Conductome

What I learned from the conference talks so far…

Applications to Obesity and Metabolic Disease of the Physiolome and Conductome

Christopher R. Stephens

C3 – Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México

Third International Summer Institute

Lake Como School of Advanced Studies – 24-29 July 2022

Highlights from the first talk…

How many things can I say about a system?

1. An ”omic” perspective is associated with the concept of having a totality of things
   a) Structures (direct or indirect) – genes (genome), proteins (proteome), organs-tissues-etc (Physiolome),…
   b) Behaviours (ethome)
   c) Predictors of a behaviour (Conductome)
2. Two ways to exhibit relationships in an “ome” are:
   a) A network perspective - $P(X_1, X_2,\ldots, X_i,\ldots, X_N)$ $\{X_1, X_2,\ldots, X_i,\ldots, X_N, E(X_1, X_2), E(X_1, X_3),\ldots, E(X_i, X_j),\ldots, E(X_{N-1}, X_N)\}$
   b) A “niche” perspective - $P(C|X_1,X_2,\ldots,X_N)$ (can also be represented as a “network”)
3. Almost anything that concerns living beings is immensely multi-factorial from any perspective
   a) Whether we can ignore it is a question of scale and model accuracy, as well as “subjective” criteria
   b) Behaviours are extremely multi-factorial, with predictors ranging from the genetic to the political-cultural
   c) Accounting for and representing this multi-factoriality is an immense challenge in data acquisition, data integration, data modeling and data visualization. It requires Hybrid Intelligence (Human Intelligence + Artificial Intelligence)
4. Conducts and their adverse consequences are at the heart of every major human problem
   a) Obesity and most chronic disease, global warming, poverty, war and violence, discrimination,…

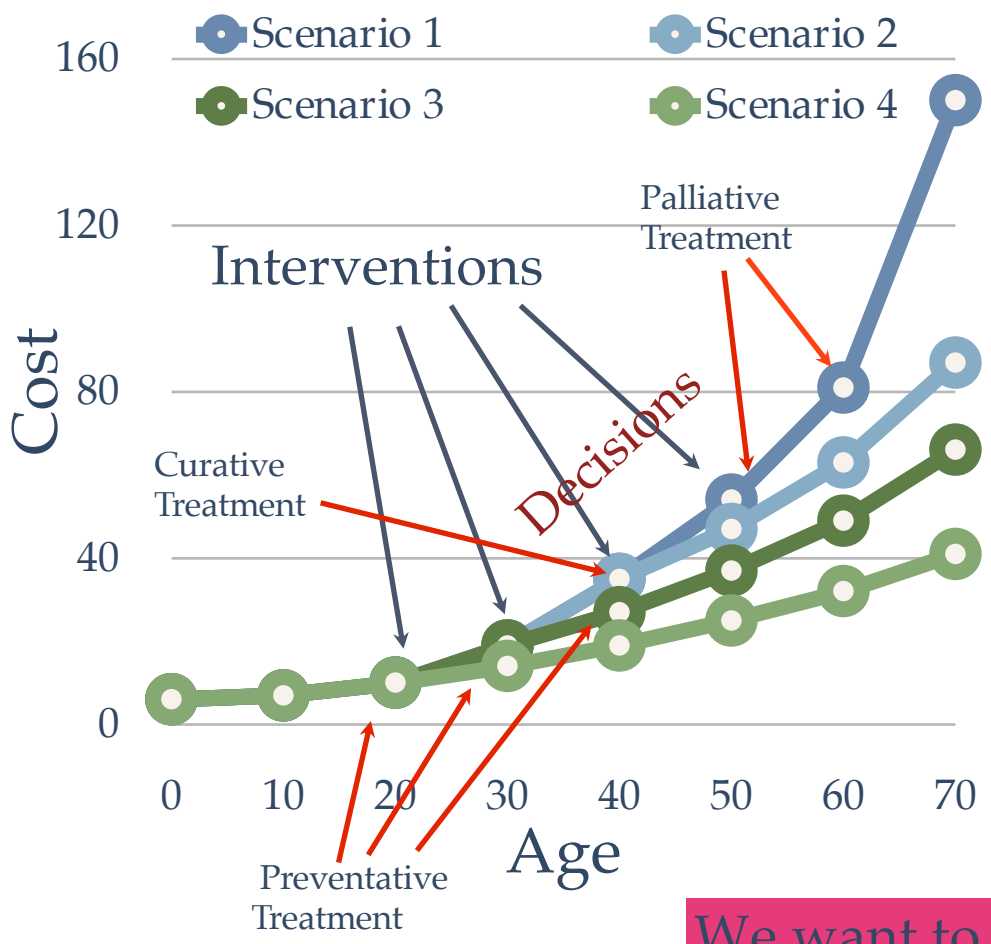
We don’t take multi-factoriality seriously (enough)

Decision making in a Complex Adaptive System…

At any given point in time, we can define a clinical or epidemiological state C(t) and calculate P(C(t)|**X**(t')) using a Bayesian ML algorithm.

The main challenge is getting the data to represent the multitude of Xs that are relevant.

A second challenge is how to distinguish correlation from causation and identify those actionable factors that can be used to reduce P



Scenario 1: Onset of obesity at 20
Continued obesity and onset of metabolic syndrome at 40
Onset of diabetes at 50
Onset of renal failure at 60
Death at 70

Scenario 2: Onset of obesity at 20
Continued obesity and onset of metabolic syndrome at 40
Onset of diabetes at 50
Adequate control and treatment of comorbilities at 60
Bad health and high cost at 70

Scenario 3: Overweight at 20
Obesity at 40
Onset metabolic syndrome at 50
Onset of diabetes at 60
Continued diabetes but no serious comorbilities at 60
Ill health and moderate cost at 70

Scenario 4: Overweight at 30
Obesity at 50
Onset of metabolic syndrome at 60
Onset of diabetes at 70 but relative health

We want to predict and understand "histories"

Why is there obesity?

