Nettle, Andrews & Bateson: Food insecurity as a driver of obesity in humans: The insurance hypothesis

Online Appendix A: Theoretical model

This appendix presents in detail the theoretical model outlined in section 3 of the main paper. Section A1 gives details of the modelling methods and assumptions. Section A2 demonstrates that the model converges to give an optimal eating policy. Section A3 examines the consequences of varying each of the principal parameters of the model. Section A4 gives details of implementation, code availability and instructions for running the model.

A1. Modelling framework

As described in the main paper (section 3), we consider individuals who carry some level of stored fat reserves r ($0 \le r \le 50$) with an energy requirement $e(r)$ in each time period as follows:

$$e(r) = a + br$$

Here, a is a fixed lean-mass metabolic requirement, and b is a parameter controlling how strongly the energy requirement increases with increasing body weight. All results unless otherwise stated use the simplest possible scenario of $a = 1$ and $b = 0$. That is, the energy requirement is assumed to be a fixed 1 unit of energy in each time period. Sections 3.1 and 3.2 of this document explore the consequences of varying a and b.

The individual is deemed to have starved to death if energy reserves fall to zero. Where reserves are greater than zero, survival each period is given by:

$$s(r) = \frac{1}{1 + e^{-x(r-w)}} - \mathrm{yr}$$

As we can see, this function has two additive components. First, there is a logistic function increasing in reserves. This models the probability of avoiding death by starvation, and the logistic function is used to capture the intuition that death by starvation in a time period, which is certain at $r = 0$, rapidly becomes very unlikely as long as the individual has a threshold level of reserves $r = w$. We use $w = 1$ throughout unless specifically stated otherwise. The logistic function is controlled by a steepness parameter x. Where x is large (e.g. $x = 10$), the probability of avoiding starvation in a time period approximately follows a step function: 0 at reserves of 0, 0.5 at reserves of 1, 1 at reserves of 2. All results use $x = 10$ unless otherwise stated. The second component of the survival function is a linear decrease in survival with increasing resources. This models the increase in morbidity and decrease in mobility as the individual becomes heavier. It is controlled by another steepness parameter y, which represents the survival cost of each additional unit of weight. We use $y = 0.01$ throughout except where otherwise stated. We explore the consequences of varying w, x and y for optimal eating and weight regulation in sections 3.3 and 3.4 of this document.

Each time period, the individual finds food with probability p. If food is available, the individual consumes n units ($0 \le n \le N$; $N = 10$ unless otherwise stated); its reserves will thus change by $n - e(r)$ units. If it does not find food, its reserves will change by $-e(r)$. We consider the consequences of varying N, the maximum amount of energy that can be consumed in a time period, in in section 3.5 of this document.

To compute the optimal policy, we consider a run of *T* time periods (the value of *T = 100* is used for all results presented; see section 2). We need to determine the survival function *F(r, t)*, which gives the maximum probability of survival to time *T* for an individual with fat reserves *r* at time *t.* First, we note that if the individual is still alive at the final time period, then they have survived. Hence:

$$F(r,T) = \begin{cases} 1 & r > 0 \\ 0 & otherwise \end{cases}$$

For each earlier time step, we can write down the dynamic programming equation:

$$F(r,t) = \begin{cases} \max_n (s(r)[p(F(r+n-e(r),t+1) + (1-p)(F(r-e(r),t+1)]) & r > 0 \\ 0 & otherwise \end{cases}$$

Note that the metabolic requirement in each time period is rounded to the nearest full unit, and the amount eaten also has to be an integer number of units. Given that we now have a survival function for the final time period, and for each time period given the next step, we can work backwards to time 1 by repeated application of the dynamic programming equation. This allows us to determine *n**, which is the amount to eat that produces the maximum survival probability for each possible value of *r* when *t = 1*. This is what is referred to throughout as the optimal policy; the amount to eat that maximizes survival into the distant future for every possible level of current fat reserves.

A2. Model convergence

Figure A1 shows the optimal amount to eat, as determined by application of the dynamic programming equation, for each possible level of current reserves at each time period, with *T* = 100. In the final few time periods, individuals maximise survival by stopping eating in order to reach the terminal time point with no excess. However, as long as the terminal time point is moderately distant, an optimal eating policy emerges that depends on current reserves only and not on time. This convergence occurs for all values of *p* within fewer than 20 time periods. Thus, using *n** at *t = 1* with *T* = 100 as the eating policy that maximizes survival into the distant future appears justified.

