

# Multiple testing: to adjust or not to adjust

**Yudi Pawitan and Arvid Sjölander**

Department of Medical Epidemiology and  
Biostatistics

Karolinska Institutet



**Karolinska  
Institutet**



## LENA SUNDSTRÖM

"Vi sörjer de döda i terrorattentatet. Sorgen förenar oss för en stund i en stilla värdighet och gör oss alla lika."

## Allt inför Mellan.

Läs DN digitalt i en månad för 1 kr!

Beställ här



# Du kan också göra en prognos för hela din pension här.

minpension.se

PENSIONS  
MYNDIGHETEN



## KARIN BOJS VETENSKAPSKRÖNIKA

# Karin Bojs: Falsk matematik om farmors matvanor

Publicerad 2014-03-02 14:43

Under de senaste tjugo åren har jag gjort många

ANNONS:

Prata



RESEARCH ARTICLE

Open Access

# Change in paternal grandmothers' early food supply influenced cardiovascular mortality of the female grandchildren

Lars Olov Bygren<sup>1,2\*</sup>, Petter Tinghög<sup>3</sup>, John Carstensen<sup>4</sup>, Sören Edvinsson<sup>5</sup>, Gunnar Kaati<sup>1</sup>, Marcus E Pembrey<sup>6</sup> and Michael Sjöström<sup>1</sup>

## Abstract

**Background:** This study investigated whether large fluctuations in food availability during grandparents' early development influenced grandchildren's cardiovascular mortality. We reported earlier that changes in availability of food - from good to poor or from poor to good - during intrauterine development was followed by a double risk of sudden death as an adult, and that mortality rate can be associated with ancestors' childhood availability of food. We have now studied transgenerational responses (TGR) to sharp differences of harvest between two consecutive years' for ancestors of 317 people in Överkalix, Sweden.

**Results:** The confidence intervals were very wide but we found a striking TGR. There was no response in cardiovascular mortality in the grandchild from sharp changes of early exposure, experienced by three of the four

**Results:** The confidence intervals were very wide but we found a striking TGR. There was no response in cardiovascular mortality in the grandchild from sharp changes of early exposure, experienced by three of the four grandparents (maternal grandparents and paternal grandfathers). If, however, the paternal grandmother up to puberty lived through a sharp change in food supply from one year to next, her sons' daughters had an excess risk for cardiovascular mortality (HR 2.69, 95% confidence interval 1.05-6.92). Selection or learning and imitation are unlikely explanations. X-linked epigenetic inheritance via spermatozoa seemed to be plausible, with the transmission, limited to being through the father, possibly explained by the sex differences in meiosis.

**Conclusion:** The shock of change in food availability seems to give specific transgenerational responses.

**Keywords:** Epidemiology, Food change, Environmental shock, Human transgenerational response, Cardiovascular mortality, Överkalix

---

**Table 1 Grandparents' childhood experience of drastic change in food availability, from one year to the next year, by descendants' cardiovascular mortality\***

	Men	Women
Paternal grandfather	0.87 (0.46-1.64)	0.91 (0.43-1.96)
Paternal grandmother	0.64 (0.32-1.29)	<b>2.69 (1.05-6.92)</b>
Maternal grandfather	1.26 (0.68-2.34)	1.32 (0.58-3.04)
Maternal grandmother	0.69 (0.35-1.36)	0.56 (0.22-1.49)
Index cases	151	126

\*Associations are presented as Hazards ratios (HRs) with 95% confidence intervals (CIs 95%). The models are adjusted for birth cohort, mother's literacy, father's death before the index person attained 13 years of age. Bold numbers indicate significant associations,  $p < 0.05$ .

P-value=0.04

Q: Adjust for multiple testing?

# Statistical inference is not simple!

BASIC AND APPLIED SOCIAL PSYCHOLOGY, 37:1–2, 2015  
Copyright © Taylor & Francis Group, LLC  
ISSN: 0197-3533 print/1532-4834 online  
DOI: 10.1080/01973533.2015.1012991



## Editorial

David Trafimow and Michael Marks

*New Mexico State University*

The *Basic and Applied Social Psychology* (BASP) 2014 Editorial emphasized that the null hypothesis significance testing procedure (NHSTP) is invalid, and thus authors would be not required to perform it (Trafimow, 2014). However, to allow authors a grace period, the Editorial stopped short of actually banning the NHSTP. The purpose of the present Editorial is to announce that the grace period is over. From now on, BASP is banning the NHSTP.

