ARTICLES
https://doi.org/10.1038/s43016-020-0096-8

nature food

Check for updates

Animal-sourced foods are required for minimum-cost nutritionally adequate food patterns for the United States

Sylvia M. S. Chungchunlam[1]✉, Paul J. Moughan[1], Daniel P. Garrick[2,3] and Adam Drewnowski[4]

The amounts of animal-sourced foods required to achieve a least-cost nutritious diet depend on the food prices prevalent in each country. Using linear programming, we determine least-cost dietary patterns in the United States and the constituent amounts of animal-sourced foods. We considered local foods and prices from 2009–2010, and the average energy and nutrient requirements of adults. Nutrient-adequate food patterns were estimated at US$1.98 per day and included animal and plant products. Limiting nutrients were α-linolenic acid, potassium, choline, and vitamins C, D, E and K. The prices of animal-based foods had to be increased by 2–11.5 times to be excluded from the modelled food pattern, with the least cost of a plant-only diet at US$3.61. Given relative food prices in the United States, we show that animal-based foods are needed to secure adequate nutrition at the lowest cost, underscoring the role of price and market mechanisms in the choice of nutrient-adequate, sustainable diets.

The Food and Agriculture Organization of the United Nations has defined sustainable diets relative to four principal domains: nutrition, the environment, society and economics[1]. Guided by these domains, food patterns need to be nutrient adequate, sparing of natural resources and the environment, culturally acceptable and affordable[2]. Food choices should consider all of these aspects; for example, the environmental impact of animal food production may in some cases be higher than that of plant food production, but environmental costs that are not reflected in the price (that is, externalities) need to be assessed against attributes such as nutrient density and cultural and social value.

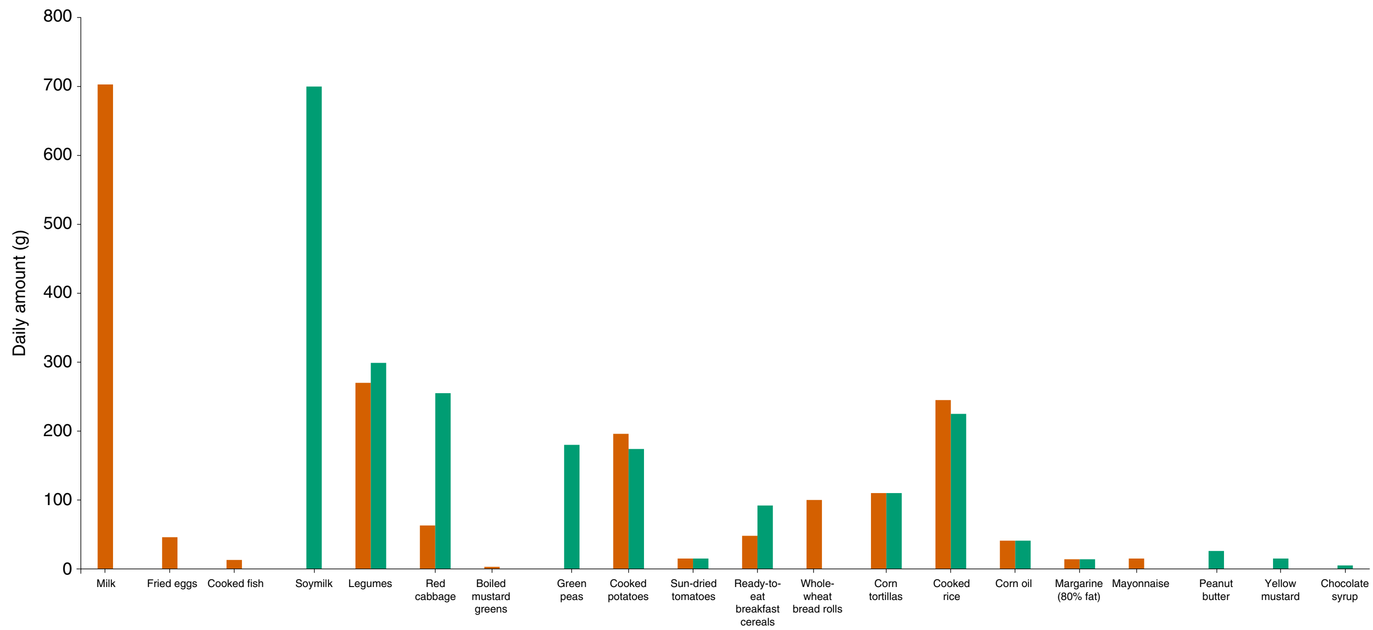
This study, which is focused on the economic domain, seeks to identify the best combination of food groups to minimize daily dietary cost while meeting energy and nutrient requirements in the United States. We used the most up-to-date, comprehensive and reliable food composition data (2016)[3] and national food prices (2009–2010)[4] available for the country. Although the detail and robustness of the dataset is a strength, that the data pertain only to the United States is also a limitation in terms of the generalizability of the findings to other regions. Yet, the price hierarchy of different food items has been shown to be the same in countries as diverse as France, the United Kingdom, the Netherlands, China, India and Australia[5,6].