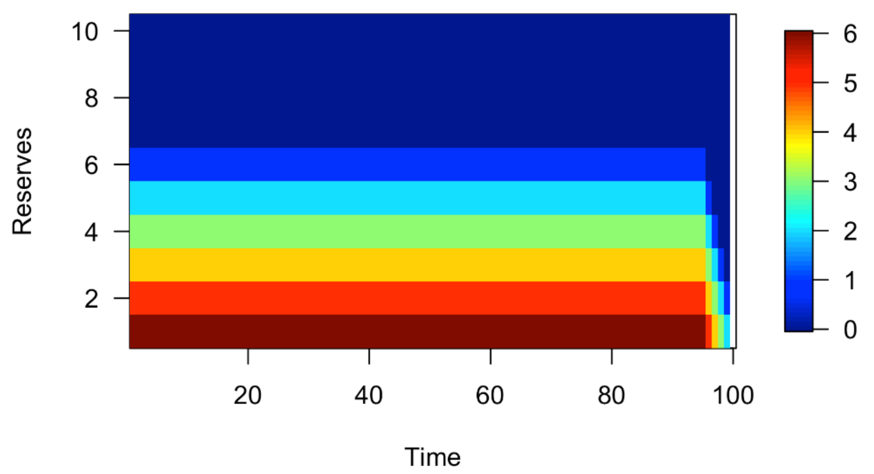

Figure A1. Amount to eat that maximises survival to time period 100 (colours), in relation to current time period (horizontal axis) and current reserves (vertical axis), for *p* = 0.5. All other parameters have their default values.

A3. Varying model parameters

This section investigates the effects of changing the values of each of the principal parameters of the model on the relationship between food security and fat reserves.

A3.1 Varying the fixed metabolic requirement

In section 3 of the main paper, the fixed component a of the metabolic requirement per time period is set to 1. Here we investigate the consequences of varying it. Repeating figure 1B of the main paper but with a set successively to 1, 2 and 3 produces the results shown in figure A2. Increasing a leads, unsurprisingly, to optimal policies that eat more per time period, but decreasing food security p still produces the same qualitative effect: it leads to eating more at a given level of reserves. In fact, the impact of p on amount eaten and hence reserves maintained becomes more marked when a is larger.

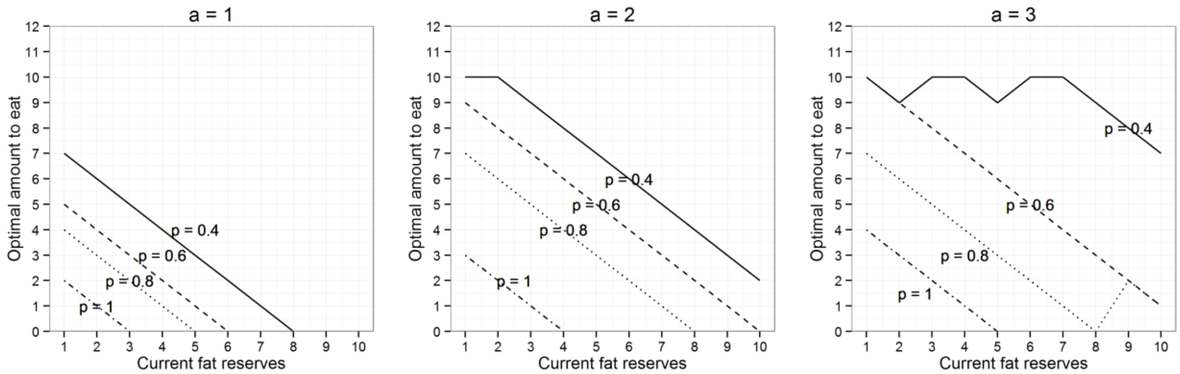

Figure A2. Figure 1B of the main paper repeated with different values of a, the fixed component of the metabolic requirement per time period. All other parameters have their default values.

A3.2 Varying the mass-dependent metabolic requirement

In section 3 of the main paper, b (the mass-dependent component of metabolic requirement) is set to zero in order to illustrate the simplest possible scenario. Here we investigate the consequences of giving this parameter a non-zero value, though we only consider cases where $b < a$, since lean mass has a larger influence on metabolic rate than fat mass (Garby et al., 1988; Johnstone, Murison, Duncan, Rance, & Speakman, 2005). Holding a at 1, we consider three values of b, 0, 0.2, and 0.5. For each level of b, we plot the optimal policy for four values of p (0.4, 0.6, 0.8 and 1), as in figure 1B of the main paper. Figure A3 shows the results.

