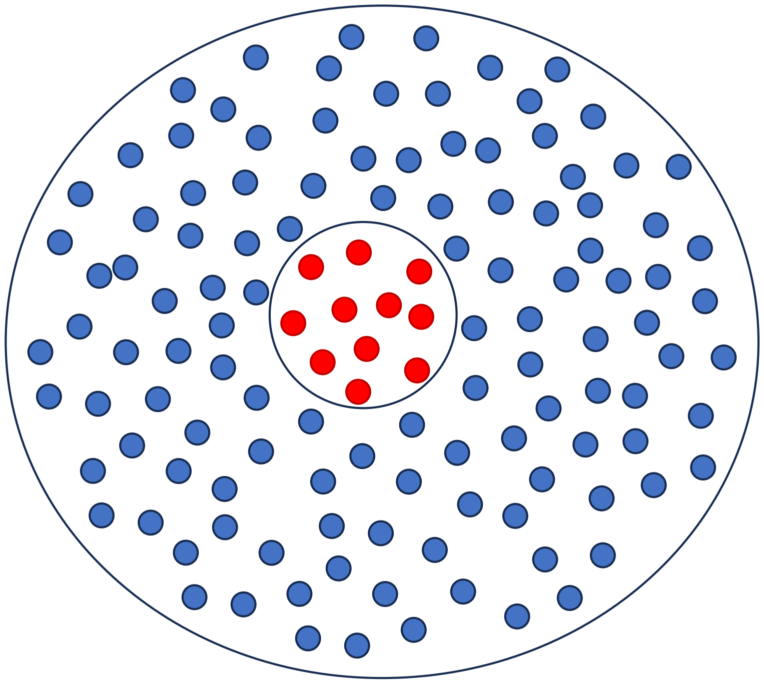
Now imagine repeating the measurement many times, on random samples of the population:



Explanation of the standard error of the mean

The whole population has true mean μ

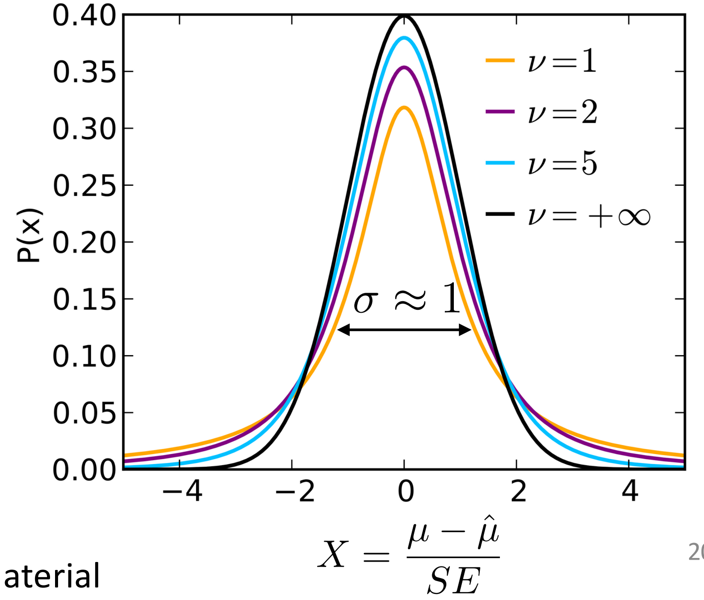
The sampled population has mean $\hat{\mu}$



The difference between true and estimated mean, normalized by the standard error, $\frac{\mu - \hat{\mu}}{SE}$

follows the Student's t-distribution*:

ν : degrees of freedom (N-1)



* See calculations in supplementary material

Beer and Statistics

William Sealy Gosset (1876–1937)

Famous as a statistician, best known by his pen name *Student* and for his work on Student's t-distribution.

Guiness had a blanket ban on publications.
So Gosset used his pen name Student.