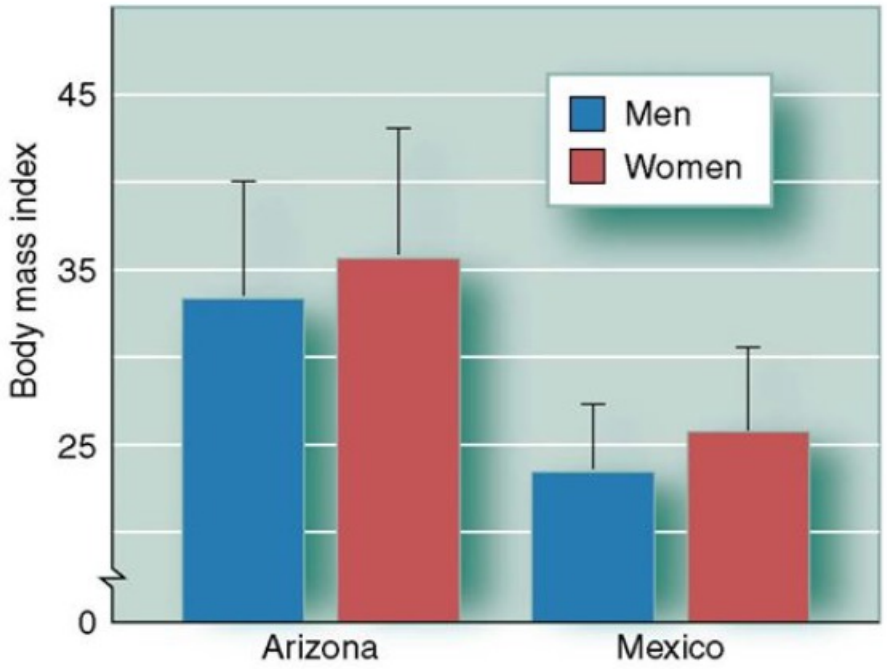
Nature versus nurture versus environment

Nature?

TABLE 2
Intraclass twin correlations (and 95% CIs) for BMI and waist circumference SD scores[1]

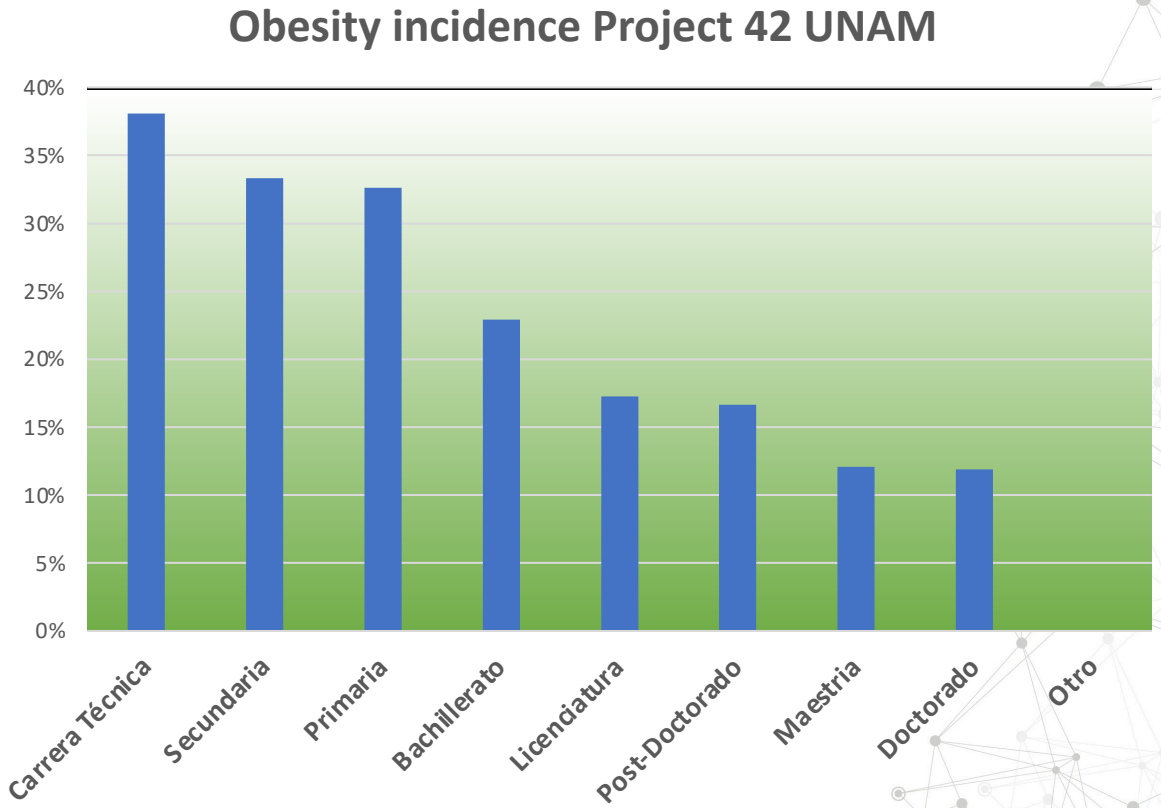
Measure	MZall (n = 1813)	DZall (n = 3279)	DZss (n = 1658)	DZos (n = 1621)
BMI	0.86 (0.85, 0.87)	0.49 (0.47, 0.51)	0.51 (0.48, 0.53)	0.47 (0.45, 0.50)
Waist	0.85 (0.84, 0.86)	0.48 (0.46, 0.50)	0.51 (0.49, 0.54)	0.45 (0.42, 0.48)

[1]MZall, monozygotic twins; DZall, dizygotic same-sex and opposite-sex twins; DZss, same-sex dizygotic twins; twins; DZM, dizygotic male twins; DZF, dizygotic female twins. All values were significant, $P < 0.001$.



The BMI for Pima Indian men and women living in Arizona and in northern Mexico (2006)

Nurture + environment?



Also applies to COVID, cancer,…

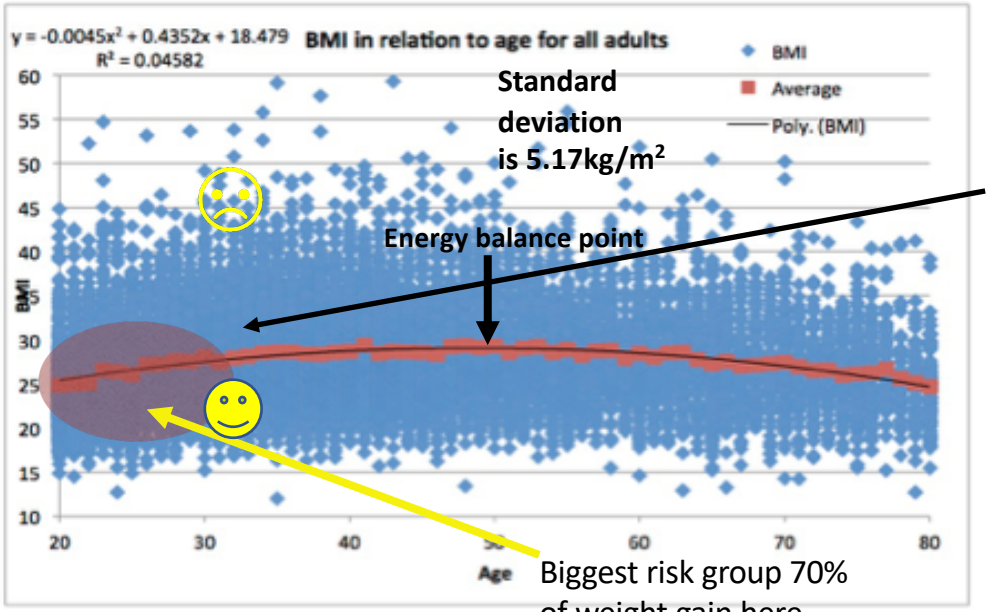
You can't gain weight without an associated set of decisions/actions that correspond to a behaviour