In linear programming, diet formulation is determined by minimizing or maximizing (that is, optimizing) a given function while subjecting this function to several constraints[7,8]. In the present study, a linear programming algorithm was applied to verify the amounts of animal- and plant-based foods in least-cost diets and in the face of numerous constraints (such as energy and nutrient requirements, upper limits to nutrient intakes, food serving sizes, and available foods and their relative pricing). This allows a rational assessment to be made as to whether—from an economic perspective alone—animal-based foods need to be included in mixed diets for adult humans.

The linear programming analysis gives a unique solution of the combination of foods that meet all nutrient requirements at the lowest cost. The analysis is limited to the effect of food costs; the resultant food pattern has not been optimized from a health viewpoint, nor does it address other relevant aspects of food production, such as greenhouse gas emissions, natural resource use and environmental pollution. Although linear programming has been used in previous studies to evaluate diets for humans[9,10], the objective of those studies was to assess the impact of cost constraints on food choices, and to evaluate altered food intake and nutrient patterns. The linear programming modelling exercise is illustrative only, and does not purport to formulate a balanced recommended diet in a public health sense. Consistent with this, the combinations of different food items in the modelled diets are at times referred to as modelled food patterns. This study applies linear programming to derive economically optimal food patterns and identify food groups that need to be included in a nutritionally adequate modelled diet, to ensure that all nutrient requirements are met at the lowest cost. In other words, the hypothesis to be tested was whether animal-based foods, due to their high nutrient density, would be found in least-cost modelled diets for adult humans, given foods and food prices in the United States.

In total, five linear programming analyses were conducted. Linear programming analysis 1 investigated a dietary scenario whereby a modelled food pattern that met the total energy requirement of 2,600 kcal and the requirements for all key macronutrients and micronutrients (28 in total) was formulated at the lowest cost. The subsequent linear programming analyses 2 and 3 examined the effects of incremental changes in animal food prices. Linear programming analysis 4 considered a nutritional scenario whereby the requirements for the vitamins would be met by dietary supplements. Linear programming analysis 5 utilized a lower potassium recommended intake level of 3,400 mg d^{-1} (ref. [11]). Linear programming includes inherent sensitivity analysis features

[1]Riddet Institute, Massey University, Palmerston North, New Zealand. [2]Theta Solutions, Atascadero, CA, USA. [3]AL Rae Centre for Genetics and Breeding, School of Agriculture and Environment, Massey University, Hamilton, New Zealand. [4]Center for Public Health Nutrition, University of Washington, Washington, DC, USA. ✉e-mail: sylvia.lawrence.17@gmail.com



	Milk	Fried eggs	Cooked fish	Soymilk	Legumes	Red cabbage	Boiled mustard greens	Green peas	Cooked potatoes	Sun-dried tomatoes	Ready-to-eat breakfast cereals	Whole-wheat bread rolls	Corn tortillas	Cooked rice	Corn oil	Margarine (80% fat)	Mayonnaise	Peanut butter	Yellow mustard	Chocolate syrup
LP diet 1	703	46	13	–	270	63	3	–	196	15	48	100	110	245	41	14	15	–	–	–
LP diet 3	–	–	–	700	299	255	–	180	174	15	92	–	110	225	41	14	–	26	15	5

Fig. 1 | Daily intake of the 15 foods included in food pattern modelling for linear programming analysis 1 and the 14 foods for linear programming analysis 3. Food components of the least-cost modelled food patterns for linear programming analysis 1 (LP diet 1) and linear programming analysis 3 (whereby the prices of animal-based foods were increased to formulate a plant-only dietary pattern; LP diet 3).

(https://www.juliaopt.org/JuMP.jl/v0.20.0/). The sensitivity of the objective function to changes in constraints is represented by the shadow price. Specifically, the shadow price for a given constraint is the expected change of the objective function value (that is, the daily cost of the modelled diet) for an infinitesimal relaxation of the linear constraint. Sensitivity analyses were generated for all of the linear programming analyses. The sensitivity analyses, along with the shadow prices for all of the constraints that were met at minimum or maximum, for linear programming analysis 1 and linear programming analysis 3 are shown in Supplementary Table 1. Sensitivity analyses using shadow prices were undertaken for linear programming analysis 1 to understand the effect of varying constraint values.

Results

Linear programming analysis 1—least-cost nutrition under current conditions. The cost of the least-cost modelled food pattern (objective function) was US$1.98 per day and the 15 foods selected are presented in Fig. 1. Table 1 shows how the least-cost modelled food pattern provided the required nutrients. The modelled dietary pattern provided daily 2,600 kcal (10.9 MJ), 89.4 g of protein, 90.8 g of total fat and 367.2 g of carbohydrate. The linear programming diet was deemed to be nutritionally adequate as the energy amounts derived from protein (13.5%), fat (30.9%) and carbohydrate (55.6%) were within the acceptable macronutrient distribution ranges of 10–35% for protein, 20–35% for fat and 45–65% for carbohydrate, respectively[12,13]. The major nutrients found to be first limiting (100% of the recommended dietary allowance (RDA) or adequate intake) were the essential fatty acid α-linolenic acid, potassium, vitamin C, vitamin D, vitamin E, vitamin K and choline. Other nutrients providing close to their minimum required levels were vitamin A and calcium.