Definition: confidence interval

- Given a dataset of N measurements $X = \{x_1, x_2, ..., x_N\}$

- A ***confidence interval*** is a range of value you can be xx% confident contains the true mean of the population. It is given by:

$$CI = \mu \pm t_{N-1} * SE$$

where t_{N-1} is a value that depends on the number of samples and the % confidence desired.

- If N>10 and we are looking for the 95% confidence interval, $t_{N-1} \approx 2$

How to find t_{N-1} ? Use a Student t-distribution abacus

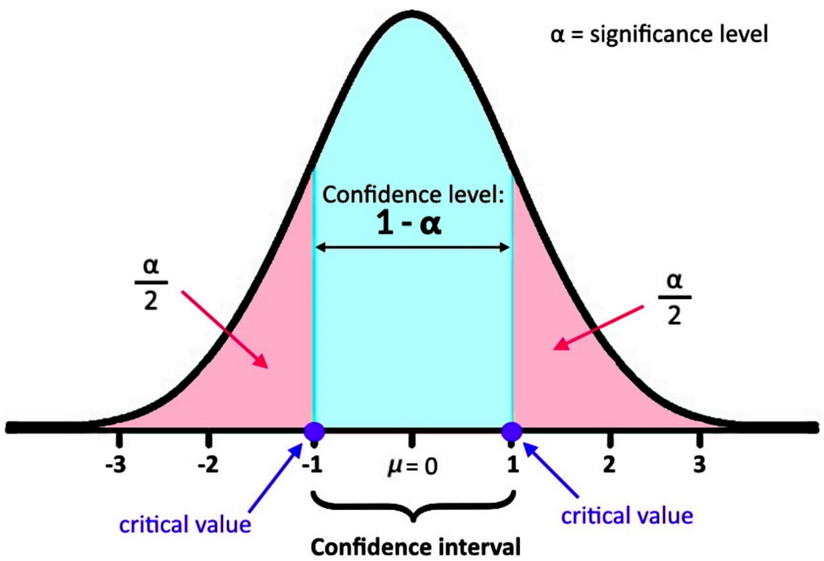
desired significance level

degrees of freedom (N-1)

ν \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	[illegible]	2.583	2.921	3.686	4.015
17	1.333	1.740	[illegible]	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460

t_{N-1}

Student t-distribution



Recap: standard deviation (SD), standard error of the mean (SE) and confidence interval (CI)