With the banning of the NHSTP from BASP, what are the implications for authors? The following are

a strong case for rejecting it, confidence intervals do not provide a strong case for concluding that the population parameter of interest is likely to be within the stated interval. Therefore, confidence intervals also are banned from BASP.

Bayesian procedures are more interesting. The usual problem with Bayesian procedures is that they depend on some sort of Laplacian assumption to generate numbers where none exist. The Laplacian assumption is that when in a state of ignorance, the researcher should assign an equal probability to each possibility. The

The *Basic and Applied Social Psychology* (BASP) 2014 Editorial emphasized that the null hypothesis significance testing procedure (NHSTP) is invalid, and thus authors would be not required to perform it (Trafimow, 2014). However, to allow authors a grace period, the Editorial stopped short of actually banning the NHSTP. The purpose of the present Editorial is to announce that the grace period is over. From now on, BASP is banning the NHSTP.

**Question 2.** *What about other types of inferential statistics such as confidence intervals or Bayesian methods?*

**Answer to Question 2.** Confidence intervals suffer from an inverse inference problem that is not very different from that suffered by the NHSTP. In the NHSTP, the problem is in traversing the distance from the probability of the finding, given the null hypothesis, to the probability of the null hypothesis, given the finding. Regarding confidence intervals, the problem is that, for example, a 95% confidence interval does not indicate that the parameter of interest has a 95% probability of being within the interval. Rather, it means merely that if an infinite number of samples were taken and confidence intervals computed, 95% of the confidence intervals would capture the population parameter. Analogous to how the NHSTP fails to provide the prob-



# Primary concern

- Too easy to get  $P\text{value} < 0.05$
- Hence too many false positives
- BASP 'solution': ban statistical inference!



**Karolinska  
Institutet**

# Multiplicity issues arise in...

- Hypothesis testing (formal)
  - More tests = more possibilities of errors
- Estimation (less formal)
  - More estimates = higher probabilities for extreme results
- Modelling, data analyses (informal)
  - Definition of variables
  - Subgroup analyses, un-documented search
  - Model selection, regression analysis: p-values after model selection are not meaningful.
- **Actually: even in a single test!**



# False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science

XX(X) 1–8

© The Author(s) 2011

Reprints and permission:

[sagepub.com/journalsPermissions.nav](http://sagepub.com/journalsPermissions.nav)

DOI: 10.1177/0956797611417632

<http://pss.sagepub.com>



Joseph P. Simmons<sup>1</sup>, Leif D. Nelson<sup>2</sup>, and Uri Simonsohn<sup>1</sup>

<sup>1</sup>The Wharton School, University of Pennsylvania, and <sup>2</sup>Haas School of Business, University of California, Berkeley

## Abstract

In this article, we accomplish two things. First, we show that despite empirical psychologists' nominal endorsement of a low rate of false-positive findings ( $\leq .05$ ), flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates. In many cases, a researcher is more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not. We present computer simulations and a pair of actual experiments that demonstrate how unacceptably easy it is to accumulate (and report) statistically significant evidence for a false hypothesis. Second, we suggest a simple, low-cost, and straightforwardly effective disclosure-based solution to this problem. The solution involves six concrete requirements for authors and four guidelines for reviewers, all of which impose a minimal burden on the publication process.

## Keywords

methodology, motivated reasoning, publication, disclosure

Received 3/17/11; Revision accepted 5/23/11

**Table 1.** Likelihood of Obtaining a False-Positive Result

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ( $r = .50$ )	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

Note: The table reports the percentage of 15,000 simulated samples in which at least one of a set of analyses was significant. Observations were drawn independently from a normal distribution. Baseline is a two-condition design with 20 observations per cell. Results for Situation A were obtained by conducting three  $t$  tests, one on each of two dependent variables and a third on the average of these two variables. Results for Situation B were obtained by conducting one  $t$  test after collecting 20 observations per cell and another after collecting an additional 10 observations per cell. Results for Situation C were obtained by conducting a  $t$  test, an analysis of covariance with a gender main effect, and an analysis of covariance with a gender interaction (each observation was assigned a 50% probability of being female). We report a significant effect if the effect of condition was significant in any of these analyses or if the Gender  $\times$  Condition interaction was significant. Results for Situation D were obtained by conducting  $t$  tests for each of the three possible pairings of conditions and an ordinary least squares regression for the linear trend of all three conditions (coding: low =  $-1$ , medium =  $0$ , high =  $1$ ).