1. What are some of those behaviours?
2. How do we quantify/measure them?
3. What are risk factors for those behaviours?
4. How plastic are they?
5. How do we model them?
6. How do we change them?

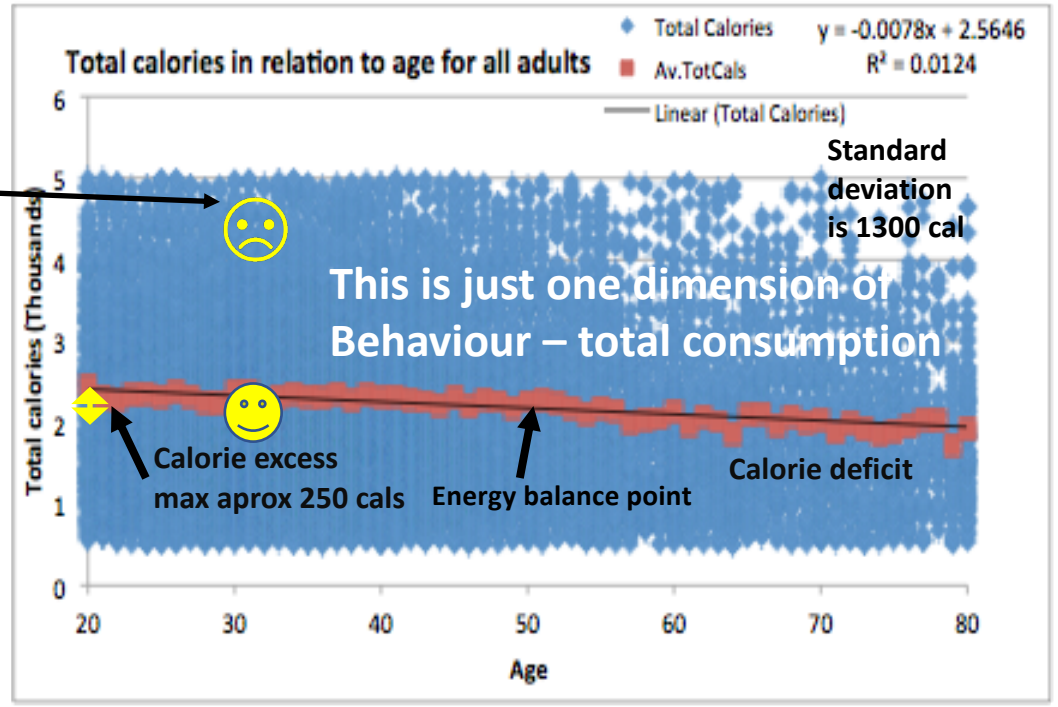
1) What are some of those behaviours? Consumption and Exercise

1. Consumption...

You aren´t what you eat you become what you eat



This isn't noise its multifactoriality



Epidemiological data from ENSANUT 2006

Regression of BMI change versus calorie excess

	Variable(s)	Unstd. B	Std. Error	t	f	R^2	Sig	Lower	Upper
Moving Av.					29.236	0.343	0		
BMI Change	Constant	-1.954	0.362	-5.392			0	-2.68	-1.228
ALL	Total_Cals	0.904	0.167	5.407			0	0.569	1.239
	Variable(s)	Unstd. B	Std. Error	t	f	R^2	Sig	Lower	Upper
Moving Av.					13.397	0.193	0.001		
BMI Change	Constant	-1.625	0.444	-3.656			0.001	-2.515	-0.734
Men	Total_Cals	0.724	0.198	3.66			0.001	0.328	1.121
	Variable(s)	Unstd. B	Std. Error	t	f	R^2	Sig	Lower	Upper
Moving Av.					22.429	0.286	0		
BMI Change	Constant	-1.754	0.372	-4.711			0	-2.5	-1.008
Women	Total_Cals	0.833	0.176	4.736			0	0.481	1.185

This gradually decreasing calorie excess seems to be the motor for the population level increase in BMI

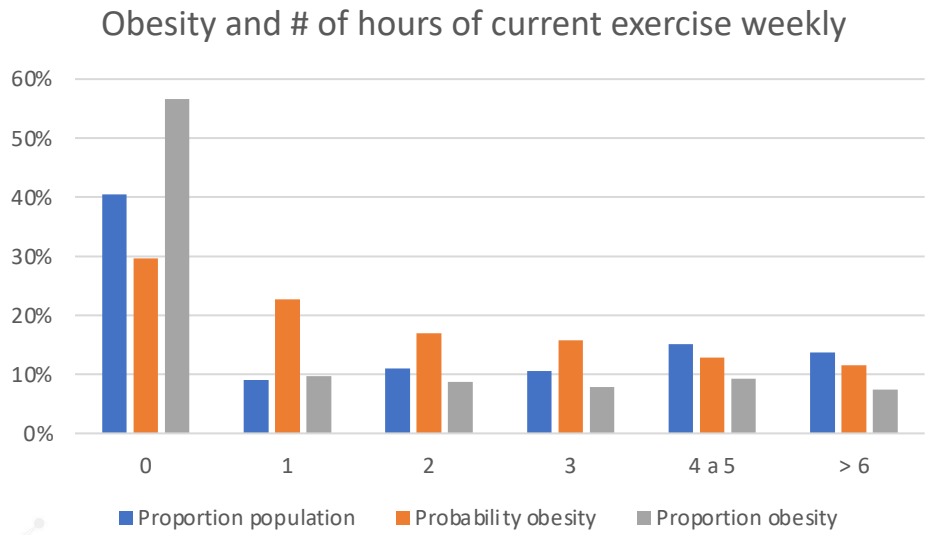
You aren't what you eat you become what you eat
Stephens, Easton and Sicilia
medRxiv https://doi.org/10.1101/2021.08.23.21262191

