**Investigating fetal movements in reaction to auditory stimuli in late pregnancy: results of the ANNA trial.**

Matteo Quartagno1, Anna Quartagno1,2 and Thea Ilaria Piovano2

1 Department of Fatherhood, Institute of Child Production, London, UK

2 Department of Motherhood, Institute of Child Production, London, UK

E-mail: m.quartagno@ucl.ac.uk

Received 30 Dec 2019

Accepted for publication 16 Sept 2020

Published 22 Sept 2020

Abstract

**Background.** The fetal auditory system apparently develops around 25 weeks of pregnancy, although I do not really guarantee that this is correct. What is less clear is whether fetuses can actually react to their parents talking to them.

**Methods.** We designed a randomised trial to compare the movement patterns of Anna, our daughter, when her father talked towards her or not. The primary analysis model was a Poisson regression model fully adjusted for all collected baseline variables thought to be associated with the outcome.

**Results.** There was no evidence that Anna moved more when talked at (IRR: 1.06 [0.87,1.31]) nor that she moved more often (OR: 0.96 [0.51, 1.8]). There was a highly significant interaction (p<0.0001) with pregnancy month, suggesting the baby might be more responsive in later pregnancy.

**Conclusion.** We could not prove whether our baby moved differently when talked at in the womb, but future research will investigate whether this was due to us starting the experiment too early.

**Keywords:** Baby, hearing, father, does not care at all

1. Background

The first pregnancy is known to be perhaps the most important moment for a couple. Every *first* pregnancy has itself a lot of *first* things. The *first* positive test, the *first* scan, the *first* kg of body weight gained by the mother and lost by the father to stress, the *first* kick...

All of these things were experienced by the first and last author of this paper (henceforth the father and the mother), while waiting for the second author (henceforth, Anna). In particular, the first fetal movements felt by the mother led to the father talking to her belly, in the hope Anna would listen.

It is well known that the fetal auditory system develops in the first part of pregnancy and becomes fully functional 25 weeks into the pregnancy [1]. What is less known, is whether the baby is actually able to really react to inputs from the outside world. Of course, most families assume that they do. It is common for first time fathers to be told “ohhhh, she really does hear you down there!” as soon as a single well placed kick deforms the mother’s belly seconds after talking to her, similarly to what huge worms did in Tremors [2]. Most fathers are generally happy with this explanation, which makes for a good story, increases the self-esteem of the father, who has otherwise finished his active job and merely acts as an esquire from then on, and makes everybody happy. But most fathers are not statisticians.

1.1 Is it true?

Statisticians are ugly beasts. They are those classmates that make you realise you made a mistake when the test is over and you thought you did great. They are those referees who disallow the most incredible of goals. They are the Michele Scalseggi who tell you that human stomach is itself digestible.

Marrying a statistician is arguably the *non plus ultra* of masochism. Rather than being happy of a wondeful story, a statistician will dig into it, until they are *absolutely certain* that the story is true. Pardon me, until they are *reasonably confident* that the story is true.

So, the first author wanted to test whether it was true that his baby was listening to his voice and responding by kicking after the 25th week of pregnancy, as everybody implies.

To investigate this, there are two options: an observational study or an interventional one. Since we had no damn idea what sort of DAG to expect from this problem, we just decided to use the magic of randomisation. [3] Hence, we designed a randomised controlled trial: ANNA (Ante-Natal Auditory system assessment).

**2. Methods**

 The full trial protocol was written when the trial had already started, as this allowed us to slightly tweak it to favour our research hypothesis. It is available on the GitHub page of the father. [4]

2.1 Eligibility and Intervention

Eligibility criteria included being at least 25 weeks old and living in the mother’s belly. There is a single participant, our second author Anna. Therefore this is technically a N-of-1 trial. However, I never studied N-of-1 trials and have no idea if they are that different to design compared to standard two-arm trials. Therefore, I completely ignored the literature [5] and just designed the trial as a simple two-arm study, where instead of randomising participants I randomised observations.

Observations were randomised to one of two treatments. The control consisted in the father simply putting his hands on top of the mother’s belly for 30 seconds and counting kicks and punches. The active treatment consisted instead of additionally talking towards the belly for the full 30 seconds. As discussed in the protocol, the content of the belly talks remain secret.

*2.2 Randomisation*

Randomisation was performed with a very sophisticated technique which is described in the trial protocol. Observations were randomised with a 1:1 allocation ratio to either of the two arms. Simple randomisation was used. We aimed to randomise around 4-5 observations each day. However, in early July we went on holiday in Italy and, honestly, as geek as I can be, there are better things to do in the Alps than randomised experiments.

Figure 1: Movement patterns over time aggregating the two arms

*2.3 Outcome measures*

The primary outcome measure was the number of movements in a 30 seconds window. A secondary outcome was binary movement, which was able to answer the question “Do my words make her move more often?” rather than “Do my words make her move more?”.