# Basic concepts

- Single test of  $H_0$  using statistic  $Z$ 
  - $\alpha = P(\text{rejection} | H_0) = P(\text{false rejection})$
  - P-value =  $P(\text{more extreme } Z | H_0)$
- Several hypotheses:  $H_{01} \dots H_{0m}$ 
  - P-values:  $p_1 \dots p_m$
  - FP = number of false positives/rejections



# Outcomes from multiple tests

		Test Result		
		NoRej	Reject	Total
True	Null	90	<b>FP=5</b>	<b>95</b>
	NonNull	FN=1	TP=4	5
Total		91	D=9	100

## Discovery terminology:

- D = significant results, discoveries
- FP = number of false positives/discoveries
- TP = number of true positives/disc
- FP/D = false discovery rate



# Classical procedures

- Family-wise error rate (FWER) =  $P(\text{FP} > 0)$ 
  - Control:  $\text{FWER} \leq \alpha$  (e.g. 0.05)
  - Or: guarantee  $\text{FP} = 0$  with large probability
- General approach: adjust the individual p-values or alpha level (critical value)
- Simplest: Bonferroni method
  - adjust p-values:  $m * p_i$  (less significance)
  - adjust level:  $\alpha/m$  (harder to reject)



- Large literature: Sidak, Holm, Westfall&Young's maxT, minP, Hochberg, Troendle, etc
- More recent: false discovery rate (FDR)



**Karolinska  
Institutet**



# Bygren et al study

- $M=8$ ,  $\text{min.Pvalue} = 8 \times 0.04 = 0.32$
- Cannot reject null hypothesis?
- **Logical question:  $M=?$** 
  - Limit to all tests from one table?
  - one paper? Can split paper? Collection of papers?
  - one researcher over a life time?
  - How to account future papers?
- Call this **multiple-testing dilemma.**



# Strict application: clinical trials

- Even with single hypothesis (treatment=placebo)
  - Interim analyses: m group sequential tests
  - Many methods
- FDA 46-page doc “Guidance for Industry: E9 Statistical Principles for Clinical Trials”
  - “adjustment should always be considered ...
  - Otherwise explain.
- FDAAA 801 Requirement: registration
- **Main idea: in assessing evidence, intention (eg interim analyses) matters.**



# Other than clinical trials

- No FDA rule/guideline
- Rothman (Epidemiology, **1990**): "No adjustments are needed for multiple comparisons"
- Rothman (J Gen Intern Med, **2014**): " Six Persistent Research Misconceptions"
  - No.5: One should always ... adjust for multiple comparisons
- What is the fundamental difference (with clinical trials)?



# Example 1

**Table 1. p-values in a study of a metastatic cancer drug vs placebo for ten patient characteristics.**

<b>Variables</b>	<b>p-value</b>	<b>10<sup>*</sup>p-value</b>
1 Karnofsky index	0.007	0.07
2 Body weight	0.013	0.13
3 Tricep skin-fold	0.091	0.91
4 Hemoglobin concentration	0.236	2.36
5 Erythr sedimentation rate	0.350	3.50
6 Albumin in serum	0.525	5.25
7 Creatinine in serum	0.535	5.35
8 Bilirubin in serum	0.662	6.62
9 S-alkaline phosphatase	0.823	8.23
10 Alanine aminotransferase	0.908	9.08

- is there evidence of benefit?



# Different scenarios

- A. **We know nothing** before the study.
  - Order is posthoc. Adjustment is reasonable.
- B. Karnofsky index is the main interest, **stated in advance**. No adjustment is needed.
- C. We are in A, but **another group** is only interested in the index.
  - If they only ask for that variable, for them adjustment is **not needed?**
  - What if they choose to collaborate and co-author the paper? (after the data collection)



## Example 2: single test?

- a client comes with a study (say  $n=100$ ); an analysis **obtains  $z=2.1$**
- Is it significant ( $\alpha=0.05$ )? (p-value=0.036)
- Yes, but...
- What did he plan to do if the result were not significant?
  - Collect more data.



- Actual procedure:
  - Collect  $n=100$  and test if  $|z| > c$ , if significant stop.
  - Otherwise, collect 100 more, and test if new  $|z| > c$ .
  - To get FWER  $\alpha=0.05$ , must use  $c=2.18$
- So, observed  $z=2.1$  is not significant!
- Significance is affected by the intention of the scientist, not just the data. Is this ok?



- Suppose the scientist collected more data (final  $n=200$ ), and got final  $z=2.1$ . This is not significant ( $p\text{-value}=0.055$ ).
- Then the full dataset is put on the web, and another scientist downloads it.
  - He will get  $z=2.1$
  - For him, is it a significant result? ( $p\text{-value}=0.036$ ).