2. Exercise ...

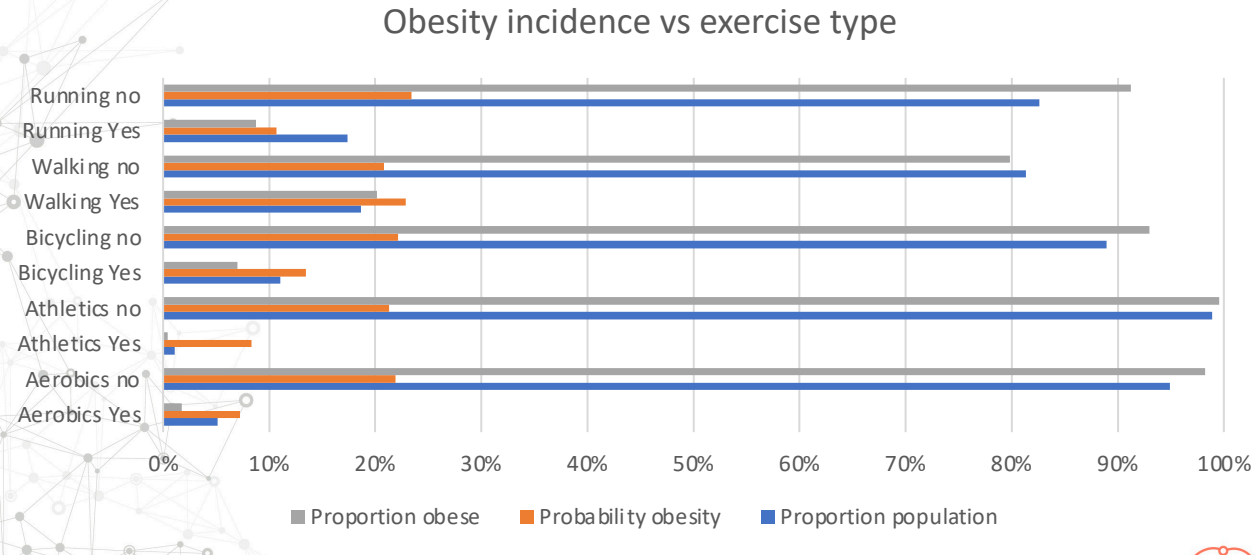
Obesity % versus historical exercise behavior
A > recommended exercise, B < recommended, * don't care;
(30y, 20y, 10y, 5y, 1y, now)

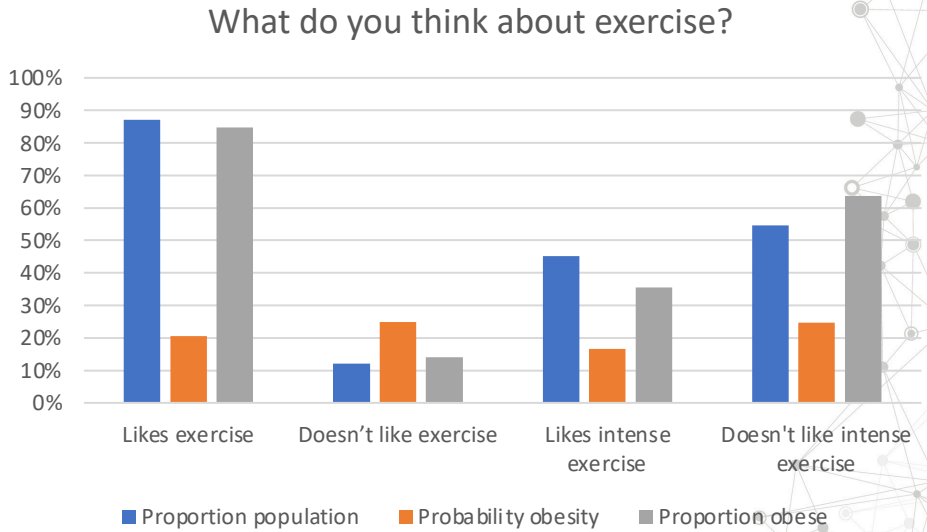
History	ϵ	N_x	N_{cx}	%	score
A*A*BB	3.56	94	38	40.43	0.73
AAA*B	3.55	91	37	40.66	0.74
AA**BB	3.53	113	44	38.94	0.67
AA**B*	3.40	131	49	37.40	0.60
A***BB	3.23	137	50	36.50	0.57
*A***A	-3.27	157	21	13.38	-0.75
**AAA	-3.27	157	21	13.38	-0.75
AA**AA	-3.51	103	10	9.71	-1.11
A**AA	-3.61	134	15	11.19	-0.95
***AA	-3.76	193	25	12.95	-0.79

Its worse to have had good habits and lost them than never to have had them

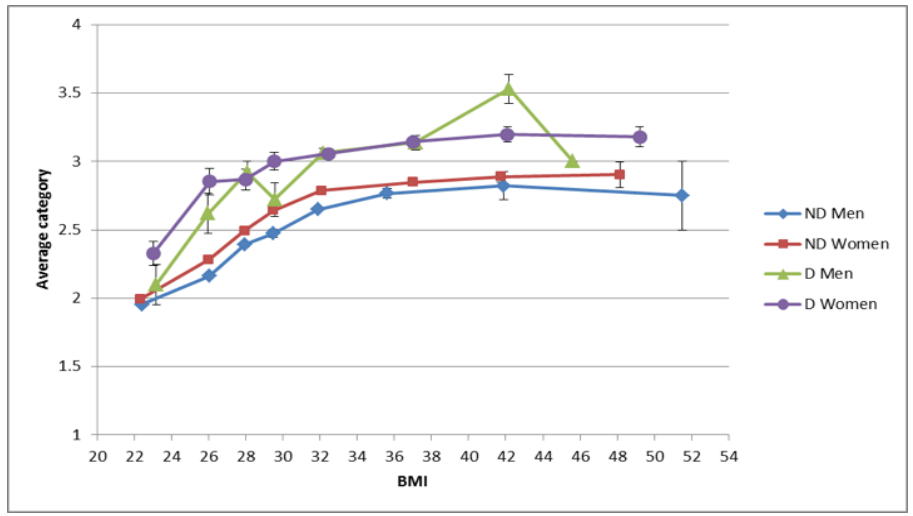


How many dimensions do we need to describe our decision making/actions and behavior with respect to exercise?





