Statistical Lab

No different differences

Mariia Chernieva , PhD-2 Life sciences

Statistical significance

Some hypothesis H

-what is a significant result

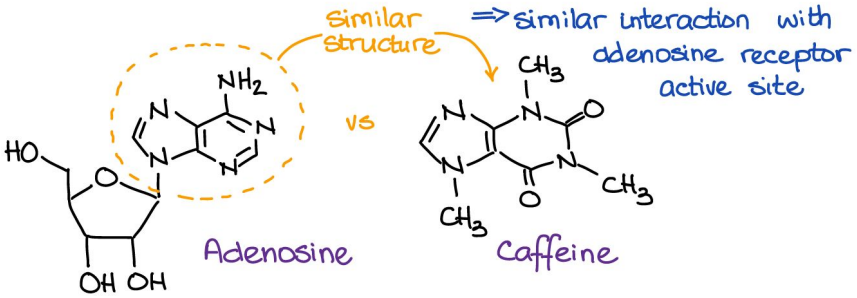
Some hypothesis H The probability of such data (observed statistics), or something even more extreme-lookgin, is low, if hypothesis H were true.

-what question would actually like to answer

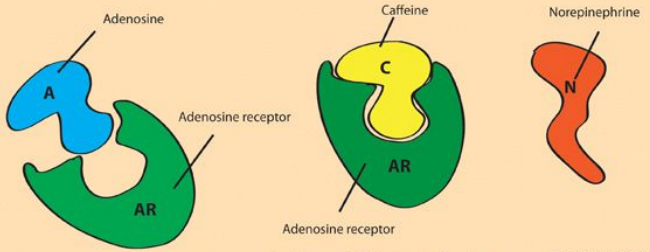
Is hypothesis H true

Here is some evidence against hypothesis H

Do you drink coffee?



Caffeine binding



Adenosine is a chemical in the brain that binds to specific receptors. Its job is to make humans feel tired.

Caffeine, which has a similar shape to adenosine, can also bind to these receptors when introduced to the brain. Caffeine blocks the binding of adenosine and stops the feeling of tiredness.

Norepinephrine is another chemical in the brain. It also binds to receptors; it acts as a stimulant.

Long-term effects

With caffeine blocking the effects of adenosine, norepinephrine may be reduced and the brain relies on caffeine to remain alert.

Statistical significance vs practical significance

Ochsner J. 2023 Summer;23(2):152–158. doi: 10.31486/toj.22.0073

Impact of Coffee Consumption on Cardiovascular Health

Michael F Mendoza [1,✉], Ralf Martz Sulague [2], Therese Posas-Mendoza [3], Carl J Lavie [4,5]

Author information · Article notes · Copyright and License information

PMCID: PMC10262944 PMID: 37323518

Abstract

Background: Coffee is a widely available beverage that is enjoyed by individuals of many cultures. The publication of new studies prompts a review of the clinical updates regarding the association between coffee consumption and cardiovascular disease.

Methods: We present a narrative review of the literature related to coffee consumption and cardiovascular disease.

Results: Recent (2000-2021) studies have shown that regular coffee consumption is associated with a decreased risk of developing hypertension, heart failure, and atrial fibrillation. However, results are inconsistent with regard to coffee consumption and risk of developing coronary heart disease. Most studies show a J-shaped association, wherein moderate coffee consumption resulted in decreased risk of coronary heart disease and heavy coffee consumption resulted in increased risk. In addition, boiled or unfiltered coffee is more

who did not show a statistically significant reduction in relative risk.[8] In a pilot crossover randomized study by Revuelta-Iniesta and Al-Dujaili (2014), green coffee was found to significantly reduce systolic blood pressure (P=0.02) by a mean of 2.65 ± 1.37 mm Hg compared to baseline.[9]

DRINK Coffee =)

Statistical significance vs practical significance

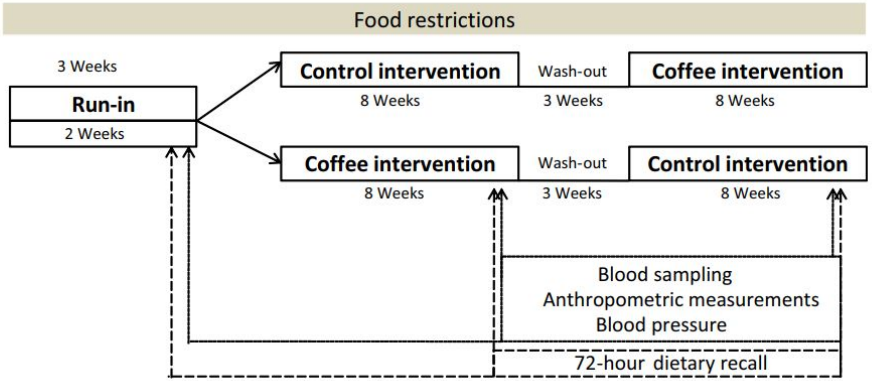
Home › European Journal of Nutrition › Article