The central result—that the optimal point to start eating and the optimal amount to eat increase as p decreases—is not altered by increasing the value of b up to 0.5. In fact, the increase in food consumption as p decreases becomes more marked with increasing b. This is because the eating policies for lower p have to eat to fund the additional metabolic requirements of the buffer they will need to build up, whereas the policies for higher p will build up little buffer and hence experience little increased metabolic cost.

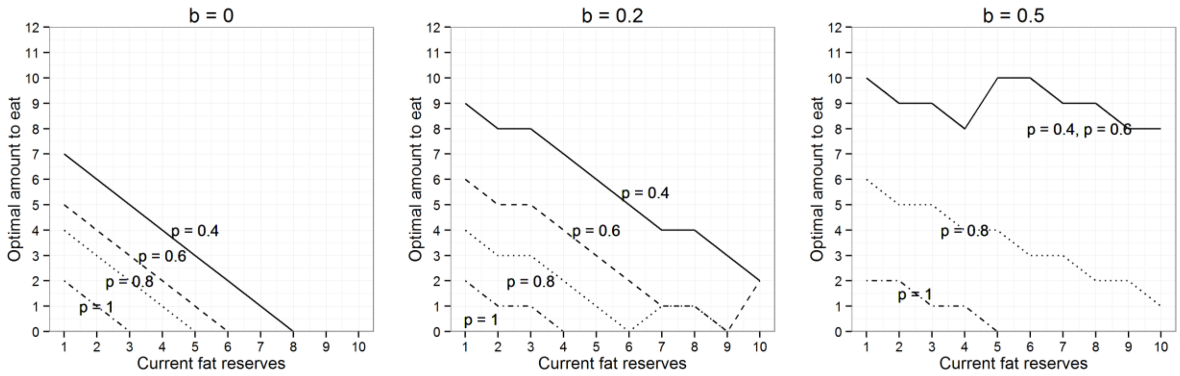

Figure A3. Figure 1B of the main paper repeated with different values of *b*, the parameter controlling how strongly metabolic requirements per time period increase with increasing weight. All other parameters have their default values.

A3.3 Varying the parameters of the logistic component of the survival function

In results presented thus far, the survival function in each time period steps up abruptly around *r = 1*, so that starvation, which is certain at *r* = 0, is essentially impossible at *r* = 2. We can vary this assumption in two ways. First, we can increase the value of the location parameter *w* in the logistic component of the survival function. This has the effect of moving the fitness cliff-edge to the right without changing its steepness (figure A4, panel A). Figure A4 panel B shows the consequences of increasing *w* for the relationship between food security and steady-state target reserves (i.e. the level at which the individual will stabilise under the optimal policy if it finds enough food to do so). As the figure shows, and unsurprisingly, increasing *w* increases the steady-state target level of fat reserves at any value of *p*. It does not however, change the gradient of the relationship between *p* and steady-state fat reserves. Thus, increasing *w* should be expected to produce individuals who are fatter at all levels of *p*, but no more or less responsive to their level of food insecurity.