3. And some risk factors… being short, being non-academic, looking at the world through rose-tinted glasses,…

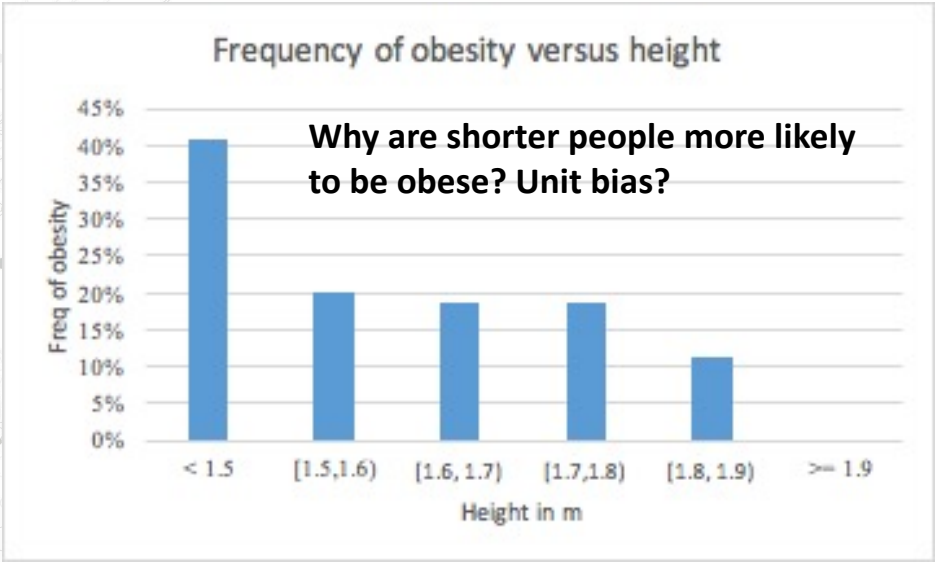


Probability to be an academic versus historical exercise behavior
A > recommended exercise, B < recommended, * don't care;
(30y, 20y, 10y, 5y, 1y, now)

History	ϵ	N_x	N_{cx}	%	score
*A***A	5.55	157	85	54.14	0.86
A**AA	5.21	134	73	54.48	0.88
AA**A	5.13	135	73	54.07	0.86
A*A*A	5.06	129	70	54.26	0.87
*A**A	4.97	165	85	51.52	0.76
*BBB**	-4.32	197	37	18.78	-0.77
**BB*	-4.40	267	55	20.60	-0.65
BBB	-4.41	207	39	18.84	-0.76
**BBB	-4.41	245	49	20.00	-0.69
**B*B	-4.55	260	52	20.00	-0.69

Effect of cognitive biases

- Self-serving
- Anchoring
- Unit



Big Mac meal for a large person

Big Mac meal for a short person

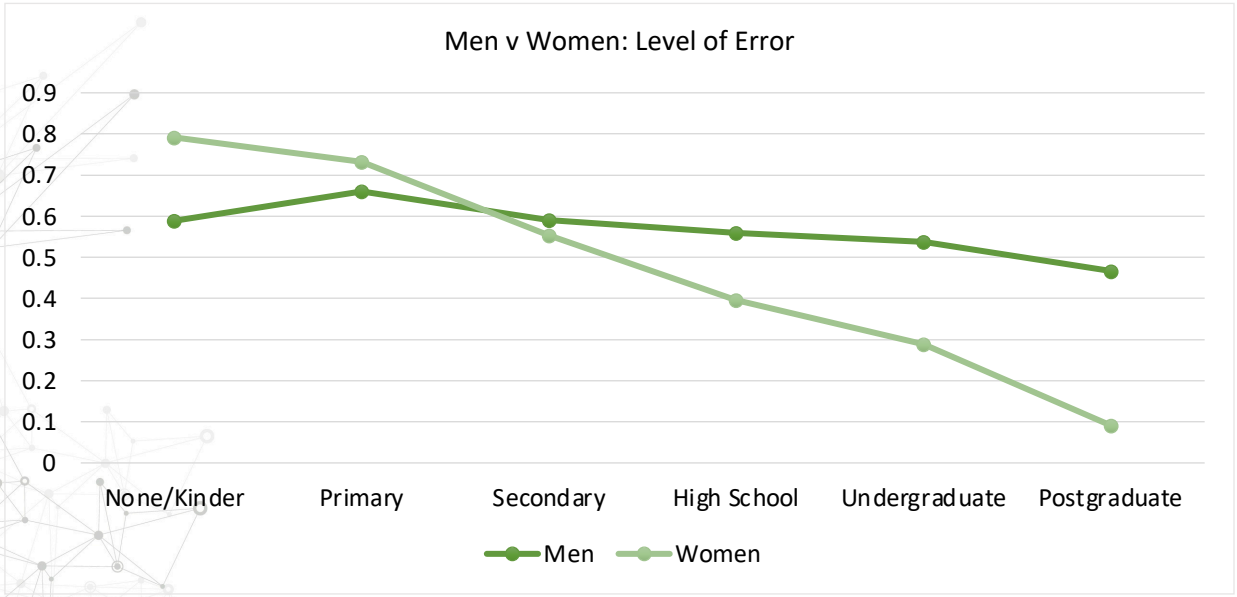
Torta Cubana for a large person

Torta Cubana for a short person

Easton, J.F., Stephens, C.R. & Román Sicilia, H. The effect of a medical opinion on self-perceptions of weight for Mexican adults: perception of change and cognitive biases. *BMC Obes* **4,** 16 (2017). https://doi.org/10.1186/s40608-017-0152-6

But we don’t even realize that overeating and sedentariness has adverse consequences

The difference between them depends on many factors, e.g. educational level
And has consequences…



1) “In the last year have you lost or gained weight?”
2) “Was this weight loss intentional?”