Regularly consuming a green/roasted coffee blend reduces the risk of metabolic syndrome

Original Contribution | Published: 13 October 2016
Volume 57, pages 269–278, (2018) Cite this article

European Journal of Nutrition
Aims and scope →
Submit manuscript →

Study design



Results

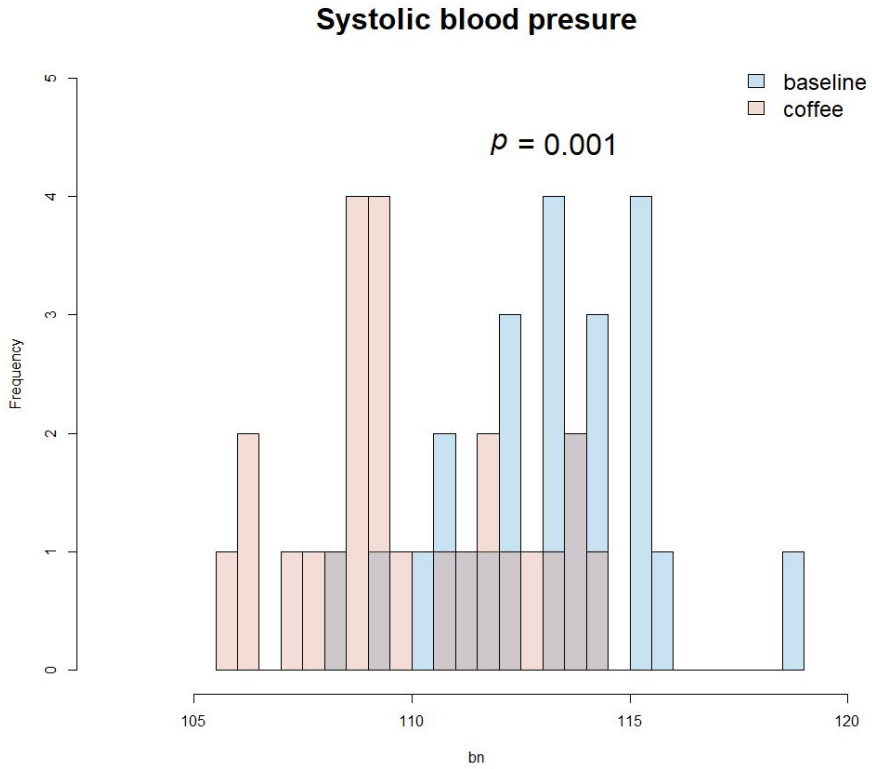
Systolic and diastolic blood pressure decreased ($p = 0.001$ and $p < 0.001$, respectively) in both groups as well as %body fat ($p = 0.001$) which may be related to the lower leptin ($p = 0.001$), PAI-1 ($p < 0.001$) and resistin ($p = 0.034$) levels in the two groups after coffee consumption. Glucose concentration ($p = 0.030$) and insulin resistance ($p = 0.011$; HOMA-IR) also decreased, as well as triglyceride levels ($p = 0.017$), so that the reduction was much greater in the hypercholesterolaemics (group effect, $p = 0.027$).