Applying either the shadow prices or re-running the linear programming model with modified constraints had only a small effect on the objective value, and led to similar outcomes and the same overall conclusions for linear programming analysis 1. As the constraint for the allowable amount of milk to be included in the modelled food pattern was found to be met at its maximum, the negative shadow price implied that an increase in the amount of milk to more than three servings per day would result in a decrease in the daily diet cost. Inversely, when the three servings (703 g) were reduced to two servings (468 g) and one serving (234 g), which is a further restriction of the constraint rather than a relaxation, the objective value increased, as implied by the shadow prices, from US$1.98 to US$2.04 (versus US$2.01 as predicted by the shadow price) and US$2.15 (versus US$2.05 predicted) per day, respectively. In all three milk serving scenarios tested, the maximum allowable amount of milk was included in the resulting least-cost dietary pattern, and as the maximum allowable amount of milk decreased, the amount of other animal-based food products (fried eggs and cooked fish) increased. When the allowable amount for all of the specified food groups was limited to no more than three servings per day per food group, the modelled food pattern included similar food types as those for linear programming analysis 1—particularly milk (703 g; three servings), fried eggs (42 g), cooked fish (11 g), bread rolls (150 g; three servings) and corn tortillas (165 g; three servings)—and had a slightly lower diet cost of US$1.89 per day. Overall, the conclusions based on linear programming analysis 1 were robust, and not greatly influenced by varying the respective constraint values.

Linear programming analysis 2—5–20% price increments for animal-based foods. Increasing the prevailing market prices of animal-sourced foods by 5, 10, 15 or 20% resulted in a gradual small increase in the daily cost of the modelled diet and a slight change in the foods selected (Table 2). When the prices of animal-based food products were increased by 15 or 20%, a small amount of seeds was included in the least-cost modelled food patterns and the amount of fish was reduced (Table 2). The nutrients potassium, vitamin C, vitamin D, vitamin E, vitamin K and choline remained first limiting (100% of RDA or adequate intake) for all four dietary scenarios. Compared with being first limiting (100% of adequate intake) for the baseline dietary pattern of linear programming analysis 1, and when there was a 5% increase in the original prices for the animal-based foods, the essential fatty acid α-linolenic acid was close to its minimum required level when the original costs of animal-based foods increased by 10, 15 and 20% (124% of adequate intake). As the baseline prices of animal-based foods increased from 5 to 20%, sodium, pantothenic acid and iron were closer to having their daily minimum required levels met (101–146% of the RDA or adequate intake).

Linear programming analysis 3—50% price increments for animal-based foods. The original prices of animal-based food products were increased by increments of 50% until no animal-based foods were included in the resulting least-cost modelled food pattern. A dietary pattern containing no animal-based food items became economically optimal only after an increase in the price of milk by eight times, eggs by 11.5 times, fish by 6.5 times, mayonnaise and animal-based salad dressings by five times, bread rolls and buns (which included milk and eggs) by 4.5 times, beef by 5.5 times, chicken by five times, sausages by three times, turkey by three times, cheese by three times, pork by 2.5 times, cold cuts and cured meats by twice, cooked egg noodles by twice, ice cream by twice, yogurt by 2.5 times and mashed potatoes (which included milk and/or butter) by twice their original costs, respectively. This resulted in a relatively expensive least-cost modelled food pattern with a daily cost of US$3.61, and containing 14 foods (Fig. 1). The least-cost plant-only dietary pattern provided 2,600 kcal and 77.2 g of protein, 77.3 g of total fat and 413.5 g of carbohydrate, contributing to 11.6, 26.2 and 62.2% of energy, respectively. The energy levels derived from protein, fat and carbohydrate were found to be within the recommended acceptable range of 10–35% for protein, 20–35% for fat and 45–65% for carbohydrate, respectively[13]. The nutrients that were first limiting (100% of the RDA or adequate intake) were α-linolenic acid, potassium, vitamin D, vitamin E and choline. Another nutrient that was comparable between both (animal- and plant-food-containing) least-cost modelled food patterns was pantothenic acid (vitamin B-5), which was found to be supplied at its minimum required level by the plant-based least-cost dietary pattern (100% of adequate intake). When animal-based foods became too costly to be part of the least-cost modelled dietary pattern, the requirements for vitamin C and vitamin K were met at 175% of the RDA and 234% of adequate intake, respectively, and the dietary supplies of selenium (126% of the RDA), calcium (133% of the RDA) and protein (152% of the RDA) were closer to their minimum requirements.

