Group 1: Elisa, Jacqueline e Carmen

We found your hypothesis/approach to obstetric selection very interesting.
Thanks!
Our perception was that the theory, although carrying many simplifying assumptions, could be an important step forward in this field. We have been constantly thinking that it would be valuable to test these ideas empirically. However, the article states that D cannot be inferred by usual clinical measurements, limiting the viability of testing the proposed model. (1) In what other way, given the difficulty in measuring D, can we test this model?

One way to explore models systematically, is to test their assumptions. Another is to test the predictions they make. We have found evidence for example that indeed the dystocia appears to be running in female lineages in families, so there is some heritability to it (PNAS 2017). On the prediction side, there are some indications that the incidence of the need for C-sections is increasing. Of course the difficulties in testing such trends are numerous, as the C sections are strongly on the rise and it is hard to know what part of those is due to mismatch, and also because human size is changing with nutrition, and such effect is particularly important when we are talking about 2 generations. Here something for you to think about: what happens when mothers grow up under food restriction, but their fetuses grow in economically much stronger conditions….

Besides that, (2) what do you consider "usual clinical measurements", which are lacking to obtain D? If you consider them to be the measures obtained by prenatal exams, why can't we use those to estimate D? Could we obtain the D distribution of a society from prenatal data? Would it help us to validate your model?

It would help to validate the model if one could estimate the D, yes. There are several obstacles for now. D is a theoretical concept, a combination of maternal and fetal size. However, even if the classical literature shows the problem as if it is about getting a ball through a ring, human birth is actually much more like getting a bit amorphous thing through a narrow, uneven and a bit winding channel, and "the thing" in addition is turning. The clinical measurements taken on mother are several, and usually the narrowest one is considered most important- and yet the predictive power appears to be relatively poor. In fact one of my coauthors, Barbara Fisher, is just trying to get funding to figure out what measurements would be the best predictors- then we might be able to go back and test the model on the base of those.

2. In model 2, you consider that the male fitness increases with D indefinitely. However, we thought of scenarios under which this assumption may not be true. (a) **Monogamy**: it seems that in a monogamous society males which produce a mate/offspring with large D will suffer a decrease in their fitness due to time and resources they spend in the process, and lack of other opportunities. **Fetus selection**: a high D decreases fetus fitness for both male and female fetuses, regardless of their sex. **Fetal viability**: iif D becomes too high due to male driven evolution, even male fitness will decline, since he will have fewer offspring. Do you think these factors explain why the actual selection scenario may vary between models I and II?

You mean that the constraint only affects female fitness, but does not affect male fitness? Those are good points. In general, I agree that one could think of different scenarios to what extent the male fitness is affected. I am actually not a great fan of the ideas of male genes driving this (like conflict theory etc.) – as I think of fetus not only in male + female, but rather as an interaction. So there is much to do here….

Yes, male and female fetuses are affected, but selection is on (interaction with) female there. One thing we reported in a follow up paper (PNAS 2017) is that there appears to be heritability on maternal side (daughters born with difficulties themselves have them when giving birth, but the children of sons born with difficulties, do not)
To the last one- the fetal viability is of course also bound by other things, not only size at birth, so mothers would likely not be able to support such fetuses already much before the birth, so this is beyond what this model is supposed to address.

If the models used in your study do not take these scenarios into account, what would you propose that they be taken into consideration?

Sure there are many ways to extend the model, once the model has been formulated.

Your article states that factors such as flexibility of the pelvic ligaments, orientation of the neonate and efficiency of uterine contractions influence the success of the labor, but are statistically independent of D. However, our intuition is that these factors may be related to measures underlying D. Is it safe to assert that they are statistically independent?

Literally, such question can only be answered by making the test. I think in first approximation the size relationship is likely not directly dependent on many of these things, but I will point you to one biological reason that may cause association and that we have been pursuing in later works. (Keep in mind what the theoretical models are for, namely to explore the most important factors, rather than duplicate the reality. Since 2016, several publications develop then various more or less connected paths of inquiry - most recent summary in Pavlicev, Romero, Mitteroecker, Am J Obstetrics and Gyn 2019)

Would it be worth to try to find data that tests for relationship between these traits?