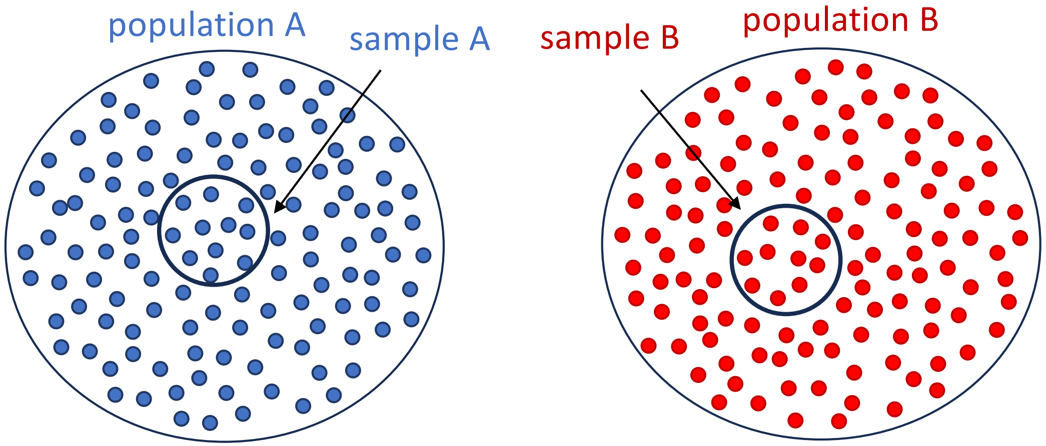
Table I. **Common error bars**

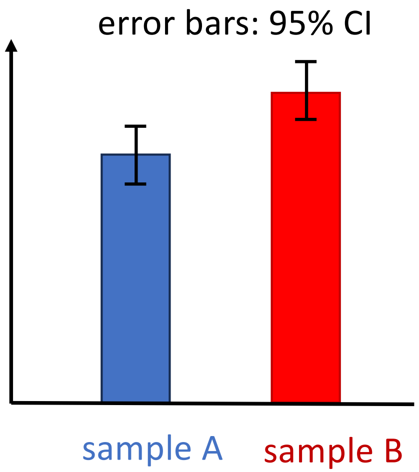
Error bar	Type	Description	Formula
Range	Descriptive	Amount of spread between the extremes of the data	Highest data point minus the lowest
Standard deviation (SD)	Descriptive	Typical or (roughly speaking) average difference between the data points and their mean	$SD = \sqrt{\frac{\sum (X - M)^2}{n - 1}}$
Standard error (SE)	Inferential	A measure of how variable the mean will be, if you repeat the whole study many times	$SE = SD/\sqrt{n}$
Confidence interval (CI), usually 95% CI	Inferential	A range of values you can be 95% confident contains the true mean	$M \pm t_{(n-1)} \times SE$, where $t_{(n-1)}$ is a critical value of t. If n is 10 or more, the 95% CI is approximately $M \pm 2 \times SE$.

Bessel correction for an unbiased estimation of the standard deviation

Definition: t-test

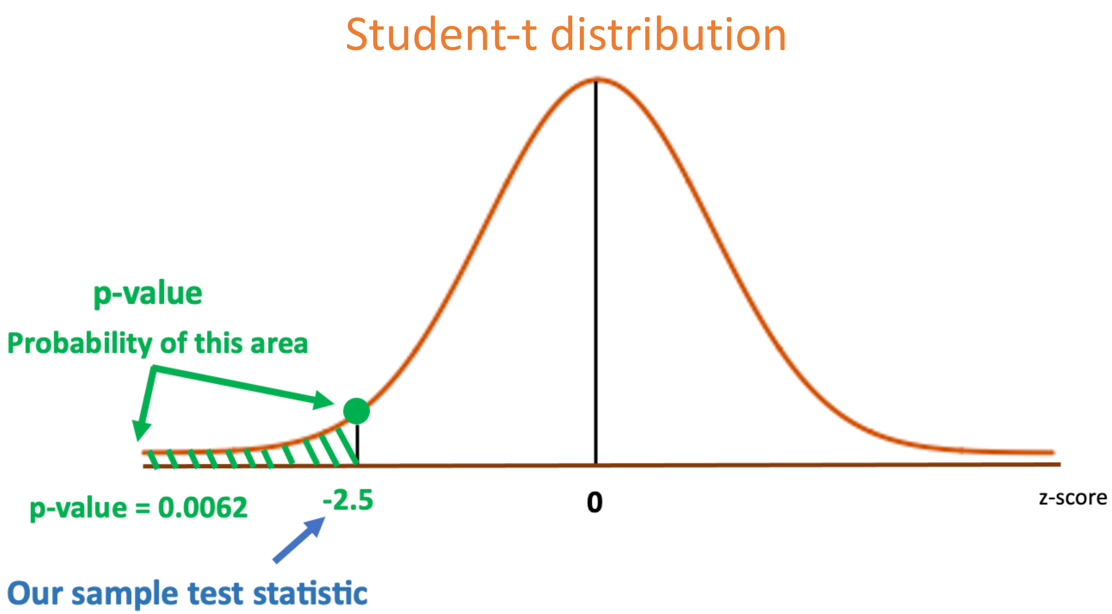
- A ***t-test*** is a statistical test to compare the means of two populations from their samples.





Definition: p-value

- The *__p-value__* is the probability that the difference in means observed would occur by chance.
  example: A p-value of 0.05 implies a chance of 95% that the difference observed is statistically significant.



The different types of t-test

• One-sample t-test: compares the mean of a sampled population to a predefined value.

example: are trees in Finland taller than the (known) average tree size?