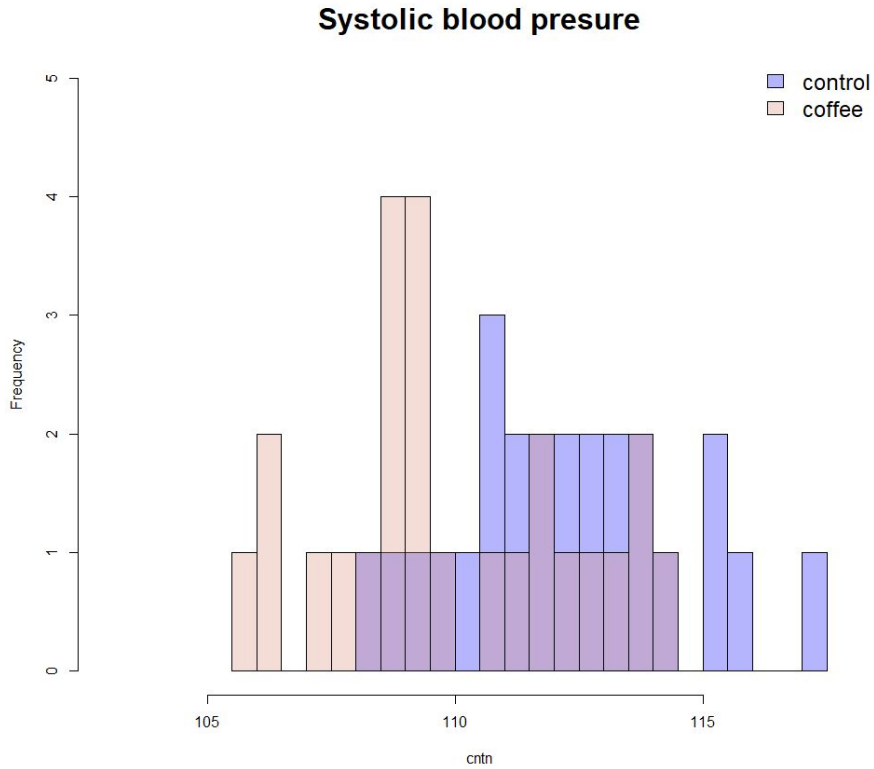
Coffee consumption, 2016

Table 3 Effect of the consumption of the roasted/green coffee on components of the metabolic syndrome

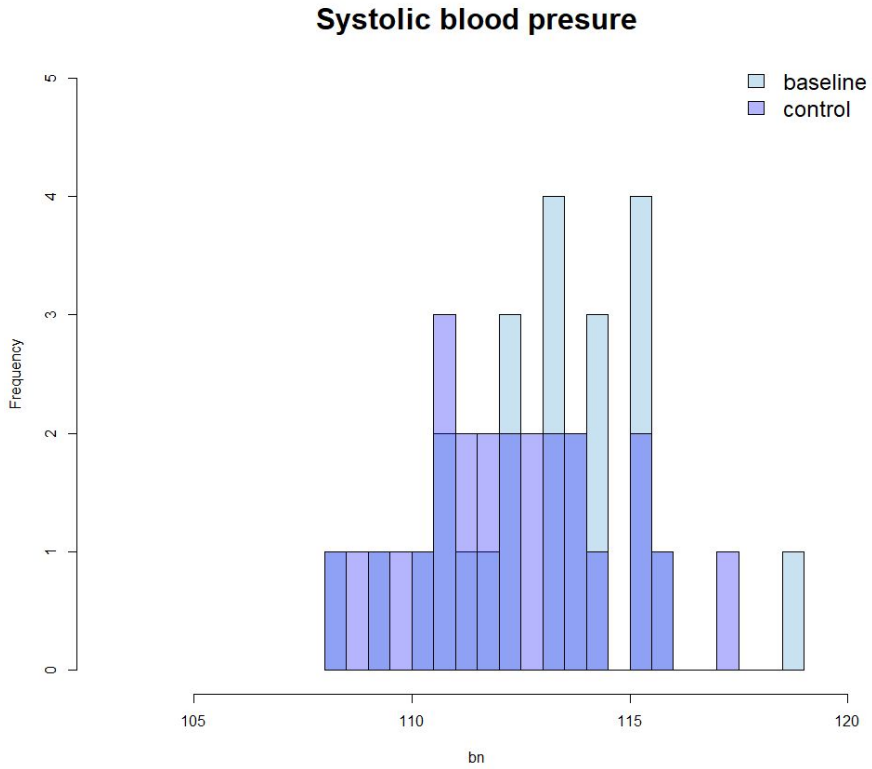
cm	Normocholesterolaemic ($n = 25$)			Hypercholesterolaemic ($n = 27$)			P*
	Baseline	Control	Coffee	Baseline	Control	Coffee	
Waist circumference	70.4 ± 1.3	70.6 ± 1.4	70.9 ± 1.4	76.8 ± 2.5	76.6 ± 2.5	75.6 ± 2.5	N.S.
mmHg							
Systolic BP	113.3 ± 2.1	112.6 ± 2.0	109.9 ± 2.1	119.4 ± 2.9[a]	115.8 ± 2.5[ab]	114.2 ± 3.1[b]	0.001
Diastolic BP	69.5 ± 1.2	69.2 ± 1.0	67.2 ± 1.2	76.8 ± 2.2[a]	73.3 ± 1.9[b]	71.2 ± 2.2[b]	<0.001
mg/dL							
Triglycerides	71.3 ± 7.1	70.0 ± 5.2	71.0 ± 5.7	103.3 ± 7.5[a]	86.6 ± 5.5[ab]	82.9 ± 6.1[b]	0.017
HDL-C	55.9 ± 2.9	56.7 ± 2.4	58.2 ± 2.7	63.1 ± 3.1	59.4 ± 2.5	62.7 ± 2.9	N.S.
Glucose	74.0 ± 1.5	74.9 ± 1.4	71.0 ± 1.6	76.95 ± 1.6	74.57 ± 1.4	73.20 ± 1.7	0.030