Linear programming analysis 4—no vitamins in the least-cost diet formulation. It was found that many food items readily available on the US market are enriched with vitamins and it was considered that this may have biased the outcomes by securing their inclusion in the least-cost modelled diets. Moreover, it is possible that vitamins could be supplied by dietary supplements. Therefore, a linear programming analysis was undertaken to minimize daily diet cost while meeting the recommended requirements for energy, the energy-providing nutrients and ten minerals. The nutrient requirements for young American adults were met with 12 food items, for a total energy value of 2,600 kcal and a daily diet cost of US$1.45. The latter cost does not include the additional cost of the necessary dietary vitamin supplements. The least-cost dietary pattern included a lower amount of milk (376 g), a higher amount of baked potatoes (330 g) and cooked rice (280 g), and no fish, eggs, cabbage, mustard greens, breakfast cereals or margarine. The amounts of boiled pinto beans, sun-dried tomatoes, whole-wheat bread rolls, corn tortillas, corn oil and mayonnaise remained the same as for linear programming analysis 1. Potassium, calcium and sodium were observed to be the first limiting nutrients (contents were 100% of the RDA or adequate intake). Zinc (112% of the RDA), iron (121% of the RDA) and the essential fatty acid α-linolenic acid (126% of adequate intake) were found to approach their minimum requirement levels.

Table 1 | Daily nutrients required by the average young adult man or woman, and provided by the least-cost modelled food pattern for linear programming analysis 1

Nutrient	Average adult requirement		Linear programming analysis 1	
	RDA or AI (amount per day)	Upper limit (amount per day)	Amount in modelled food pattern	Amount as % of RDA or AI
Carbohydrate (g)	130	–	367.2	282
Total dietary fibre (g)	31.5 (AI)	–	48.4	154
Linoleic acid (18:2 *n*-6 c,c) (g)	14.5 (AI)	–	30.3	209
α-linolenic acid (18:3 *n*-3 c,c,c) (g)	1.35 (AI)	–	1.35	100
Protein (g)	50.8	–	89.4	176
Calcium (mg)	1,000	2,500	1,343	134
Copper (mg)	0.9	10	1.8	202
Iron (mg)	13	45	20.7	159
Magnesium (mg)	355	–	524	148
Manganese (mg)	2.05 (AI)	11	6.2	302
Phosphorus (mg)	700	4,000	2,141	306
Potassium (mg)	4,700 (AI)	–	4,700	100
Selenium (µg)	55	400	140	255
Sodium (mg)	1,500 (AI)	2,300	2,300	153
Zinc (mg)	9.5	40	16.8	177
Vitamin A (RAE) (µg)	800	3,000	1,060	132
Thiamin (mg)	1.15	–	2.1	186
Riboflavin (mg)	1.2	–	3.1	256
Niacin (mg)	15	35	23.6	158
Pantothenic acid (mg)	5 (AI)	–	8.5	171
Vitamin B-6 (mg)	1.3	100	3.5	271
Vitamin B-12 (µg)	2.4	–	7.5	312
Folate (DFE) (µg)	400	1,000	1,000	250
Choline (mg)	487.5 (AI)	3,500	487.5 (AI)	100
Vitamin C (total ascorbic acid) (mg)	82.5	2,000	82.5	100
Vitamin D (IU)	600	100	600	100
Vitamin D (D2 + D3) (µg)	15	4,000	15	100
Vitamin E (α-tocopherol) (mg)	15	1,000	15	100
Vitamin K (phylloquinone) (µg)	105 (AI)	–	105	100

Average adult requirements are listed as both the RDA or adequate intake (AI) and the tolerable upper intake limit. Nutrient requirement values were sourced from the Institute of Medicine of the National Academies[12,14,35–39]. DFE, dietary folate equivalent; RAE, retinol activity equivalent.

Table 2 | Daily amounts (g) of food groups included in the least-cost modelled food patterns for linear programming analysis 1 compared with linear programming analysis 2

Food group	Linear programming analysis 1	Increase in price of animal-based foods			
		5%	10%	15%	20%
Milk	703	703	703	703	703
Fried eggs	46	46	47	47	47
Cooked fish	13	13	9	7	7
Boiled legumes	270	270	270	270	270
Vegetables	277	277	286	289	289
Breakfast cereals	43	43	52	56	56
Whole-wheat bread rolls	100	100	100	100	100
Corn tortillas	110	110	110	110	110
Cooked rice	245	245	237	218	218
Corn oil	41	41	41	41	41
Margarine	14	14	–	–	–
Vegetable oil spread	–	–	14	14	14
Mayonnaise	15	15	15	15	15
Seeds	–	–	–	2	2
Daily diet cost	US$1.98	US$2.02	US$2.06	US$2.10	US$2.14

Daily diet costs for each scenario are included in the final row.

Linear programming analysis 5—reduced daily recommended intake of potassium. It was noted that potassium was limiting in all of the least-cost modelled diets, and the recommended intake for potassium of 4,700 mg d^{-1} (ref. [14]) may be too high[11,15,16]. A recent report from the National Academies of Sciences, Engineering, and Medicine[11] suggested that the adequate intake recommendations for potassium be decreased to 3,400 mg d^{-1} for American adults aged over 18 years. When the requirement for potassium was set to 3,400 mg d^{-1} (while all other nutritional variables remained as for linear programming analysis 1), the dietary pattern was marginally cheaper, costing US$1.80 per day. Nonetheless, the results were not substantially different from those from linear programming analysis 1, whereby the modelled food pattern also included the same animal-based food products (milk, fried eggs and cooked fish), and potassium at the minimum daily intake requirement level of 3,400 mg remained first limiting (100% of adequate intake).