• Independent two-sample t-test: compares the means of two sampled populations.

example: are pine trees taller than spruce trees on average?

• Paired two-sample t-test: compares the means of two conditions in which the same individuals participated.

example: are pine trees taller in winter than in summer?

Fun fact: a paired two-sample test is technically just a one-sample t-test performed on the variable defined by the difference between the two conditions for each individual.

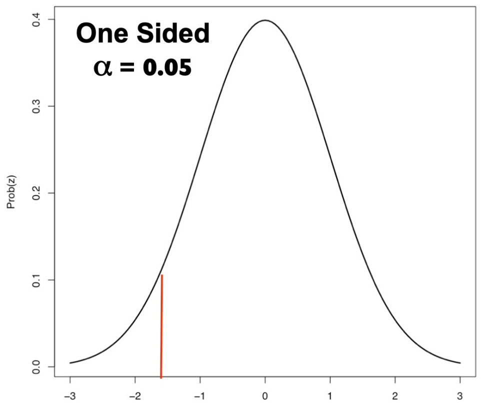
One-tailed vs. two-tailed t-test

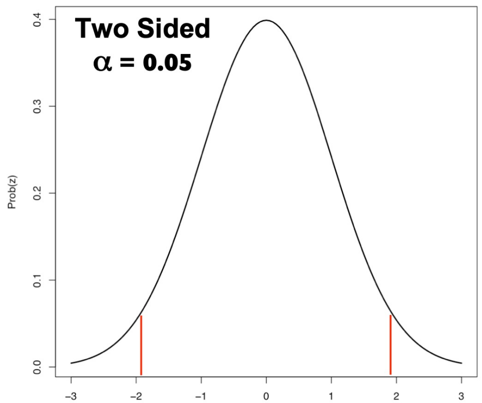
• one-tailed t-test: use when you know in which direction to expect a difference

example: do students sleep less in class after a good night sleep?

• two-tailed t-test: use when you <u>don't</u> know in which direction to expect a difference

example: are pine trees taller or smaller than spruce trees on average? I don't have a specific hypothesis.





How to calculate a t-test in practice?

- The t-test function exists in many programing languages (e.g., Excel, MATLAB, Python, R).
- You just need to choose the right t-test (one-tailed vs. two-tailed, one-sample vs. two-sample, paired vs. unpaired) and feed your data samples to it.

Language/Program	Function	Notes
Microsoft Excel pre 2010	`TTEST(array1, array2, tails, type)`	See [1]
Microsoft Excel 2010 and later	`T.TEST(array1, array2, tails, type)`	See [2]
Apple Numbers	`TTEST(sample-1-values, sample-2-values, tails, test-type)`	See [3]
LibreOffice Calc	`TTEST(Data1; Data2; Mode; Type)`	See [4]
Google Sheets	`TTEST(range1, range2, tails, type)`	See [5]
Python	`scipy.stats.ttest_ind(a, b, equal_var=True)`	See [6]
MATLAB	`ttest(data1, data2)`	See [7]
Mathematica	`TTest[{data1,data2}]`	See [8]
R	`t.test(data1, data2, var.equal=TRUE)`	See [9]
SAS	`PROC TTEST`	See [10]
Java	`tTest(sample1, sample2)`	See [11]
Julia	`EqualVarianceTTest(sample1, sample2)`	See [12]
Stata	`ttest data1 == data2`	See [13]

Correction for multiple comparisons

1 Why is multiple testing a problem?

Say you have a set of hypotheses that you wish to test simultaneously. The first idea that might come to mind is to test each hypothesis separately, using some level of significance α. At first blush, this doesn't seem like a bad idea. However, consider a case where you have 20 hypotheses to test, and a significance level of 0.05. What's the probability of observing at least one significant result just due to chance?

$$\begin{aligned} \mathbb{P}(\text{at least one significant result}) &= 1 - \mathbb{P}(\text{no significant results}) \\ &= 1 - (1 - 0.05)^{20} \\ &\approx 0.64 \end{aligned}$$

So, with 20 tests being considered, we have a 64% chance of observing at least one significant result, even if all of the tests are actually not significant. In genomics and other biology-related fields, it's not unusual for the number of simultaneous tests to be quite a bit larger than 20... and the probability of getting a significant result simply due to chance keeps going up.