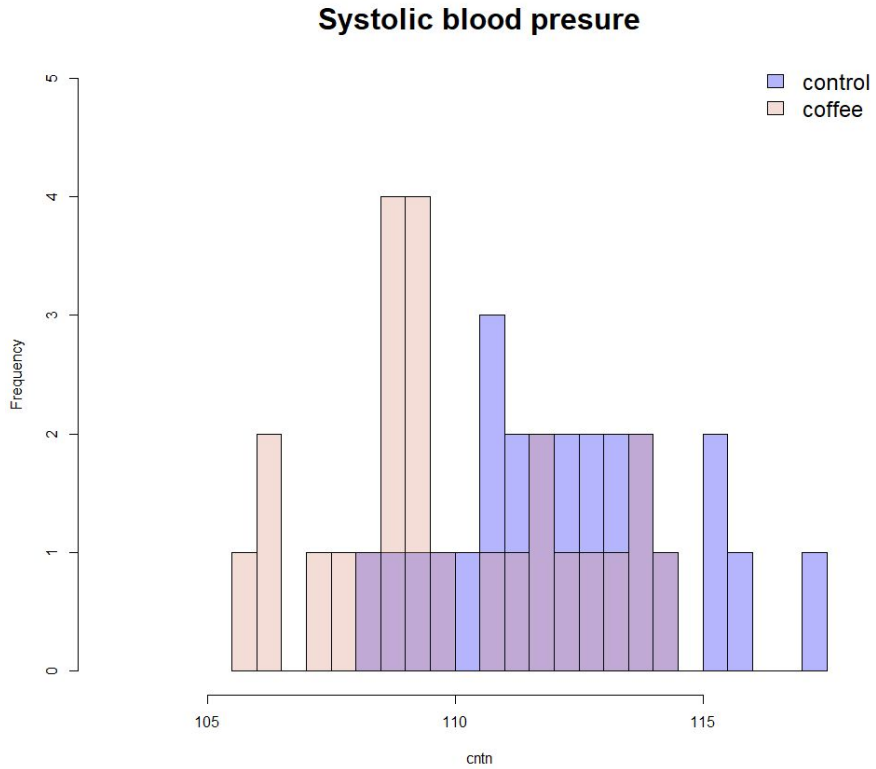
Coffee consumption, 2016.
Normocholesterolaemic (n=25)