Discussion

The production of meat, eggs and dairy products for human consumption has been viewed by some as an inefficient use of natural resources[17]. Moreover, the recent EAT–*Lancet* report has recommended a global shift towards a diet that is more plant based, citing as part of its reasoning the environmental impact of the production of animal-based foods[18]. However, such conclusions are often based on crude comparisons between the efficiency of production of plant versus animal foods, and do not fully account for the high nutritional quality of animal-sourced foods and other factors[19]. Our aim was to assess the place of animal-sourced foods in least-cost food patterns for the United States, under current national prices.

When linear programming was applied to commonly available foods and food prices prevailing in the United States (2009–2010), animal-based foods (milk, eggs and fish meat) were selected for the formulation of dietary patterns that met the energy and nutrient requirements of healthy adults at the lowest cost. This shows that—at least for US foods and food prices in 2010—animal-based foods are required for nutritionally adequate least-cost food provision. Vitamins and minerals were the main limiting factors in the least-cost dietary patterns, and nutrient availability needs to be taken into account in future work.

The linear programming approach remains to be extended to the elderly, pregnant or lactating women, and growing children, who are more at risk with inadequate nutrient intakes and increases in food prices[5,6,9,15,16,20–22]. Yet, the present analysis illustrates the usefulness of linear programming to determine the influence of different food groups on the lowest cost of a dietary pattern. As this analysis is restricted to the US economy and to a single time period (2009–2010), the work should be repeated for other economies, especially where the relative prices of food groups may be different[6,23]. In particular, the results based on US data may not be general in that there may be food price distortions caused by animal food subsidies imposed by the US Government. However, the effects of subsidies are complex and there is evidence that subsidies have little effect on retail prices of animal-based foods[24].

The set of linear programming analyses included increasing the prices of animal-based foods in relation to plant-based foods, to evaluate to what extent the price of animal foods could rise before such foods become too costly to be included in the least-cost food pattern. This gives an indication of the margin for potential inclusion of externalities, or costs arising from the removal of subsidies, into the costs of these animal foods. Increasing market prices of animal-based foods by up to 20% had little effect on the food composition of the least-cost modelled diets. To formulate a nutritionally adequate diet at the lowest cost that no longer included any animal-based foods, the price of all animal-derived foods had to be increased by 200–1,150% of their baseline costs. The resulting plant-only diet was relatively expensive (US$3.61 compared with US$1.98 for linear programming analysis 1). A change in diet cost may mean greater or lesser affordability for particular groups of the US population. Given that the average per-capita income in 2010 was US$26,558 (ref. [25]), 5% of the annual income ((US$3.61 (daily diet cost) × 365 d) / US$26,558 (income)) would be spent on plant-only food products, relative to 2.7% of the annual income ((US$1.98 (daily diet cost) × 365 d) / US$26,558 (income)) spent on a diet that included animal- and plant-based food products. In particular, a one-person American household at the poverty threshold of US$11,139 in 2010[25] would spend 11.8% of the annual income on foods based only on plants, compared with 6.5% of the annual income on foods originating from plants and animals. When relative affordability is considered, the food pattern that included animal- and plant-sourced foods would be more affordable than the modelled dietary pattern that only consisted of plant-based foods.

A potential limitation of the present study is that the wider costs (both social and environmental) associated with animal production and increased consumption of animal-based foods are not considered. Plant food production also has associated externality costs, but these may be greater for some systems of animal production. The differences in such costs between agricultural production systems and the effects of such cost differences on the least-cost modelled diets are beyond the scope of the present study, but should be taken into account in future work. In any case, it is apparent from the present study that, when price elasticities were considered, the 2009–2010 US food retail prices of animal-based foods had to be substantially increased before the animal-derived foods disappeared from the modelled dietary pattern. This would offset to some extent any potential relatively higher externality costs associated with animal food production, compared with plant food production.

The least-cost modelled food patterns selected by the linear programming exercise take no account of the holistic properties of foods in healthy diets[26] and food attributes other than the main nutrients, and are not meant to be realistic diets to be recommended in practice. Rather, they highlight the food groups needed to be included in a dietary pattern at the lowest cost. Based on the criterion of minimum food cost alone, the animal-based food group appears to have a role in optimum cost-minimized dietary patterns.

Methods

Linear programming. The linear programming model was developed and solved using the Julia programming language[27], together with the JuMP mathematical optimization library[28], due to their flexibility and performance. The linear programming model takes into account the dietary supply of all nutrients simultaneously to meet stated nutrient requirement levels while minimizing the cost of the modelled food pattern (that is, the objective function)[7,8,29–33]. The objective function to be minimized in the linear programming model (a linear function) was formulated with the dependent variables as the quantities of each food item to be consumed and the coefficients set as the corresponding cost of each food item. Linear constraints were defined using energy and nutritional requirements and, where applicable, also included known upper nutrient intake limits and maximum daily food serving sizes. The constraints applied were as follows:

(1) Daily amounts of each food item needed to be either null or positive;
(2) The maximum quantity of a food item to be consumed per day was restricted to be no more than three times the recommended reference amount customarily consumed (RACC) at one eating occasion[34] (≤3× RACC);
(3) The energy content of the modelled food pattern was set to meet the daily estimated energy requirement of 2,600 kcal for an active young adult human[12], such that the constraint on energy was set to be 2,600 kcal d^{-1};
(4) The daily amount of each nutrient provided by the modelled food pattern was equal to or above the minimum, defined as either RDAs or adequate intakes[12,14,35–39];
(5) The daily intake of each nutrient was equal to or below the tolerable upper intake level[14,35–39].