# How to explain?

- Pvalue is objectively meaningful under replications of the experiment
  - Precise setup/specs needed
  - Replications are (mostly) hypothetical
- 1 study (test) can be embedded in different hypothetical replication studies.
  - Different experimenters may have different hypothetical replications in mind
  - Properties of the hypothetical studies may be different. One implies adjustment, the other does not.
  - Example 1: scenarios A, B and C
  - Example 2: first vs second reaction
  - **Whose perspective is correct?**



**Karolinska  
Institutet**

- In real life this is not so strange: a person may belong to different groups with different properties
  - Case: Arvid, a young male looking for car insurance
    - Insurance company: young male  $\in$  {**reckless drivers**}
    - Arvid claims  $\in$  {**safety-conscious**}
    - **Both are valid**



**Karolinska  
Institutet**

- How to reconcile different perspectives?
  - Choose one by decree/law (eg FDA, insurance obligation), no dilemma for decision.
  - Not take decision when we do not have to: reserve judgement. Good: open-minded, bad: undecisive.
- **Seemingly our only problem:** should we reject null hypothesis [based on this dataset]?



# Key differences

- **Clinical trials:** Decision must be made.
  - If reject null → \$\$\$
  - Future monitoring to do, but does not affect the need to decide now.
- **Scientist's attitude:** all results are provisional
  - maybe there is something here, interesting to investigate further,
  - let's think of an experiment to investigate further
  - maybe a larger study or a **distinct functional/biological validation**
  - Of course, we could be wrong!



**Karolinska  
Institutet**

- In science, the dilemma occurs if we assume we must make a decision (about the true state of nature).
- **Does not mean** we do not need to make multiplicity adjustment!
  - There are obvious situations where we do want to limit false positives.
  - Other methods than Bonferroni.



**Karolinska  
Institutet**

- Easy trade-off: False positive vs False negative
- Example:
  - Search with true discovery rate 0.000 001 ( $10^{-6}$ ).  
Would you do it?

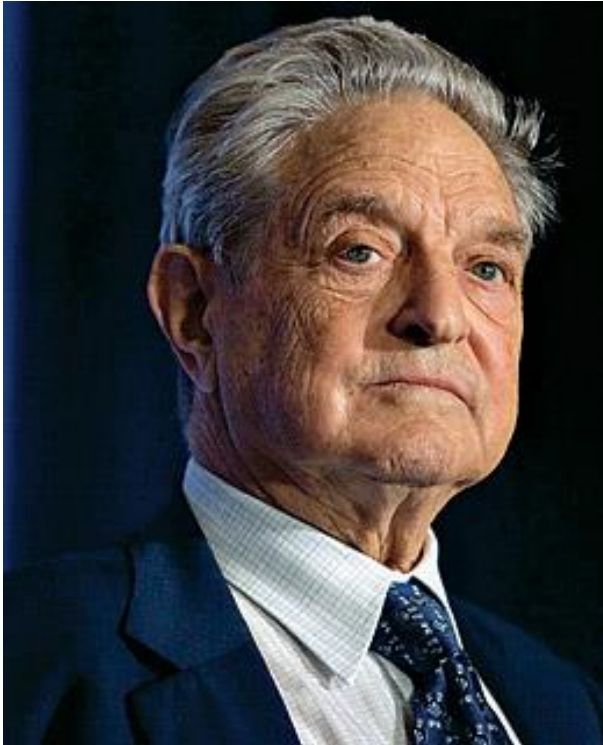


**Karolinska  
Institutet**



# George Soros

billionaire-investor, philosopher



- It is not whether you are right or wrong that is important, but
  - how much money you make when you are right, and
  - how much you lose when you are wrong.



- How much false positive rate can/should a scientist or a scientific field endure?
  - Tradition, level of activity,
  - cost of past failures
  - number of leads to potential discoveries
  - **value** of potential discoveries
- Case: molecular epidemiology



**Karolinska  
Institutet**

# Molecular epidemiology

- Candidate gene era (1980s-early 2000s)
  - Motivated by biology
  - Most discoveries were not validated in subsequent studies:
  - winner's curse in an active research field: under the null, 1 in 20 independent research groups can still **legitimately** produce/publish false discovery.



Karolinska  
Institutet

# Molecular epidemiology

- Genome-wide association study (GWAS) era
  - ~1 million tests per study
  - Standard Bonferroni correction ( $p < 10^{-8}$ )
  - Subsequent analyses to find lower-ranking signal
  - A large study eg Framingham: 100s phenotypes (different papers, different researchers), not adjusted for multiplicity.