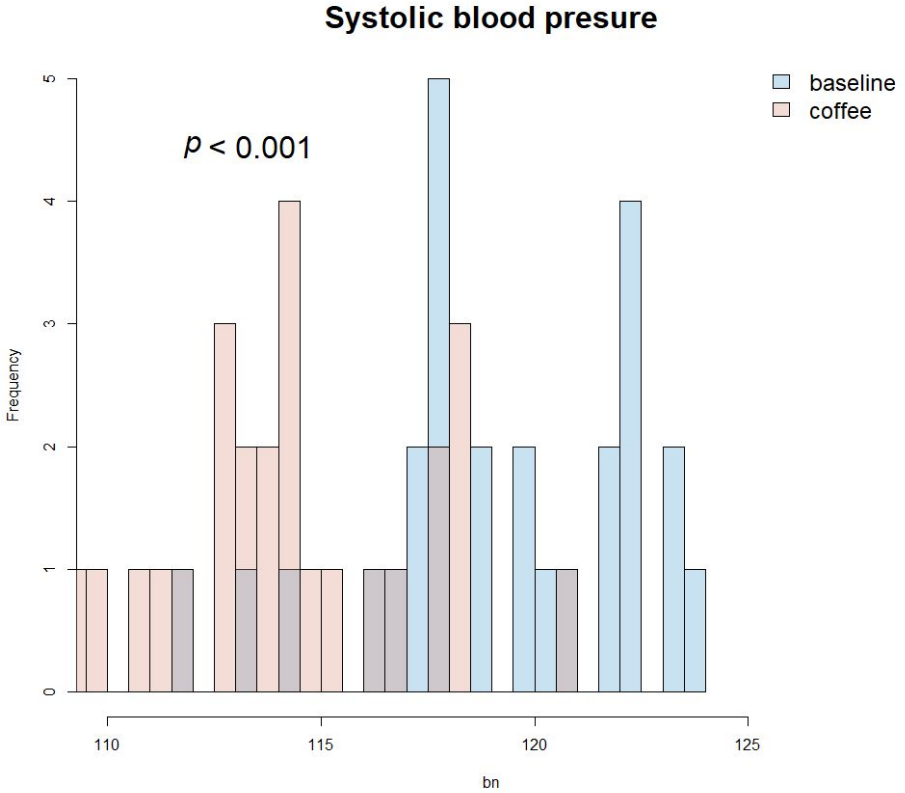


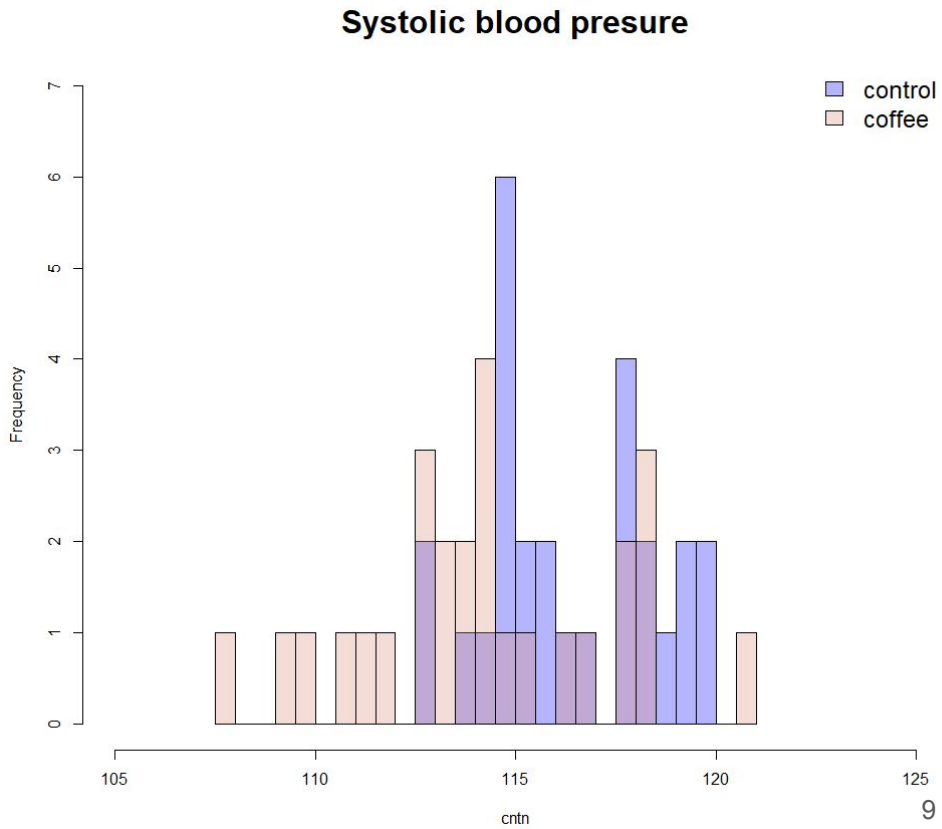
Coffee consumption, 2016.
Normocholesterolaemic (n=25)