The linear programming model described above is summarized mathematically as follows:

Minimize:

$$f(x) = \sum_{i=1}^{N_f} c_i x_i$$

subject to:

$$x_i \geq 0 \ (i = 1, 2, \ldots, N_f) \quad (1)$$

$$x_i \leq 3r_i \ (i = 1, 2, \ldots, N_f) \quad (2)$$

$$\sum_{i=1}^{N_f} e_i x_i = E \quad (3)$$

$$\sum_{i=1}^{N_f} n_{ij} x_i \geq m_j \ (j = 1, 2, \ldots, N_n) \quad (4)$$

$$\sum_{i=1}^{N_f} n_{ij} x_i \leq u_j \ (j = 1, 2, \ldots, N_n) \quad (5)$$

where N_f is the number of food items included in the linear programming analysis, N_n is the number of nutrient quantities included in the linear programming analysis, x_i is the number of units of food item i to consume, c_i is the cost per unit of food item i, r_i is the recommended daily intake serving size for food item i, e_i is the energy per unit of food item i, E is the daily energy target to meet, n_{ij} is the amount of nutrient j per unit of food item i, m_j is the daily minimum required amount of nutrient j, and u_j is the daily maximum intake level of nutrient j.

The above constraints were first applied to all individual food items included in the linear programming model. Initial analyses using the above constraints led to the inclusion of multiple similar types of food items, such as several different types of bread roll, each being selected up to their allowed three servings (3× RACC) quantity in the resulting least-cost modelled food pattern. Such diets are impractical. To address this, some additional constraints as analogous to constraint (2) described above were applied to specific food groups, on the basis of the MyPlate guidelines (https://www.choosemyplate.gov/). Milk, milk substitutes, legumes, tomatoes, potatoes, vegetable oils and sugars were limited to no more than three servings per day per food group. Yeast-based breads (including bread rolls and buns), tortillas and rice were restricted to no more than two servings per day per food group. Margarine and vegetable spreads, peanut butter, mayonnaise and salad dressings were limited to no more than the recommended RACC once (1× RACC) per day per food group.

Foods. The United States Department of Agriculture What We Eat in America database was supplemented with data from the National Nutrient Database for Standard Reference 28, to provide a comprehensive list of foods and their nutrient contents[3]. A subset of foods was selected for the linear programming model that excluded mixed dishes and most composite foods and included most key foods, based on the 2011–2012 food consumption survey data from the United States Department of Agriculture National Health and Nutrition Examination Survey and that provide 75% of the nutrient intake of the total US population for selected nutrients of public health importance, as identified by Haytowitz et al.[40–42]. The 962 selected food items are commonly available and consumed in the United States. The nutrient composition of the 962 selected foods, given per 100 g of edible portions[3], was based on either raw or cooked foods and different methods of food storage (such as canned, frozen or dried). For each food item, daily intake serving sizes were based on the RACCs at one eating occasion[34]. The prices of the foods were those prevalent in the US market in 2009–2010, as given by the Center for Nutrition Policy and Promotion, and were the most up-to-date, reliable and comprehensive food prices available[4].

Energy and nutrient requirements. The daily requirements for energy and nutrients for a healthy American adult, aged between 19 and 50 years, were based on the recommendations reported by the Institute of Medicine of the National Academies (Table 1). The daily estimated energy requirement for a healthy 63.5-kg active young adult was given as 2,600 kcal, based on the average of the energy requirements for a low active and moderately active man aged 19 years, weighing 70 kg and being 1.77 m in height, and a low active and moderately active woman aged 19 years, weighing 57 kg and being 1.63 m in height[12,43]. Recommendations for daily nutrient intakes were given as either the RDA, which met the nutrient requirement of almost all (97.5%) healthy individuals, or adequate intakes, based on observed determined estimates of nutrient intake. The RDA for total available carbohydrates (sugars and starches) was 130 g d^{-1} based on the average minimum amount of glucose utilized by the brain[12]. The RDA value for protein was based on the reference equation of 0.8 g of protein per kg body weight per day and the reference body weight of 70 kg for a man and 57 kg for a woman[12,43]. Adequate intake values were used for total dietary fibre and the two polyunsaturated fatty acids required in the diet (namely, linoleic acid and α-linolenic acid)[12]. Daily nutrient requirement data for key minerals and vitamins[14,35–39] were also included (Table 1). The RDA for magnesium was given as the average intake level for an adult aged between 19 and 30 years, based on the recommended 400 mg for a man and 310 mg for a woman, respectively[35]. The RDA for vitamin A was presented as the retinol activity equivalent to account for the different bioactivities of retinol and carotenoids[38], while folate was reported as the dietary folate equivalent[36]. The daily recommended nutrient intakes for manganese, potassium, sodium, pantothenic acid, choline and vitamin K (phylloquinone) were given as adequate intakes[14,36,38]. Available maximum micronutrient intake levels[14,35–39] were also included as tolerable upper levels (Table 1).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data used and generated during the current study are available from an online resource (https://gitlab.com/thetasolutionsllc/naturefood-19100372).