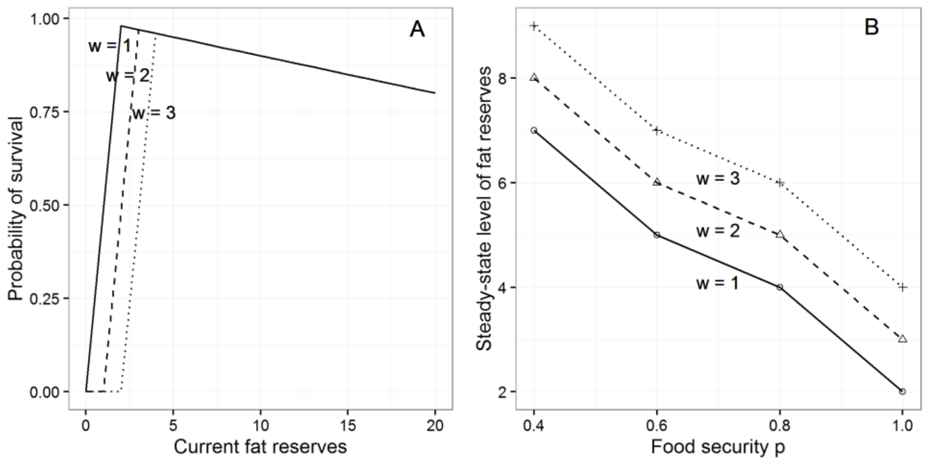

Figure A4. A. The probability of survival against current fat reserves for three values of the logistic location parameter *w.* B. The steady-state level of fat reserves that optimally-behaving individuals reach if they find food in every period against *p*, the level of food security, for three values of *w.* All parameters other than *x* have their default values. This figure is reproduced as part of figure 3 of the main paper.

Second, we can change the steepness parameter *x*, and hence make the increase in survival more gradual with increasing reserves (figure A5, panel A). Figure A5, panel B, shows the effect of food security on steady-state target reserves (i.e. the level at which the individual will stabilise under the optimal policy if it finds enough food to do so), for different values of the steepness parameter *x*. As the figure shows, decreasing *x* leads to individuals optimally carrying more fat. This is unsurprising since the effect of decreasing *x* is to move the point of maximal survival to the right (see figure A5, panel A). However, a secondary effect of reducing *x* is that individuals become somewhat less responsive to food insecurity; the difference in optimal reserves between $p = 0.4$ and $p = 1$ is 5 units when $x = 10$, but only 2 units when $x = 0.5$. Thus, a survival function that is less step-like at the lower end leads to individuals carrying more fat, but also being somewhat less responsive to their food security.

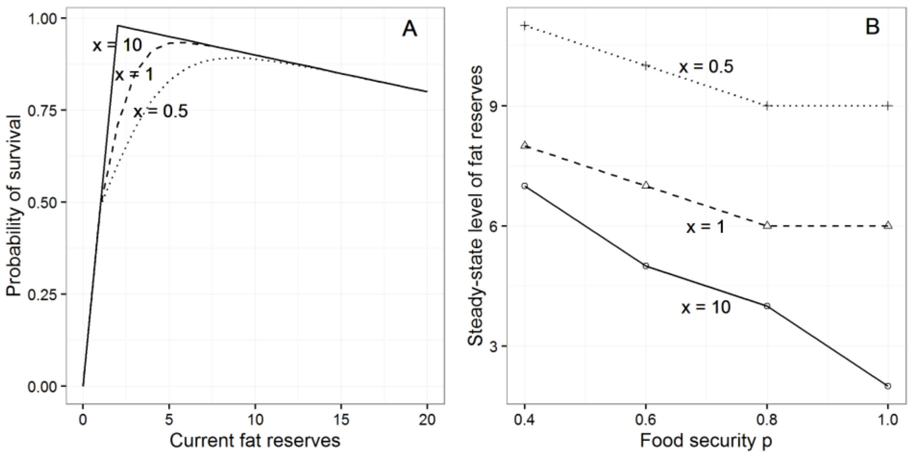

Figure A5. A. The probability of survival against current fat reserves for three values of the logistic parameter *x.* B. The steady-state level of fat reserves that optimally-behaving individuals reach if they find food in every period against *p*, the level of food security, for three values of *x.* All parameters other than *x* have their default values. This figure is reproduced as part of figure 3 of the main paper.

A3.4 Varying the steepness of the linear component of the survival function

The slope of the right-hand arm of the survival function is controlled by the parameter *y* in the model. In this section we investigate the consequences of making *y* larger, and hence the survival cost of every extra unit of weight greater. We do this by considering the consequences of using *y* values of 0.05 and 0.2 as well as the 0.01 used until this point (whilst keeping all other parameters, including *x*, at their usual values). This produces the three survival functions shown in figure A6, panel A.

We now compute the steady-state target reserves (i.e. the level at which the individual will stabilise under the optimal policy if it finds enough food to do so), for different values of *y* (figure A6, panel B). As the figure shows, a higher *y* leads to individuals carrying less fat at all levels of *p* except $p = 1$. Moreover, increasing *y* also makes individuals less responsive to food insecurity. The difference in steady-state reserves between $p = 1$ and $p = 0.4$ is 5 units for $y = 0.01$. It is only 1 units for $y = 0.2$. This is a logical result: making the carrying of a buffer more costly reduces the size of buffer individuals should carry.