BMI Obese	Education level (n; %)					
Intention to lose	None/Kinder	Primary	Secondary	High School	Undergraduate	Postgraduate
All	17; 6.3	100; 7.1	61; 9.2	28; 10.9	24; 15.7	2; 25.0
Men	2; 3.9	23; 8.0	10; 6.4	10; 12.2	10; 16.1	1; 25.0
Women	15; 6.8	77; 6.8	51; 10.0	18; 10.3	14; 15.4	1; 25.0

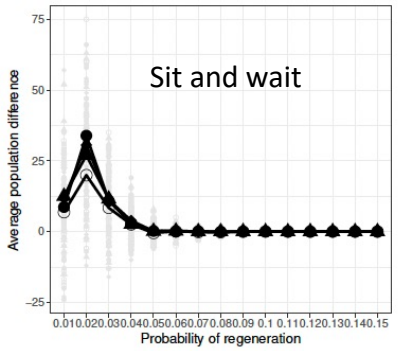
Conductome

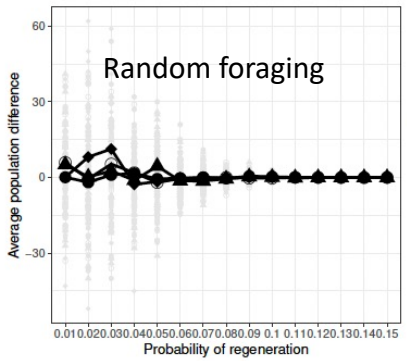
Why are we sedentary and overeat?

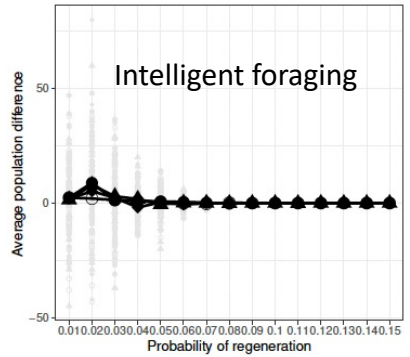
Hypothesis: there is an evolutionary advantage to overconsumption and sedentariness in certain food environments

Test: Agent based model where agents adopt different foraging strategies in an environment with a random distribution of food resources that regenerate with a fixed probability and where they have energy accumulation (fat storage) limits (testing the “thrifty gene”).

Conductome

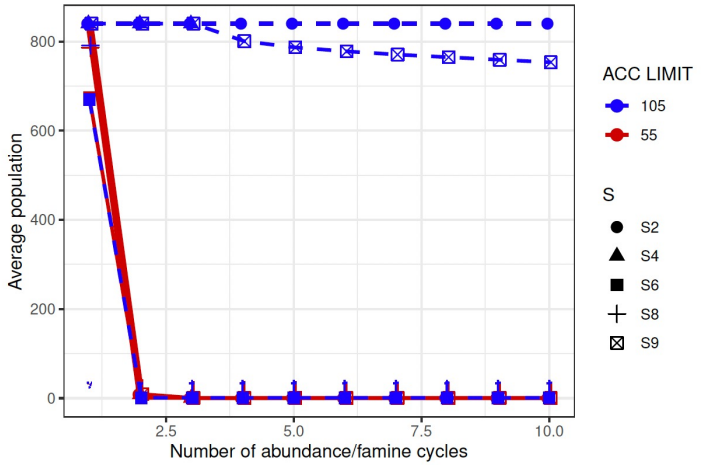




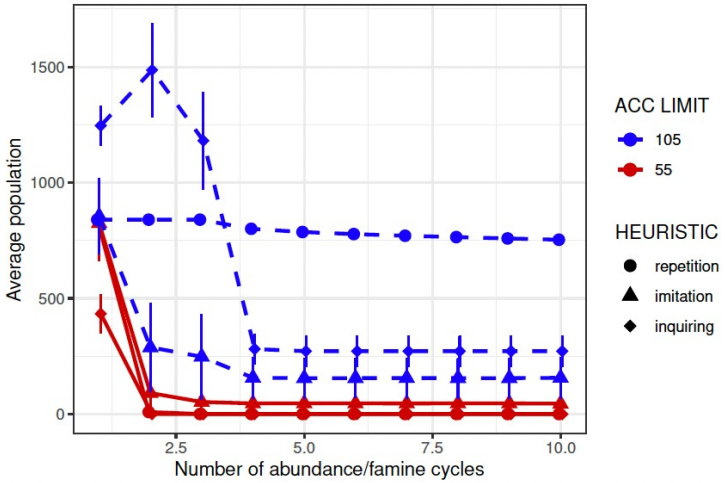


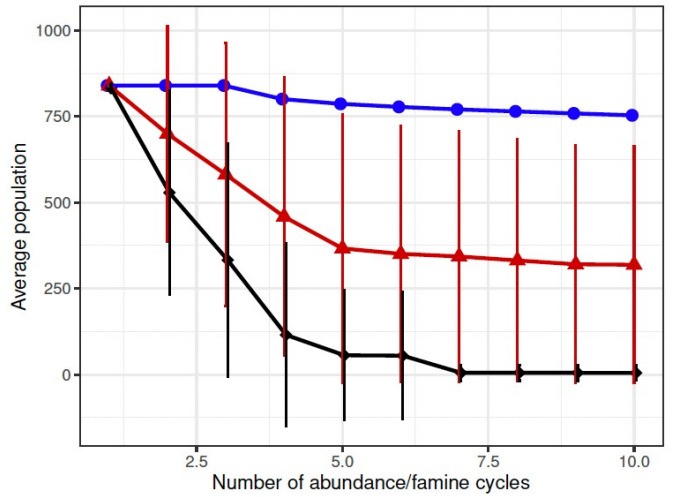
What happens in predictable feast-famine cycles?

There's no advantage to a higher energy accumulation limit (a "thrifty" gene) in environments with ample but neither is there an advantage in environments with very scarce resources. In fact, there's very little advantage in any constant resource environment.



Now a thrifty gene is preferred but only for certain "sedentary" foraging strategies – don't move if you don't have to





Uncertainty in the resource environment makes life much more difficult!

A higher accumulation limit can ONLY be taken advantage of if you eat more than you need.