Of course intuitions are generally worth to try out, it is however hard for me to judge without knowing what your intuition is about. I think it is crucial to think of these data as biological data, rather than mere statistical dependencies. So the intuitions need some grounding in biology. For example, one reason that there may be associations is that smaller pelvises may have stronger pelvic floors. This could be due to an interesting trade-off that we suggested in later studies. Namely, it may be, that narrower pelvis providing a better support for long pregnancy with heavy babies, trades off with harder childbirth.

Grupo 2: Jonas, Giovanni, Matheus e Miguel

4. It occurred to us that the regular use of Caesarean sections could cause a dynamic of a Runaway-like selection process. The idea is that the extensive use of this procedure could lead to a feedback model in which the use of cesarean sections would become essential for birth.
I am not sure that this is a real runaway (runaway would involve not just interdependency but that there is an acceleration of some kind). But I think I understand what you mean.

Or do you think it is more likely that there is a certain threshold for the size of babies heads that we cannot yet identify from current data?

I think models are useful tools for thinking, but one should keep in mind that they are valid locally. This model for example doesn't consider that maternal metabolism also plays a role, because in the narrow window we look at we might be able to neglect it. But another reason for upper limit to the size of the fetuses is maternal metabolism (one of the theories for specific birth timing is that it is the timepoint where mother reaches a metabolic limit).

Mothers undergo important restructuring of their physiology during pregnancy. It is an evolved state, but certainly not with unlimited capacity. Where exactly the limit is, and whether it is in any way correlated with pelvic size, would be worth exploring.

Grupo 3: Gabriel, Fernando e Rafael

5. Do you believe it is possible to introduce to the model the mother's stage of development to infer her aptitude for childbirth? In other words, because age of mothers is a very rapidly changing feature, and differs greatly among countries and cultures, do you think that correlating the mother's childbirth age to the results presented in the study could yield new conclusions?

Of course it is likely that there is an impact of other factors, such as maternal age. However, introducing additional factors into the model requires that we also know *how* these factors affect the variables. So yes, if there is knowledge about that, one can include age as well. To answer the question- it is hard for me to speculate about that. Perhaps I can turn your question back to you, because with your reading you can likely answer such question yourself - what exactly do you have in mind? Follow through the logic of the model and think about what kind of age-related effects would result in what kind of differences in conclusion?

Grupo 4: Arthur, Giovanna, Raul e Victor

6. Analysing graph 1 in figure 1, we might expect that a smaller variance (sigma) would solve the problem of cephalopelvic disproportion, increasing the mean fitness of the population even if not changing the mean of the trait D. Is it possible to investigate if selection favored a reduction in the variance of D? Is there any reason to suspect that there is a heritable component to this trait (i.e., the variance of D), that could be selected to be reduced?

It is correct that change in the spread would minimize the cut off part of the population. Sure, given the appropriate data, providing that D can be measured, one can look at the evolution of variance of a trait. There appears to be a heritable component to D, yes (PNAS 2017).

7. The paper assumes a linear increase of fitness with D, until the cliff edge is reached. Does this make specific assumptions about how fitness is related to head size of pelvis dimensions? Would a different shape to the function of D have an impact in downstream analyses?

To be exact, it is the D which maps linearly in our model to fitness. We do not so far have the data for knowing precisely how the fitness function looks like, except that it is increasing with increasing fetal size. There have been indications of selection for narrow female pelvis as well. But in the lack of any exact idea how precisely the increasing function looks like, there is little incentive to make it more complex than linear. But I would encourage you to try to answer your speculative question yourself- one can think about it in terms of the figure 1—how would the maximal intersection between symmetrical phenotypic distribution and different fitness function look like, and what would be the consequences for the region beyond the cut-off? I think one may soon recognize that the crux is that one distribution is inherently symmetrical, while the other isn't, right?