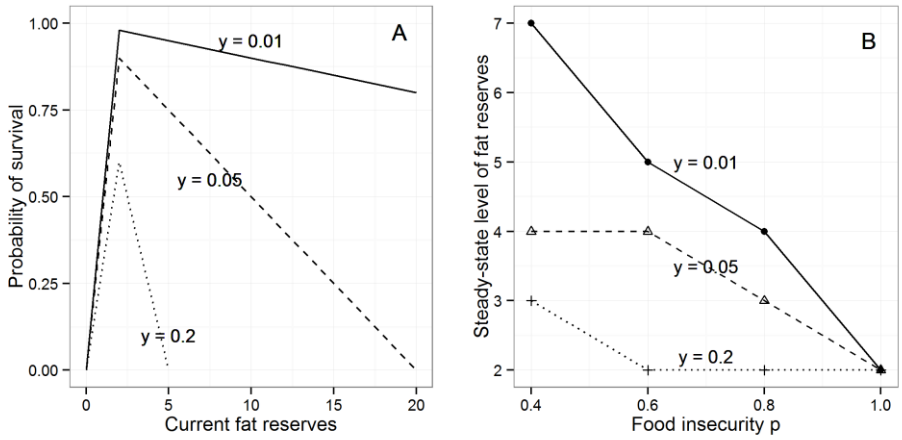

Figure A6. A. The probability of survival against current fat reserves for three values of the parameter *y.* B. The steady-state level of fat reserves that optimally-behaving individuals reach if they find food in every period against *p*, the level of food security, for three values of *y.* All parameters other than *y* have their default values. This figure is reproduced as part of figure 3 of the main paper.

A3.5 Varying the energy density of food

As a final variation of the parameters of the model, we investigate limiting *N* (the maximum energy available from food per time period assuming food can be found). The previously used value of *N = 10* effectively meant that individuals could build up a large reserve in a single time period in which food is available. We now explore setting *N* at the lower values of 3 and 2; this captures a situation where even when food is available, its energy-density is not sufficient to be able to consume many more calories than needed for metabolism in a given period. First we show the consequences of following the optimal eating policy and finding food in every period, for high food insecurity (p = 0.4), for the different values of *N* (figure A7). As the figure shows, the lower *N* is, the higher the individual's steady-state target weight (this is only true for high levels of food insecurity, $p < 0.6$). However, the lower *N* is, the more slowly the individual is able to put on weight to attain that target.

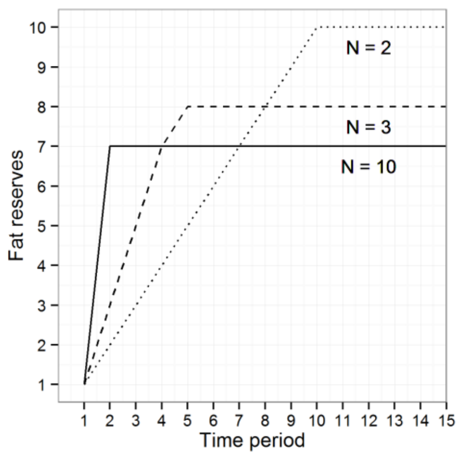

Figure A7. Level of fat reserves over time for individuals experiencing p = 0.4 who start with reserves of 1, follow the optimal eating policy, and find food every period for three values of the maximum energy available from food per period (*N*). All other parameters have their default values.

Next, we consider the consequences of decreasing *N* for the relationship between food security and average reserves/weight. We do this by repeating the simulations underlying figure 1D of the main paper, but with *N* = 2 as well as *N* = 10. The results are shown in figure A8. As the figure shows, when *N* = 2, the variance in individual weight within a value of *p* becomes substantial, particularly when *p* is low. This is because although individuals have a high steady-state target weight, in practice they are often operating well below it, because stochastic runs of foodless periods reduce their reserves, and it takes them a long time to build their reserves back up again because of the limited energy available in the periods when they do find food. The effect of this increased variability *within* groups of individuals at the same level of *p* is to attenuate the statistical relationship between *p* and mean weight. In the data underlying the left (*N* = 10) panel of figure A8 (and excluding individuals for whom *p* = 1), the value of *p* explains 77% of the variance in mean body weight. In the data underlying the right (*N* = 2) panel, the value of *p* explains only 20% of the variance in mean body weight, even though the relationship between steady-state target weight and food insecurity is actually steeper for the *N* = 2 world than the *N* = 10 world. We consider how variation in the energy-density of available food might explain differences between high- and low-income countries in terms of the relationship between food insecurity and obesity in section 6.3 of the main paper.