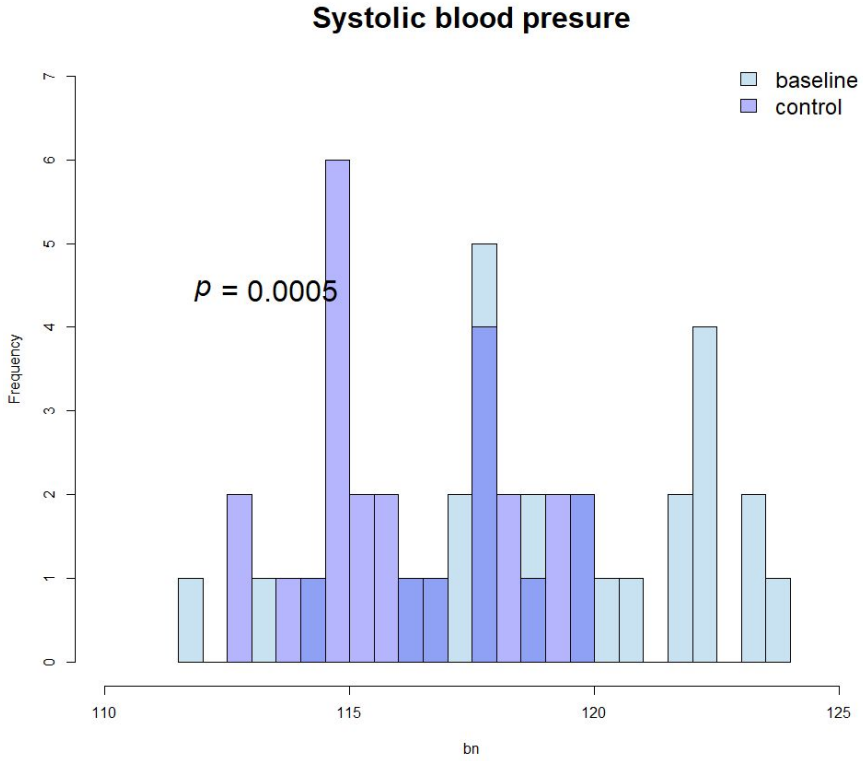


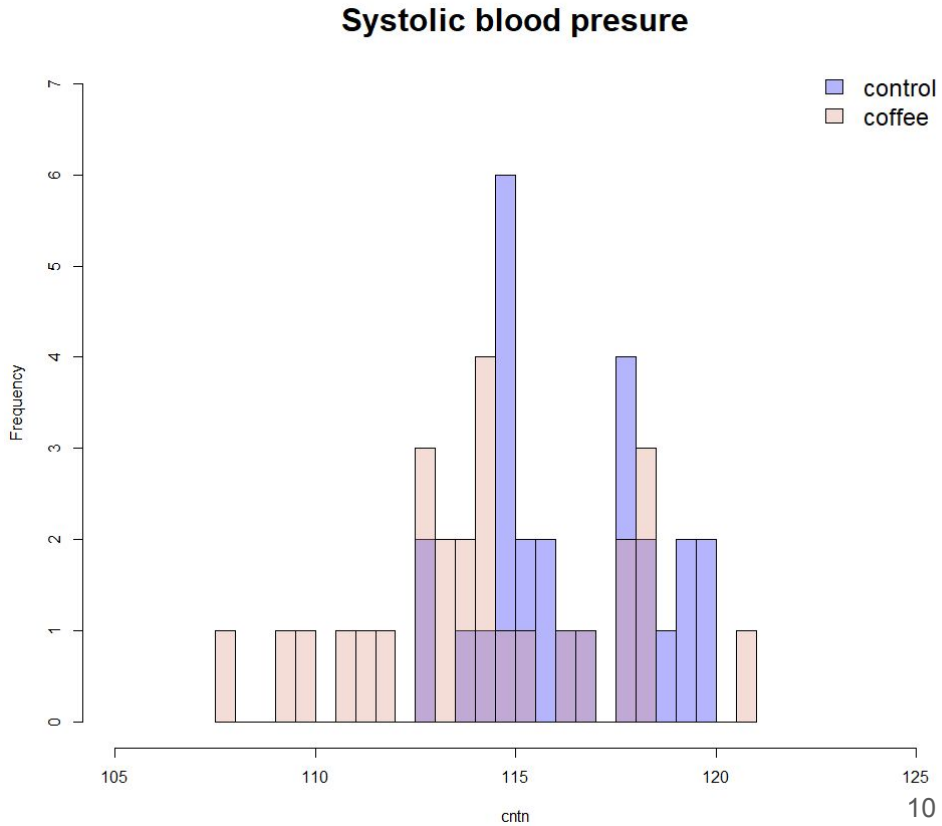
Coffee consumption, 2016.
Hypercholesterolaemic (n=27)





Coffee consumption, 2016.
Normocholesterolaemic (n=25)





Other approach: Odds Ratios

What is an odds ratio?

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. Odds ratios are most commonly used in case-control studies, however they can also be used in cross-sectional and cohort study designs as well (with some modifications and/or assumptions).

Exposure status	Outcome status +	Outcome status -
+	a	b
-	c	d

a = Number of exposed cases
b = Number of exposed non-cases
c = Number of unexposed cases
d = Number of unexposed non-cases

OR shows the strength of the association between two events, **A** and **B**

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc}$$

OR=1 Exposure does not affect odds of outcome
OR>1 Exposure associated with higher odds of outcome
OR<1 Exposure associated with lower odds of outcome

Fisher test: p-value for OR

Outcome status

Exposure status	+	-	total
+	*a*	*b*	*a* + *b*
-	*c*	*d*	*c* + *d*
totals	*a* + *c*	*b* + *d*	*n*

a = Number of exposed cases
b = Number of exposed non-cases
c = Number of unexposed cases
d = Number of unexposed non-cases

$$p = \binom{a+b}{a}\binom{c+d}{c} \Big/ \binom{n}{a+c} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!} \quad (1)$$

p-value:
1.Calculate p_1 for present table using (1)
2. Search for min value in table
3. Decrease *min* value with same sum(columns) and sum(rows)-> new table
4. Calculate p_2 for new table using (1)
5. etc. until min value table ==0

$\text{sum}(p_i)$ = p-value

OR significance. Example 1

Previously Suicidal Adolescents: Predictors of Six-Month Outcome

Brian Greenfield [1,2,✉], Melissa Henry [1,3], Margaret Weiss [4,5], Sze Man Tse [1], Jean-Marc Guile [1,6], Geoffrey Dougherty [1,2], Xun Zhang [2,7], Eric Fombonne [2], Eric Lis [1], Sam Lapalme-Remis [1], Bonnie Harnden [8]

▸ Author information ▸ Article notes ▸ Copyright and License information

PMCID: PMC2583916 PMID: 19018322

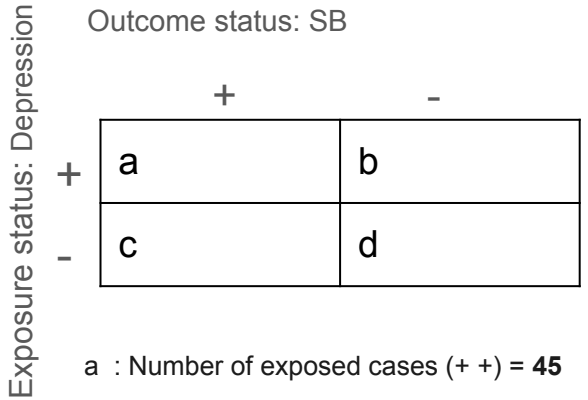
Objective

To determine the baseline variables, including borderline personality disorder (BPD), associated with the six-month outcome of previously suicidal adolescents (n=263) presenting to an emergency department and treated predominantly as out-patients.