Methods for dealing with multiple testing frequently call for adjusting α in some way, so that the probability of observing at least one significant result due to chance remains below your desired significance level.

Correction for multiple comparisons

2 The Bonferroni correction

The Bonferroni correction sets the significance cut-off at α/n. For example, in the example above, with 20 tests and $\alpha = 0.05$, you'd only reject a null hypothesis if the p-value is less than 0.0025. The Bonferroni correction tends to be a bit too conservative. To demonstrate this, let's calculate the probability of observing at least one significant result when using the correction just described:

$$\begin{aligned}\mathbb{P}(\text{at least one significant result}) &= 1 - \mathbb{P}(\text{no significant results}) \\ &= 1 - (1 - 0.0025)^{20} \\ &\approx 0.0488\end{aligned}$$

Here, we're just a shade under our desired 0.05 level. We benefit here from assuming that all tests are independent of each other. In practical applications, that is often not the case. Depending on the correlation structure of the tests, the Bonferroni correction could be extremely conservative, leading to a high rate of false negatives.

Note: there exists some less stringent correction, such as the false discovery rate (FDR)

Back to our example:

Table 2 Change in perceived anxiety before and after the theoretical training and first immersion into the cold lake water during the winter measured by Zung SAS (mean±SD)

	1st measurement	2nd measurement	P value
Intervention group (N=25)	41.1±11.0	39.3±11.2	**0.032 (0.031)**
Control group (N=24)	40.9±9.3	41.8±8.2	0.721 (0.730)

Bold significant differences between the first and second measurements: p≤0.05 (tested by Wilcoxon test; in brackets, we state the exact test as Monte Carlo).
SAS, Self-Rating Anxiety Scale.

Conlusion: small effect size, which is only mildly statistically significant on anxiety. But OK.

The effect size is small compared to the standard deviation of the data!

A p-value of 0.032 means:
there is a probablility of 0.032
that the difference observed
(or a greater difference) would
happen by chance.

In other words:
1/0.032 = 31
$\Rightarrow$ on 31 random draws,
such a difference would be observed
on expectation once.

Results (2)

Table 3 The life satisfaction before and after the periodical exposure to cold water for 8 weeks (mean±SD)

	Intervention group (N=25)			Control group (N=24)		
	Before	After	P value	Before	After	P value
HLTH	39.6±5.0	42.1±3.4	0.052 (0.051)	37.8±5.9	36.6±6.0	0.396 (0.414)
JAE	39.4±6.1	40.1±5.5	0.807 (0.816)	37.8±6.0	39.1±5.7	0.179 (0.183)
FIN	38.8±7.0	41.0±7.5	0.084 (0.083)	40.1±7.1	40.8±9.9	0.736 (0.746)
LT	35.3±7.0	36.2±7.2	0.688 (0.699)	33.4±6.6	33.0±8.0	0.749 (0.762)
MOP	39.2±4.6	40.2±5.9	0.252 (0.263)	36.6±7.6	37.5±7.1	0.106 (0.111)
SEX	38.5±5.5	40.9±5.6	**0.045 (0.046)**	38.9±7.6	39.3±8.1	0.101 (0.103)
FFR	33.2±4.2	33.8±4.0	0.713 (0.730)	33.6±5.0	33.8±5.5	0.849 (0.869)
LAH	34.2±5.1	34.6±5.7	0.566 (0.575)	33.3±6.5	33.8±5.9	0.391 (0.408)
SUM LSQ	259±28	269±29	0.178 (0.185)	254±36	258±37	0.939 (0.947)

Bold significant differences before and after the periodical exposure to cold water: p≤0.05 (tested by Wilcoxon test; in brackets, we state the exact test as Monte Carlo).
FFR, friends, family and relationships; FIN, finance; HLTH, health; JAE, job and employment; LAH, living and housing; SUM LSQ, Summary Life Satisfaction Questionnaire; LT, leisure time; MOP, my own person; SEX, sexuality.