**Karolinska  
Institutet**

# Reasonable requirements for dealing with multiplicity (Goeman & Solari, Stat Sci 2011)

- not too strict (unlike FWER approach)
  - allows false rejections
- Posthoc
  - Allows choice after seeing the data
  - 'cherry-picking'
- Flexible
  - Allows choice of whatever results to pursue (e.g. not just significant ones)



# Example: Table 1

**Table 1. p-values in a study of a metastatic cancer drug vs placebo for ten patient characteristics.**

<b>Variables</b>	<b>p-value</b>	<b>10<sup>4</sup>p-value</b>
1 Karnofsky index	0.007	0.07
2 Body weight	0.013	0.13
3 Tricep skin-fold	0.091	0.91
4 Hemoglobin concentration	0.236	2.36
5 Erythr sedimentation rate	0.350	3.50
6 Albumin in serum	0.525	5.25
7 Creatinine in serum	0.535	5.35
8 Bilirubin in serum	0.662	6.62
9 S-alkaline phosphatase	0.823	8.23
10 Alanine aminotransferase	0.908	9.08

**For top 3 variables** (using 'cherry' procedure)

- Estimated number of true discovery:  $3 - 0.91 = 2.09$
- 95% probability: number  $\geq 2$

# ABO blood-groups and cancers

G Edgren, V Kandaswamy, J Hwang (MEB)

Cancer Groups	A				AB		
	Events	IRR (95%CI)	Crude P value	FDR	Events	IRR (95%CI)	Crude p value
Lip	85	0.80 (0.62, 1.03)	0,092	0,353	13	0.10 (0.56, 1.54)	0,861
Tongue	202	1.11 (0.94, 1.32)	0,200	0,506	20	0.89 (0.59, 1.29)	0,577
Salivary glands	131	<b>1.36 (1.07, 1.74)</b>	<b>0,011</b>	0,200	16	1.33 (0.80, 2.09)	0,230
Mouth	195	1.01 (0.82, 1.23)	0,959	0,791	34	<b>1.49 (1.03, 2.01)</b>	<b>0,026</b>
Pharynx	330	<b>0.81 (0.69, 0.96)</b>	<b>0,013</b>	0,209	45	0.91 (0.64, 1.26)	0,597
Oesophageal adenocarcinoma	175	<b>0.76 (0.60, 0.96)</b>	<b>0,023</b>	0,250	23	0.81 (0.48, 1.28)	0,401
Oesophageal SCC	154	0.96 (0.77, 1.19)	0,723	0,741	18	0.88 (0.53, 1.37)	0,608
Stomach	703	<b>1.12 (1.01, 1.23)</b>	<b>0,021</b>	0,245	95	1.19 (0.97, 1.44)	0,073
Small intestine	169	<b>0.82 (0.67, 0.99)</b>	<b>0,040</b>	0,580	21	0.80 (0.52, 1.19)	0,306
Colon, incl. recto sigmoid	2514	0.99 (0.94, 1.05)	0,904	0,782	317	0.98 (0.87, 1.08)	0,667
Rectum, excl. Anus	1792	1.01 (0.95, 1.07)	0,748	0,748	236	1.06 (0.92, 1.19)	0,407
Anal cancers	111	0.87 (0.68, 1.12)	0,287	0,569	20	1.44 (0.92, 2.16)	0,090
Liver	575	1.11 (0.98, 1.25)	0,084	0,348	65	0.97 (0.74, 1.24)	0,688
Gallbladder, bil.pass., amp. Vater	225	1.01 (0.84, 1.20)	0,900	0,781	24	0.82 (0.54, 1.20)	0,346
Pancreas	1113	<b>1.37 (1.23, 1.52)</b>	<b>0,000</b>	<b>0,000</b>	137	<b>1.31 (1.05, 1.61)</b>	<b>0,012</b>
Peritoneum and unspecified	71	1.19 (0.87, 1.63)	0,272	0,560	8	1.05 (0.51, 1.93)	0,879
Nasal cavities, and sinuses	82	1.14 (0.84, 1.56)	0,394	0,617	14	1.53 (0.85, 2.57)	0,126
Larynx	221	0.94 (0.55, 1.60)	0,831	0,767	30	1.05 (0.30, 2.70)	0,926
Lung and tracheae	3353	0.99 (0.95, 1.04)	0,826	0,766	423	0.98 (0.89, 1.07)	0,643
Pleura	161	1.01 (0.81, 1.24)	0,958	0,791	9	<b>0.44 (0.21, 0.80)</b>	<b>0,014</b>
Mediastinum	157	1.03 (0.73, 1.46)	0,838	0,768	10	0.53 (0.17, 1.23)	0,196
Breast	6688	<b>1.06 (1.02, 1.09)</b>	<b>0,001</b>	0,072	892	<b>1.07 (1.00, 1.09)</b>	<b>0,047</b>
Cervix uteri	647	1.09 (0.98, 1.23)	0,122	0,410	80	1.08 (0.85, 1.36)	0,504
Corpus uteri	783	1.01 (0.92, 1.11)	0,712	0,738	128	<b>1.23 (1.03, 1.45)</b>	<b>0,018</b>
Uterus, other parts and unspec	69	1.06 (0.95, 1.17)	0,247	0,544	11	1.03 (0.83, 1.27)	0,735
Prostate	9134	1.00 (0.97, 1.03)	0,750	0,748	1148	1.02 (0.96, 1.08)	0,369