What is the OR of suicidal behaviour at six months follow-up given presence of depression at baseline?

Depression as factor for Suicidal Behaviour

- 186 of the 263 adolescents previously judged as having experienced a suicidal behaviour requiring immediate psychiatric consultation did not exhibit suicidal behaviour (non-suicidal, NS) at six months follow-up.
- Of this group, 86 young people had been assessed as having depression at baseline.
- Of the 77 young people with persistent suicidal behaviour at follow-up (suicidal behaviour, SB), 45 had been assessed as having depression at baseline.



a : Number of exposed cases (+ +) = **45**

b : Number of exposed non-cases (+ –) = **86**

c : Number of unexposed cases (– +) = **32**

d : Number of unexposed non-cases (– –) = **100**

OR = (a/c)/(b/d) = 45*100/32*86 = 1.63

It seems baseline depression is factor, but what about CI?

Depression as factor for Suicidal Behaviour?

Table 1.

Clinical characteristics of the suicidal adolescents at baseline and between-group differences

Characteristic	Suicidal (S) (n=77) % or mean±SD	Non-Suicidal (NS) (n=186) % or mean±SD	Odds Ratio) (S/NS)	p-value	95% C.I.
Age	14.7 ± 1.4	14.5 ± 1.6	1.1	0.34	0.9 – 1.3
Sex (female)	79.2	64.0	2.2	0.02	1.2 – 4.0
IFR score	46.1 ± 23.8	42.4 ± 24.7	1.0	0.27	1.0 – 1.0
Depression	58.4	46.2	1.6	0.07	1.0 – 2.8
Conduct disorder	32.5	19.8	2.0	0.03	1.1 – 3.6
Life events	11.7 ± 7.0	10.2 ± 6.6	1.0	0.10	1.0 – 1.1
Number of previous hospitalization	0.5 ± 0.7	0.3 ± 0.5	1.6	0.03	1.1 – 2.5
Borderline personality disorder	90.9	72.6	3.8	0.002	1.6 – 8.7
Previous suicide attempt(s)	89.6	71.6	3.4	0.02	1.2 – 9.6
Drug use	71.4	45.9	3.0	0.001	1.7 – 5.2
Alcohol use	63.3	48.8	1.8	0.09	0.9 – 3.6
CGAS score	38.6 ± 11.5	40.0 ± 11.0	1.0	0.37	1.0 – 1.0
Parent previous suicide attempt(s)	87.5	78.7	1.9	0.37	0.5 – 7.7

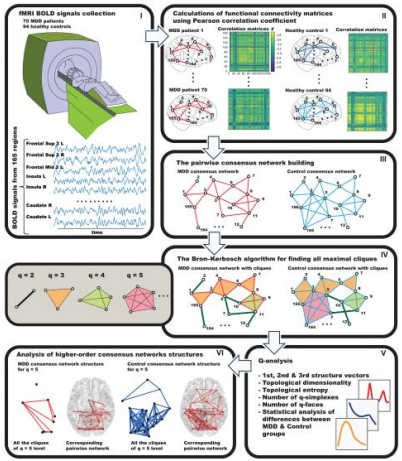
p-value = 0.07

OR =1.63 > 1 (!)

Depression biomarker, fMRI

“detect depression”
“diagnostic method”

fMRI data



Ученые научились вычислять депрессию

© Фото : nensuria

Читать ria.ru в Дзен

МОСКВА, 28 фев — РИА Новости. Инновационный математический подход — Q-анализ для диагностики клинической депрессии создали ученые БФУ имени И. Канта. По их мнению, новый метод поможет разработать более точные методы диагностики и персонализированные подходы к терапии, сообщили РИА Новости в пресс-службе Минобрнауки РФ

Depression biomarker, fMRI

TABLE 4 Statistical Significance of the Modulus of Difference Between the HC and MDD Groups of the First, Second, and Third Structure Vectors Estimated by the Permutation-Based Procedure

Measure	p-value
$\lvert \mathbf{Q}^{MDD} - \mathbf{Q}^{HC} \rvert$	0.0453
$\lvert \mathbf{N_s}^{MDD} - \mathbf{N_s}^{HC} \rvert$	0.2052
$\lvert \hat{\mathbf{Q}}^{MDD} - \hat{\mathbf{Q}}^{HC} \rvert$	0.0308

ieeexplore.ieee.org/document/10811911

Marker genes, genetic tests, (ACTN3)

https://www.genotek.ru/demo/health/risks

In this meta-analysis, we present clear associations between the R allele/RX genotype in the *ACTN3* polymorphism and elite power athlete status. Significant effects of the R allele (overall analysis, Western and female subgroups) and RX genotype (Asians and males) were for the most part, results of outlier treatment. Interaction analysis improved the female outcome. These robust findings were free of publication bias.