Code availability

The computer code required to reproduce the findings of this study is available from an online resource (https://gitlab.com/thetasolutionsllc/naturefood-19100372).

Received: 4 November 2019; Accepted: 18 May 2020;
Published online: 17 June 2020

References

1. *Biodiversity and Sustainable Diets: United Against Hunger* (FAO, 2010).
2. Drewnowski, A. in *Sustainable Nutrition in a Changing World* 25–34 (Springer, 2017).
3. *National Nutrient Database for Standard Reference* Release 28 (USDA, 2016); http://www.ars.usda.gov/nea/bhnrc/mafcl.
4. Rehm, C. D., Monsivais, F. & Drewnowski, A. Relation between diet cost and Healthy Eating Index 2010 scores among adults in the United States 2007–2010. *Prev. Med.* **73**, 70–75 (2015).
5. Darmon, N. & Drewnowski, A. Contribution of food prices and diet cost to socioeconomic disparities in diet quality and health: a systematic review and analysis. *Nutr. Rev.* **73**, 643–660 (2015).
6. Headey, D. D. & Alderman, H. H. The relative caloric prices of healthy and unhealthy foods differ systematically across income levels and continents. *J. Nutr.* **149**, 2020–2033 (2019).
7. Dantzig, G. B. *Linear Programming and Extensions* (Princeton Univ. Press, 1963).
8. Van Dooren, C. A review of the use of linear programming to optimize diets, nutritiously, economically and environmentally. *Front. Nutr.* **5**, 48 (2018).
9. Darmon, N., Ferguson, E. L. & Briend, A. Impact of a cost constraint on nutritionally adequate food choices for French women: an analysis by linear programming. *J. Nutr. Educ. Behav.* **38**, 82–90 (2006).
10. Cleveland, L. E., Escobar, A. J., Lutz, S. M. & Welsh, S. O. Methods for identifying differences between existing food intake patterns and patterns that meet nutrition recommendations. *J. Am. Diet. Assoc.* **93**, 556–560, 563 (1993).
11. National Academies of Sciences, Engineering, and Medicine *Dietary Reference Intakes for Sodium and Potassium* (National Academies Press, 2019).
12. Institute of Medicine *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (National Academies Press, 2002).
13. Institute of Medicine *Dietary Reference Intakes: The Essential Guide to Nutrient Requirements* (National Academies Press, 2006).
14. Institute of Medicine *Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate* (National Academies Press, 2005).
15. Nicklas, T. A., O'Neil, C. E. & Fulgoni, V. L. III The role of dairy in meeting the recommendations for shortfall nutrients in the American diet. *J. Am. Coll. Nutr.* **28**, 73–81 (2009).
16. *2015–2020 Dietary Guidelines for Americans* (USHHS, USDA, 2015).
17. IPCC *Special Report on Global Warming of 1.5 °C* (eds Masson-Delmotte, V. et al.) (World Meteorological Organization, 2018).
18. Willet, W. et al. Food in the Anthropocene: the EAT–*Lancet* Commission on healthy diets from sustainable food systems. *Lancet* **393**, 447–492 (2019).
19. Adesogan, A. T., Havelaar, A. H., McKune, S. L., Eilitta, M. & Dahl, G. E. Animal source foods: sustainability problem or malnutrition and sustainability solution? Perspective matters. *Glob. Food Secur*. https://doi.org/10.1016/j.gfs.2019.100325 (2019).
20. Darmon, N., Ferguson, E. L. & Briend, A. Linear and nonlinear programming to optimize the nutrient density of a population's diet: an example based on diets of preschool children in rural Malawi. *Am. J. Clin. Nutr.* **75**, 245–253 (2002).
21. Darmon, N., Lacroix, A., Muller, L. & Ruffieux, B. Food price policies improve diet quality while increasing socioeconomic inequalities in nutrition. *Int. J. Behav. Nutr. Phys. Act.* **11**, 66 (2014).
22. Briend, A. & Darmon, N. Determining limiting nutrients by linear programming: a new approach to predict insufficient intakes from complementary foods. *Pediatrics* **106**, 1288–1289 (2000).
23. Hirvonen, K., Bai, Y., Headey, D. & Masters, W. A. Affordability of the EAT–*Lancet* reference diet: a global analysis. *Lancet Glob. Health* **8**, 59–66 (2020).
24. Alston, J. M., Sumner, D. A. & Vosti, S. A. Farm subsidies and obesity in the United States: national evidence and international comparisons. *Food Policy* **33**, 470–479 (2008).
25. *Poverty Thresholds* (US Census Bureau, 2010); https://www.census.gov/topics/income-poverty/data/tables.2010.html.
26. Moughan, P. J. Holistic properties of foods: a changing paradigm in human nutrition. *J. Sci. Food Agric*. https://doi.org/10.1002/jsfa.8997 (2018).
27. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: a fresh approach to numerical computing. *SIAM Rev.* **59**, 65–98 (2017).
28. Dunning, I., Huchette, J. & Lubin, M. JuMP: a modeling language for mathematical optimization. *SIAM Rev.* **59**, 295–320 (2017).
29. Balintfy, J. L. Menu planning by computer. *Commun. ACM* **7**, 255–259 (1964).
30. Stigler, G. J. The cost of subsistence. *J. Farm Econ.* **27**, 303–314 (1945).
31. Buttris, J. L. et al. Diet modelling: how it can inform the development of dietary recommendations and public health policy. *Nutr. Bull.* **39**, 115–125 (2014).
32. Gazan, R. et al. Mathematical optimization to explore tomorrow's sustainable diets: a narrative review. *Adv. Nutr.* **9**, 602–616 (2018).
33. Gephart, J. A. et al. The environmental cost of subsistence: optimizing diets to minimize footprints. *Sci. Total Environ.* **553**, 120–127 (2016).
34. *Code of Federal Regulations*: *Title 21—Food and Drugs* Vol. 2, Section 101.12 (US FDA, 2017); https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=101.12.
35. Institute of Medicine *Dietary Reference Intakes for Calcium, Phosphorus, Magnesium, Vitamin D and Fluoride* (National Academies Press, 1997).
36. Institute of Medicine *Dietary Reference Intakes for Thiamine, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B12, Pantothenic Acid, Biotin, and Choline* (National Academies Press, 1998).
37. Institute of Medicine *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids* (National Academies Press, 2000).
38. Institute of Medicine *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc* (National Academies Press, 2001).
39. Institute of Medicine *Dietary Reference Intakes for Calcium and Vitamin D* (National Academies Press, 2011).
40. Haytowitz, D. B. et al. Key foods: setting priorities for nutrient analyses. *J. Food Composit. Anal.* **9**, 331–364 (1996).
41. Haytowitz, D. B., Pehrsson, P. R. & Holden, J. M. The identification of key foods for food composition research. *J. Food Composit. Anal.* **15**, 183–194 (2002).
42. Haytowitz, D. B. Updating USDA's key foods list for what we eat in America, NHANES 2011–12. *Procedia Food Sci.* **4**, 71–78 (2015).
43. Kuczmarski, R. J. et al. 2000 CDC growth charts for the United States: methods and development. *Vital Health Stat.* **246**, 1–190 (2002).