Batta and Stephens, submitted to Scientific Reports

Project 42 for constructing the Conductome

First we need data

Goal: Construct the deepest data base on the planet (Deep Data, not Big Data) – to be publically available

Phase 0: National Epidemiological Surveys: ENSANUT 2006, 2012; ENCOPREVENIMSS
Phase I: (03-05/2014) 1,076 academics and non-academics from 12 institutes and faculties of the UNAM
2,524 variables - Genetic, epidemiological, physiological,… **Epidemiological**: Personal (81), Personal history (130), Family History (548), Self-health evaluation (226), Nutrition (220), Lifestyle (390), Health knowledge (293); **Genetic** (772); **Anthropometric and physiological** (49).
Phase II: (03/2017-09/2018) 500 medical students from the Fac. Med UNAM; (06/17) 100 workers and teachers from the FM. Addition of psychological variables.
Phase III: (12/2018-02/2019) 150 diabetics from the ISSTE
Phase IV: (01-03/2019) Follow up on 1,076 from Phase I and new population of academics and non-academics of the UNAM. Repetition of blood analysis, addition of psychological variables, detailed tracking of daily activities, actigraphy
Phase V: (01/19-current) Construction and publication of data base associated with Phases 1-4 with a Machine learning based analysis platform (project42.c3.unam.mx)
Phase VI: (01/2020-current) Population of 800 students from public and private universities – extensive psychological profiling
Phase VII: (03/2022-current) 3500 students from UNISA and Inst. Rosario Castellanos (first phase – "light" questionaire)
Phase VIII: (09/2022-) Metabolic biomarkers from Phase VI population, MRI data for subset of 100 participants, ghrelin and leptin biomarkers.

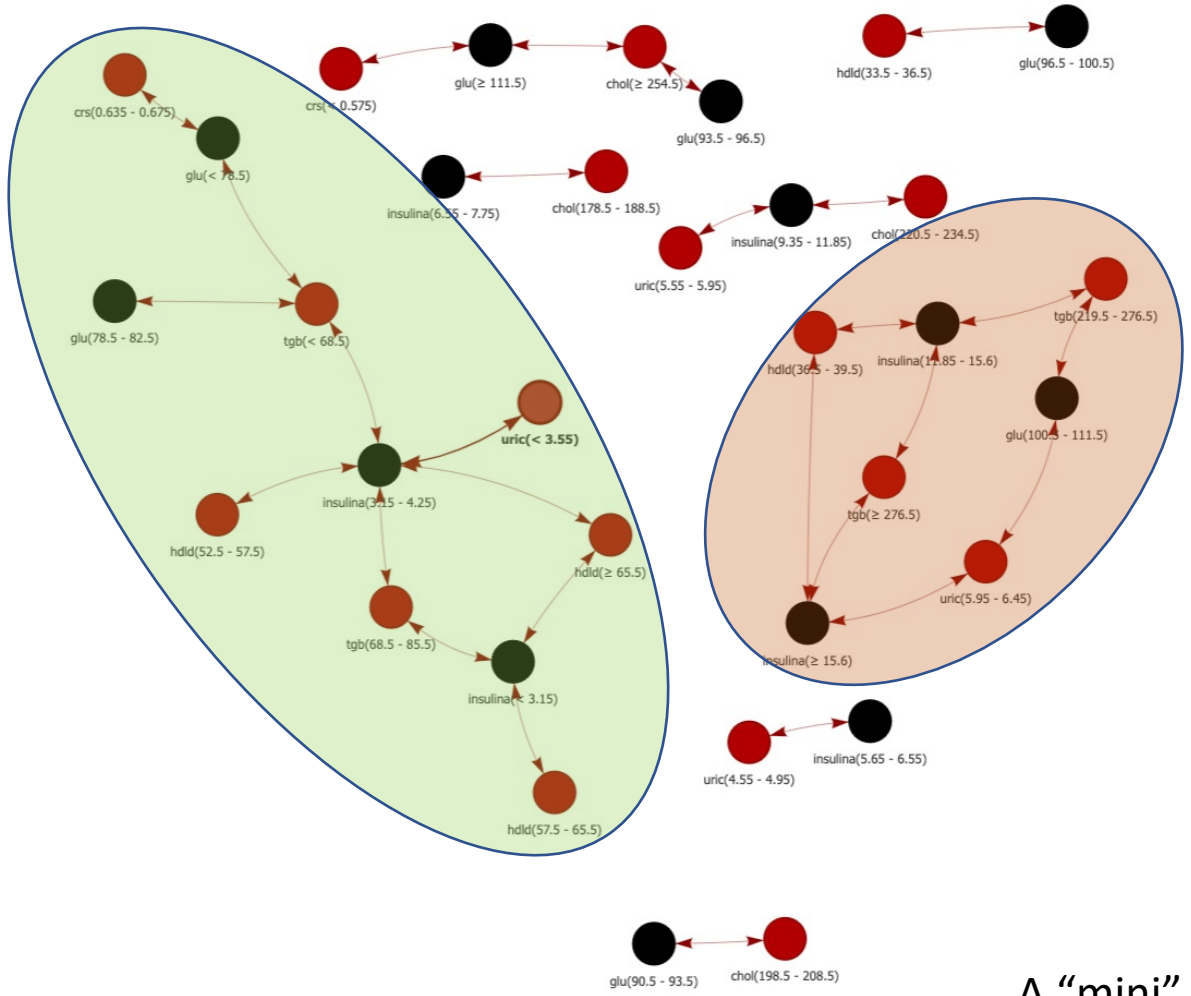
(CONACyT Fronteras 1093, CONACyT, Redes y PAPIIT, SECITI)

Conductome

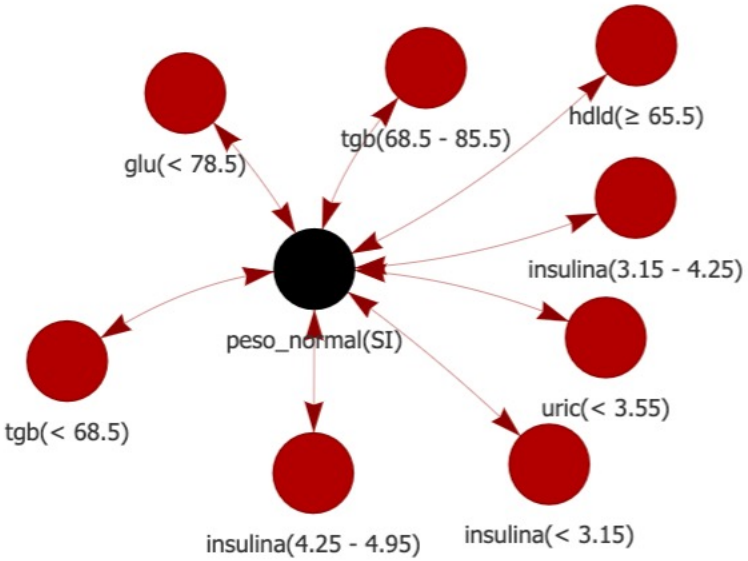
From the Phase I population:
Biomarkers – glu, tgb, chol, hdld, insulin, creatinine, uric acid; values discretized into deciles – top 10% of highest values of glu, next 10% etc.