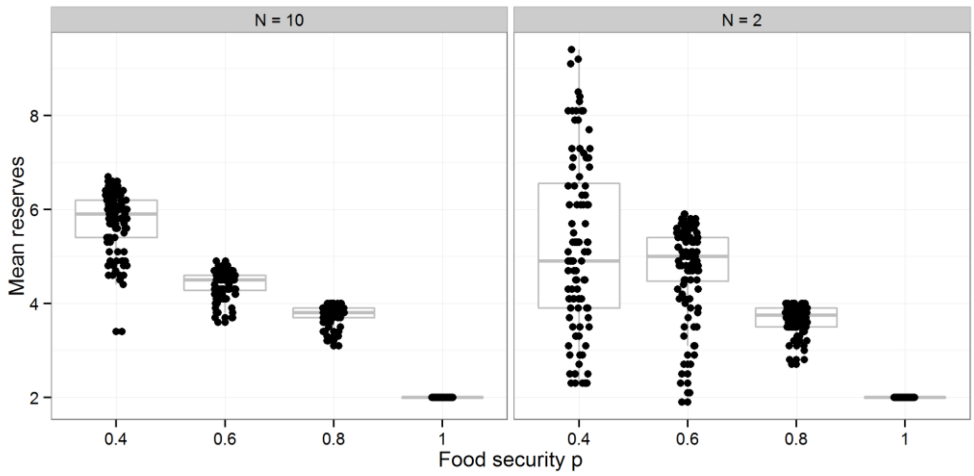

Figure A8. Mean body weight over 40 periods for simulated individuals at different levels of *p*, for two different values of the maximum energy available from food per period, *N.* Points have been jittered in the horizontal dimension to make individual data points more visible. All other parameters have their default values. This figure is reproduced as figure 4 in the main paper.

A4. Implementation and code availability

The model is programmed as a series of functions in R (R Core Development Team, 2015) and available via the Open Science Framework at https://osf.io/zarn6/. There are two R scripts. Sourcing the script 'obesity model functions.r' makes all the functions underlying the model available and allows for the user's own exploration of the model. The second script, 'obesity model for replication.r' produces all the figures and output from this document and the main paper, and relies on 'obesity model functions.r' having previously been run.

The main functions are as follows. The *policy* function gives the optimal eating policy for a given value of *p*, in the form of a vector corresponding to increasing levels of current reserves. Thus *policy(p=0.5)* will produce the a vector of 50 optimal amounts to eat corresponding to reserves of 1-50, for the specified value of *p*. In this and other functions, the parameters *w*, *x*, *y*, *a*, *b* and *N* are given their default values unless otherwise specified (i.e. w=1, x=10, y=0.01, a=1, b=0, N=10; you can vary them by specifying the required value in the function call). The *simulate.food* function gives the weight trajectory of an individual who follows the optimal eating policy for a given value of p and finds food every time period. Its output is a vector indexed by time period. For example, *simulate.food(current.p=0.8, reps=20)* gives a vector of reserve/weight levels corresponding to 20 successive time periods for an individual following the optimal policy for *p=0.8*.

The *simulate* function is the same as *simulate.food* except that food is found with probability p in each time period. Thus, every run is unique. The call *simulate(current.p=0.8, reps=20)* gives a weight history for an individual facing p=0.8 and following the optimal policy. The first ten time periods are removed to avoid initialization artefacts. *NA* means the individual has died. Finally, *run.of.simulations* automates a run of many simulations for varying values of p and saves the output (see script for details).

References

Garby, L., Garrow, J. S., Jorgenssen, B., Lammert, O., Madsen, K., Sorensen, P., & Webster, J. (1988). Relation between energy expenditure and body composition in man: Specific energy expenditure in vivo of fat and fat-free tissue. *European Journal of Clinical Nutrition*, *42*(4), 301–305.

Johnstone, A. M., Murison, S. D., Duncan, J. S., Rance, K. A., & Speakman, J. R. (2005). Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine. *American Journal of Clinical Nutrition*, *82*(5), 941–948. doi:82/5/941 [pii]

R Core Development Team. (2015). R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.