Number of tests =  $45 \times 3 = 135$

	A	B	C	D	E	F
1		Group	IRR (95%CI)	P value	AdjPval	
2	Pancreas	A.	1.37 (1.23, 1.52)	0,000	0,00	
3	Pancreas	AB.	1.31 (1.05, 1.61)	0,012	1,57	
4	Breast	A.	1.06 (1.02, 1.09)	0,001	0,18	
5	Breast	AB.	1.07 (1.00, 1.09)	0,047	6,35	
6	Bladder incl. p	A.	1.09 (1.03, 1.16)	0,005	0,69	
7	Salivary gland	A.	1.36 (1.07, 1.74)	0,011	1,49	
8	Pharynx	A.	0.81 (0.69, 0.96)	0,013	1,77	
9	Pleura	AB.	0.44 (0.21, 0.80)	0,014	1,85	
10	Corpus uteri	AB.	1.23 (1.03, 1.45)	0,018	2,48	
11	Stomach	A.	1.12 (1.01, 1.23)	0,021	2,89	
12	Oesophageal	A.	0.76 (0.60, 0.96)	0,023	3,13	
13	Mouth	AB.	1.49 (1.03, 2.01)	0,026	3,56	

	A	B	C	D	E	F	G
1		Group	IRR (95%CI)	P value	AdjPval	FDR	Validation
2	Pancreas	A.	1.37 (1.23, 1.52)	0,000	0,00	0,00	0,000
3	Pancreas	AB.	1.31 (1.05, 1.61)	0,012	1,57	0,20	0,000
4	Breast	A.	1.06 (1.02, 1.09)	0,001	0,18	0,07	0,000
5	Breast	AB.	1.07 (1.00, 1.09)	0,047	6,35	0,31	0,000
6	Bladder incl. p	A.	1.09 (1.03, 1.16)	0,005	0,69	0,16	0,487
7	Salivary gland	A.	1.36 (1.07, 1.74)	0,011	1,49	0,20	0,024
8	Pharynx	A.	0.81 (0.69, 0.96)	0,013	1,77	0,21	0,002
9	Pleura	AB.	0.44 (0.21, 0.80)	0,014	1,85	0,21	0,522
10	Corpus uteri	AB.	1.23 (1.03, 1.45)	0,018	2,48	0,23	0,237
11	Stomach	A.	1.12 (1.01, 1.23)	0,021	2,89	0,24	0,000
12	Oesophageal	A.	0.76 (0.60, 0.96)	0,023	3,13	0,25	0,037
13	Mouth	AB.	1.49 (1.03, 2.01)	0,026	3,56	0,26	0,000
14							

**Top 12 results:**

- Estimated FDR ~ 0.26 or TDR~0.74
- Validated externally: 9/12 = 0.75



# Conclusions

- Multiplicity problem has multiplicity of perspectives and solutions
- Multiplicity is not just a 'problem', leading to less-significant results, but also an opportunity to discover more
- Consider: FP, FN, adjusted p-values, FDR
- Current methods are becoming more flexible and more informative.



**Karolinska  
Institutet**

- Problem: communication of science/discoveries to the public:
  - Messy, uncertain, provisional results that could be wrong are acceptable to scientists
  - Black-white simplified views of the 'public'
- This problem is not unique to science:
  - Nearby/familiar problems allow nuances
  - Faraway problems get simplistic thoughts
  - Eg socio-political problems in far-away places: scientists are the 'public' here.
- **SO: may have to live with controversies**



**Karolinska  
Institutet**



transgenerational epigenetic inheritance



Webben Bilder Videor Nyheter Kartor Böcker

Ungefär 75 300 resultat

Alla länder

Land: Sverige

Alla språk

Sidor skrivna på svenska

När som helst

- Senaste timmen
- Senaste dygnet
- Senaste veckan
- Senaste månaden
- Senaste året

Alla resultat

Ordagrant

[Vetenskapliga artiklar med "transgenerational epigenetic inheritance"](#)

[Transgenerational epigenetic inheritance](#) - Rakyan - Citerat av 100

[Transgenerational epigenetic inheritance: prevalence ...](#) - Jablonka - Citerat av 581

[... transgenerational epigenetic inheritance via the ...](#) - Daxinger - Citerat av 209

[Transgenerational epigenetics](#) - Wikipedia, the free encyclopedia

[en.wikipedia.org/wiki/Transgenerational\\_epigenetics](http://en.wikipedia.org/wiki/Transgenerational_epigenetics) v

Main article: **Epigenetics**. Epigenetic variation may take one of four general forms . Others may yet be ...