https://pmc.ncbi.nlm.nih.gov/articles/PMC6542526/

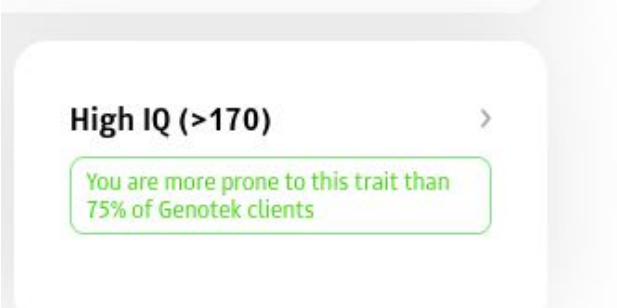
Association of the *ACTN3* R577X (rs1815739) polymorphism with elite power sports: A meta-analysis

Results

Significant, but poor effect

Significant overall R allele effects (OR 1.21, 95% CI 1.07–1.37, $P^a = 0.002$) were confirmed in the Western subgroup (OR 1.11, 95% CI 1.01–1.22, $P^a = 0.02$) and with outlier treatment (ORs 1.12–1.20, 95% CIs 1.02–1.30, $P^a < 10^{-5}$–0.01). This treatment resulted in acquired significance of the RX effect in Asian athletes (OR 1.91, 95% CI 1.25–2.92, $P^a = 0.003$). Gender analysis dichotomized the RX genotype and R allele effects as significantly higher in male (OR 1.14, 95% CI 1.02–1.28, $P^a = 0.02$) and female (OR 1.58, 95% CI 1.21–2.06, $P^a = 0.0009$) athletes, respectively, when compared with controls. Significant R female association was improved with a test of interaction ($P_{\text{interaction BC}} = 0.03$). The overall, Asian and female outcomes were robust. The R allele results were more robust than the RX genotype outcomes. No evidence of publication bias was found.

Marker genes, genetic tests (a couple of genes)



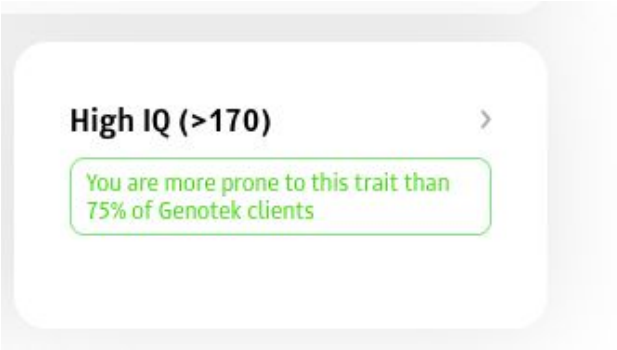
https://www.genotek.ru/demo/health/risks

Table 1 Genomic loci and lead SNPs associated with intelligence in the meta-analysis based on $n = 78,308$

rsID	Annotation	Locus[a]	Ref	Alt	RefF	z	P value	Direction[b]
rs2490272	*FOXO3* intronic	6q21	T	C	0.63	7.44	9.96×10^{-14}	++++−+++
rs9320913	Intergenic	6q16.1	A	C	0.48	6.61	3.79×10^{-11}	++++−+++
rs10236197	*PDE1C* intronic	7p14.3	T	C	0.63	6.46	1.03×10^{-10}	+++++−++
rs2251499	Intergenic	13q33.2	T	C	0.26	6.31	2.74×10^{-10}	++++++++
rs36093924	*CYP2D7* ncRNA_intr	22q13.2	T	C	0.46	−6.31	2.87×10^{-10}	?−−?????
rs7646501	Intergenic	3p24.2	A	G	0.74	6.02	1.79×10^{-9}	?++−++++
rs4728302	*EXOC4* intronic	7q33	T	C	0.60	−5.97	2.42×10^{-9}	−−−+−−+−

https://www.nature.com/articles/ng.3869

Marker genes, genetic tests (a couple of genes)



https://www.genotek.ru/demo/health/risks

5% of intelligence variance explained

We calculated the variance explained (R^2) in intelligence by the GWAS results in four independent samples, using LDpred[16] (Online Methods, **Supplementary Fig. 5** and **Supplementary Table 7**). Our calculations show that the current results explain up to 4.8% of the variance in intelligence and that on average across the four samples there is a 1.9-fold increase in explained variance in comparison to the most recent GWAS on intelligence[6].

https://www.nature.com/articles/ng.3869