1/0.029 = 34

1/0.042 = 24

1/0.045 = 22

total number of comparisons: 27

Table 4 Male and female body composition measured by the InBody 770 device before and after the periodical exposure to cold water for 8 weeks (mean±SD)

	Intervention group			Control group		
	Male (N=17)			Male (N=17)		
	Before	After	P value	Before	After	P value
Weight (kg)	85.2±9.7	84.6±9.6	0.132	83.1±9.7	83.1±8.8	0.989
BMI (kg)	25.9±2.8	25.8±2.8	0.201	25.4±2.8	25.2±2.5	1.000
WC (cm)	87.1±7.8	86.0±7.1	**0.029**	84.6±8.2	84.6±6.9	0.976
BFM (kg)	14.3±6.8	13.9±6.7	0.262	12.9±5.6	12.8±5.0	0.710
FFM (kg)	70.9±5.4	70.8±5.7	0.606	70.1±6.5	70.3±6.3	0.678
SSM (kg)	40.5±3.2	40.5±3.5	0.830	40.1±4.0	40.2±3.8	0.907
SSM (%)	47.8±3.3	48.0±3.3	–	48.5±3.2	48.5±3.0	–
PBF (%)	16.3±5.9	16.0±5.9	–	15.2±5.5	15.2±5.1	–
VFA (cm^2)	60.3±27.9	57.0±26.1	**0.042**	54.6±26.9	52.2±24.0	0.794
	Female (N=8)			Female (N=7)		
Weight (kg)	59.3±8.8	58.9±8.2	0.800	58.9±6.5	58.4±6.3	0.398
BMI (kg)	21.0±3.0	20.9±1.8	0.089	21.9±2.2	21.8±2.1	0.400
WC (cm)	67.1±4.2	67.0±3.6	0.810	69.6±7.2	69.6±6.6	0.865
BFM (kg)	11.5±3.4	11.3±3.3	0.799	12.5±4.1	12.4±4.1	0.600
FFM (kg)	47.8±5.5	47.6±5.2	0.205	46.4±5.0	46.0±4.9	0.612
SMM (kg)	26.6±3.4	26.5±3.2	0.205	25.7±3.0	25.5±2.9	0.735
SSM (%)	44.9±1.3	45.1±1.2	–	43.7±3.1	43.6±3.2	–
PBF (%)	19.0±3.0	18.9±3.0	–	21.1±5.3	21.1±5.5	–
VFA (cm^2)	43.9±14.3	41.2±12.4	0.327	49.1±16.0	47.8±15.9	0.176

Bold significant differences before and after periodical exposure to cold water: p≤0.05.
Male data were evaluated using the t-test. Due to the small sample size, female data were analysed using the Wilcoxon test.
BFM, body fat mass; BMI, body mass index; FFM, fat-free mass; PBF, percent body fat; SMM, skeletal muscle mass; VFA, visceral fat area; WC, waist circumference.

Many comparisons with no correction applied to the statistical tests:
=> methodological flaw which might invalidate the findings

Food for thought

- Maybe cold showers are actually good for you! An invalid study does not necessarily invalidate the claims of the study.

- claims reported in the media are usually not completely false, but their importance or certainty can be exaggerated, either by the media or by the authors of the study themselves.

- “statistically significant” ≠ “the effect size is significant” (the effect can be tiny but statistically significant).

Next lecture

- Fourier Transform, Wavelet Transforms, Spectrograms, High-pass, Low-pass Filters

- Give feedback at https://presemo.aalto.fi/bda2023
  Thank you.

Sources

- https://en.wikipedia.org/wiki/Student%27s_t-distribution
- https://en.wikipedia.org/wiki/Student%27s_t-test