Acknowledgements

The work reported was supported in part through funds from the National Dairy Council and Global Dairy Platform.

Author contributions

S.M.S.C. and P.J.M. were responsible for the design and analysis of the study. D.P.G. was responsible for the linear programming modelling and analysis. A.D. was responsible for the provision of databases and consulting on the linear programming models. S.M.S.C. prepared the first draft of the manuscript. P.J.M revised the first draft of the manuscript. All authors participated in interpretation of the results and have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s43016-020-0096-8.

Correspondence and requests for materials should be addressed to S.M.S.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): NATFOOD-19100372C

Last updated by author(s): May 13, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The Linear Programming model was developed using the Julia language (version 1.3.0) together with the JuMP mathematical optimization library (https://www.juliaopt.org/JuMP.jl/v0.20.0/).
Data analysis	The data were analysed using the Julia language (version 1.3.0) together with the JuMP mathematical optimization library (https://www.juliaopt.org/JuMP.jl/v0.20.0/). The software code used is available as an online resource (https://gitlab.com/thetasolutionsllc/naturefood-19100372).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The foods dataset is publicly available from the United States Department of Agriculture (http://www.ars.usda.gov/nea/bhnrc/mafcl). The data used in the study is available as an online resource (https://gitlab.com/thetasolutionsllc/naturefood-19100372).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	*Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data exclusions	*Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Replication	*Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.*
Randomization	*Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.*
Blinding	*Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National*

Research sample	Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.
Sampling strategy	Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.
Data collection	Describe the data collection procedure, including who recorded the data and how.
Timing and spatial scale	Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Reproducibility	Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.
Randomization	Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.
Blinding	Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).
Location	State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).
Access and import/export	Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).
Disturbance	Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☐	Antibodies
☐	☐	Eukaryotic cell lines
☐	☐	Palaeontology
☐	☐	Animals and other organisms
☐	☐	Human research participants
☐	☐	Clinical data

Methods

n/a	Involved in the study	
☐	☐	ChIP-seq
☐	☐	Flow cytometry
☐	☐	MRI-based neuroimaging

Antibodies

Antibodies used	Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.
Validation	Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	*State the source of each cell line used.*
Authentication	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*
Mycoplasma contamination	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*

Outcomes | *Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication. | *For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*

Files in database submission | *Provide a list of all files available in the database submission.*

Genome browser session
(e.g. UCSC) | *Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates | *Describe the experimental replicates, specifying number, type and replicate agreement.*

Sequencing depth | *Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*

Antibodies | *Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*

Peak calling parameters | *Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*

Data quality | *Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*

Software | *Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation | *Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*

Instrument | *Identify the instrument used for data collection, specifying make and model number.*

Software | *Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*

Cell population abundance | *Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*

Gating strategy | *Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*
Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*
Effect(s) tested	*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*
Correction	*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity	*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*
Graph analysis	*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.