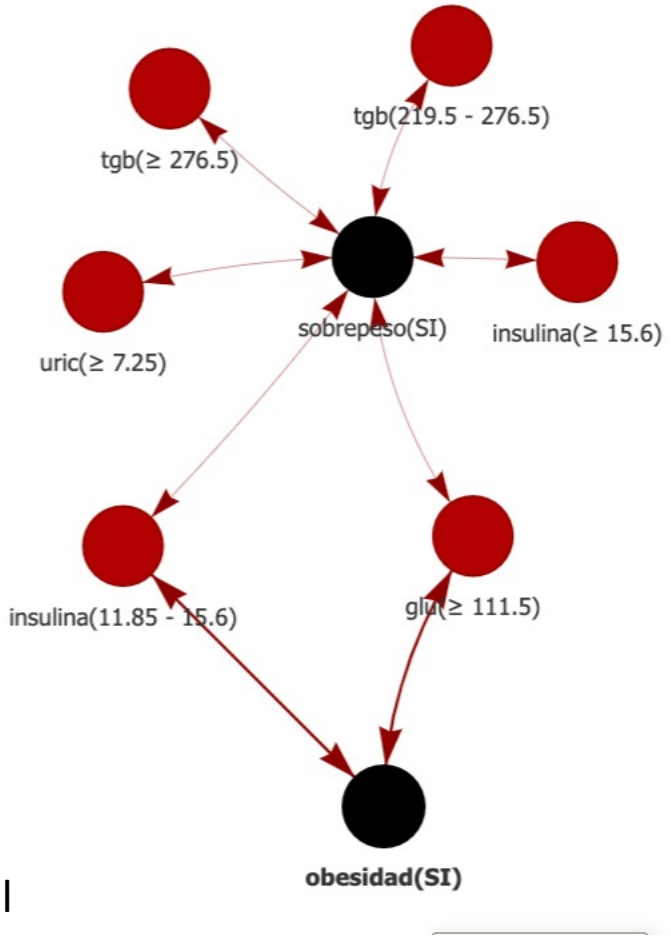
Links are a binomial test: $\varepsilon = \frac{N_X(P(C|X)-P(C))}{\sqrt{N_X P(C)(1-P(C))}}$

C is a black node and X a red one
Threshold on links = 2.62 std dev



A “mini” physiolome

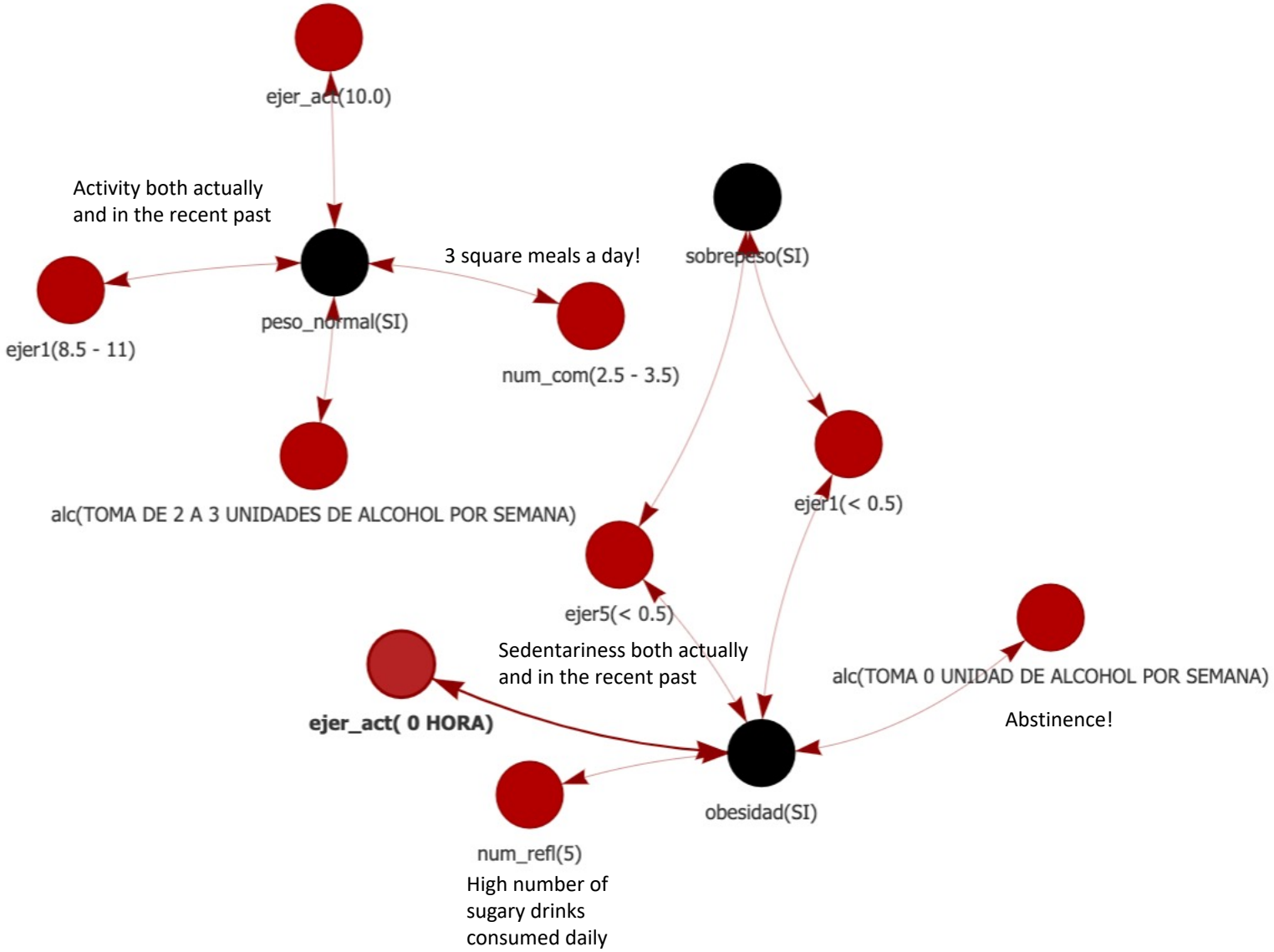




And now “anthropometrome” and the physiolome…

Epsilon threshold 3.65 std dev

Note the absence of chol



And the “anthropometrome” and the Conductome…

conductome.unam.mx