[Major controversies in the ...](#) - Origin and inheritance of epigenes

[Transgenerational Epigenetic Inheritance: Myths and Mechanisms ...](#)

[www.cell.com/abstract/S0092-8674\(14\)00286-4](http://www.cell.com/abstract/S0092-8674(14)00286-4) v

Since the human genome was sequenced, the term "epigenetics" is increasingly being associated with the hope that we are more than just the sum of our genes ...

[Transgenerational epigenetic inheritance: More questions than ...](#)

[www.ncbi.nlm.nih.gov/pmc/articles/PMC2989988/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2989988/) v

This phenomenon is referred to as **transgenerational epigenetic inheritance**. Moreover, recent evidence shows that the environment can stably influence the ...

[Bilder på transgenerational epigenetic inheritance](#)

# Starvation-Induced Transgenerational Inheritance of Small RNAs in *C. elegans*

Oded Rechavi,<sup>1,2,\*</sup> Leah Hour-Ze'evi,<sup>1</sup> Sarit Anava,<sup>1</sup> Wee Siong Sho Goh,<sup>3</sup> Sze Yen Kerk,<sup>2</sup> Gregory J. Hannon,<sup>3</sup> and Oliver Hobert<sup>2</sup>

<sup>1</sup>Department of Neurobiology, Wise Faculty of Life Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel

<sup>2</sup>Columbia University Medical Center, Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, New York, NY 10032, USA

<sup>3</sup>Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, Howard Hughes Medical Institute, New York 11724, USA

\*Correspondence: [odedrechavi@gmail.com](mailto:odedrechavi@gmail.com)

<http://dx.doi.org/10.1016/j.cell.2014.06.020>

## SUMMARY

Evidence from animal studies and human famines suggests that starvation may affect the health of the progeny of famished individuals. However, it is not clear whether starvation affects only immediate offspring or has lasting effects; it is also unclear how such epigenetic information is inherited. Small RNA-induced gene silencing can persist over several generations via transgenerationally inherited small RNA molecules in *C. elegans*, but all known transgenerational silencing responses are directed against foreign DNA introduced into the organism. We found that starvation-induced developmental arrest, a natural and drastic environmental change, leads to the generation of small RNAs that are inherited through at least three consecutive generations. These small, endogenous, transgenerationally transmitted RNAs

epigenetic information about food availability during development before the prepubertal peak in growth speed (Bygren et al., 2001; Kaati et al., 2007). In all the above-mentioned studies the effects were restricted to the immediate offspring, leaving the possibility open that rather than being transgenerationally inherited, the effects were directly exerted onto the germ cells of the exposed animals (Heard and Martienssen, 2014). Thus, a truly epigenetic effect that could transmit the somatic response to satiety or famine to the generations beyond the immediate next generation remains to be found.

The model organism *Caenorhabditis elegans* has been successfully used for the study of transgenerational epigenetic effects (Greer et al., 2011; Rechavi et al., 2011), in part because of its short generation time, *C. elegans* is particularly suited for studying transgenerational inheritance of dietary history, since worms are often faced with scarcity of nutrients in the wild. In fact, a dedicated genetic program allows worms to reversibly arrest postembryonic development in the first larval stage (L1) in the absence of food (Baugh, 2013). Similar to the situation in hu-



## FIGURE 1 | Alternative forms of transgenerational epigenetic inheritance.

FROM THE FOLLOWING ARTICLE:

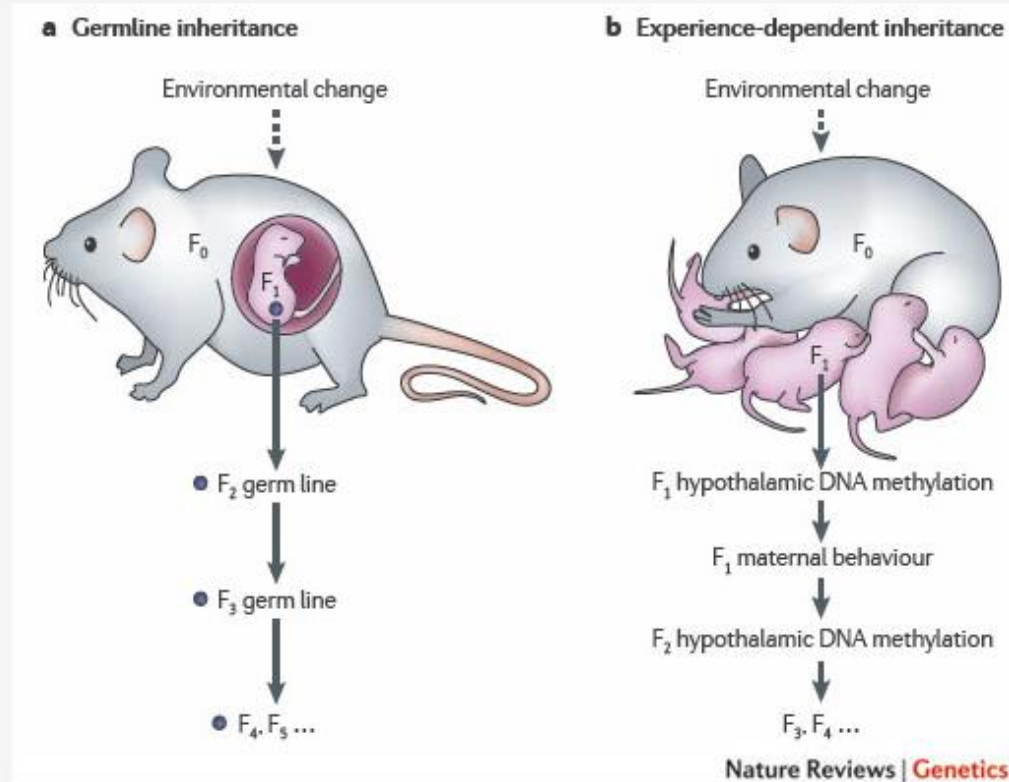
**Beyond DNA: integrating inclusive inheritance into an extended theory of evolution**

Étienne Danchin, Anne Charmantier, Frances A. Champagne, Alex Mesoudi, Benoit Pujol & Simon Blanchet

*Nature Reviews Genetics* **12**, 475-486 (July 2011)

doi:10.1038/nrg3028

< [Back to article](#) | < [Back to figures and tables](#) | [Next figure](#) >



**a** | In germline epigenetic inheritance, an environmental effect occurring during development results in an epigenetic change within the first filial generation (F<sub>1</sub>)

# Discover

SCIENCE FOR THE CURIOUS

 **SEARCH**

THE MAGAZINE | BLOGS | **HEALTH & MEDICINE** | MIND & BRAIN | TECHNOLOGY | SPACE & PHYSICS | LIVING WORLD | ENVIRO

**TOPICS** Personal Health | Aging | Sex & Reproduction | Biotechnology | Genes & Health | Vaccines & Drugs | Infectious D  
Medical Technology | Cancer | Nutrition | Mental Health | Obesity

Home » May » **Grandma's Experiences Leave a Mark on Your Genes**

FROM THE MAY 2013 ISSUE

## Grandma's Experiences Leave a Mark on Your Genes



stors' lousy childhoods or excellent adventures

**NEW ON DISCOVER**

Was Neuroscience's Most Famous Amnesiac, "HM", A Victim of Medical Error?

Introducing  
**Subscriber-Only Online Exclusives!**

Discover  
 MAGAZINE

**Subscribe Now »**

P23698

Home » May » **Grandma's Experiences Leave a Mark on Your Genes**

FROM THE MAY 2013 ISSUE

# Grandma's Experiences Leave a Mark on Your Genes

Your ancestors' lousy childhoods or excellent adventures might change your personality, bequeathing anxiety or resilience by altering the epigenetic expressions of genes in the brain.

By Dan Hurley | Tuesday, June 11, 2013

RELATED TAGS: **GENES & HEALTH**



**NEW ON  
 DISCOVER**

Was  
 Neuroscience's  
 Most Famous  
 Amnesiac, "HM",  
 A Victim of  
 Medical Error?

Praying  
 Mantises are  
 Expert Acrobats  
 in Mid-Air

Handshakes'  
 Purpose Could  
 Be to Send Scent  
 Signals

Lots of Cases of  
 Synesthesia Are  
 Based on  
 Alphabet  
 Magnets

**Discover's**

Sign up to get  
 delivered weekly

e-mail address