*2.3 Statistical analysis*

The substantive analysis model consisted in a Poisson regression model, adjusted for all collected baseline variables that were thought to be associated with the outcome. These were position of the mother (3-category variable: standing, sitting, supine), distance from the last meal (in hours), time of the day (hours only) and scaled time since start of the study. Continuous covariates were modelled with cubic splines with 3 knots.

For the secondary outcome we repeated the same analyses but using logistic regression instead of Poisson.

We estimated that a sample size of around 200 observation was needed to achieve 95% power to reject the null hypothesis of no difference between the two treatments if the average number of movements in the control arm was 0.75 and in the active arm 1.25.

**3. Results.**

Of the 200 observations collected between May 30th and August 7th, 97 were randomised to silence and 103 to talking. Good balance was achieved on most covariates, though more talking was performed in the afternoon (see Table 1). Since the primary analysis was adjusted for all baseline variables, we do not really care about these small imbalances.

*A screenshot of a cell phone

Description automatically generated*

|  |  |  |
| --- | --- | --- |
|  | Silence | Talking |
| No. | 97 | 103 |
| Part of day – no. (%) |  |  |
| Afternoon | 37 (38.1) | 54 (52.4) |
| Evening | 31 (32.0) | 29 (28.2) |
| Morning | 29 (29.9) | 20 (19.4) |
| Distance (h) from last meal (median [IQR]) | 2.00 [1.00, 3.00] | 2.00 [1.00, 3.00] |
| Position – no. (%) |  |  |
| Sitting | 39 (40.2) | 40 (38.8) |
| Standing | 10 (10.3) | 8 ( 7.8) |
| Supine | 48 (49.5) | 55 (53.4) |
| Pregmonth – no. (%) |  |  |
| Sixth | 30 (30.9) | 38 (36.9) |
| Seventh | 55 (56.7) | 51 (49.5) |
| Eighth | 12 (12.4) | 14 (13.6) |

Table 1: Baseline characteristics stratified by treatment group.

*3.1 Primary outcome*

A screenshot of a cell phone

Description automatically generatedThe average number of movements appeared to be pretty much stable across the study period (Figure 1). Two observations were wildly off (~15 movements in 30 seconds) because Anna had hiccups. Luckily, the two observations were equally divided between groups and after them we decided to amend the protocol to avoid any measurement in presence of hiccups.

The mean number of movements was 1.89 (Median [IQR] = 1 [0; 3]) in the control arm and 1.97 (2 [0; 3]) in the active arm. (Table 2)

The incidence rate ratio (IRR) from the fully adjusted model for talking vs silence was 1.06 (95% CI = [0.87, 1.31], p=0.56). This did not vary much compared to the unadjusted analysis. This provided no evidence against the null hypothesis that talking caused no difference in Anna’s moving patterns.

Figure b: Forest plot of pre-specified subgroup analyses. This is based on code from Stefano Spada that will soon be implemented in an R package.

*3.2 Secondary outcomes and subgroup analyses*

Similar conclusions were drawn when looking at the secondary outcome. When remaining silent, the baby moved 66% of the times, while she moved 67% of the times when being talked.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Analysis** | **Estimate** | **SE** | **IRR/OR** | **95% CI** | **p-value** |
| Poisson - Fully adjusted | 0.06 | 0.11 | 1.06 | [ 0.87 ,  1.31 ] | 0.56 |
| Poisson - Unadjusted | 0.04 | 0.1 | 1.04 | [ 0.86 ,  1.28 ] | 0.67 |
| Logistic - Fully adjusted | -0.04 | 0.32 | 0.96 | [ 0.51 ,  1.8 ] | 0.9 |
| Logistic - Unadjusted | -0.04 | 0.3 | 0.96 | [ 0.53 ,  1.72 ] | 0.88 |

Once again, adjusting barely made any difference. This is also because the baseline variable most predictive of the outcome (position) was also the one that ended up being more balanced (swear we did not cheat).

Figure 2 is a forest plot showing the results of all the pre-specified sub-group analyses performed. The p-value for interaction with pregnancy month was smaller than our baby, suggesting the effect of talking varied throughout the experiment. In particular, the point estimates and CIs suggest Anna became more and more responsive to outside stimuli with the progression of pregnancy.

However, there are some important caveats: first, very few observations were taken in the eighth month, while in Italy. Second, the difference between the sixth and seventh month might be partly due to the fact that the hiccup measurement in the sixth month was in the silence arm, while the one in the seventh was in the talking arm. Removing these two observations, the CI for the sixth month crosses 1, but the p for interaction remains similarly significant.   
The p for the other interactions remains somewhat larger, with the CI for the IRR in the standing position not crossing 1, but based on very few observations.

Table 2: Results of adjusted and unadjusted primary and secondary analysis models

**4. Discussion**

The ANNA trial investigated whether daddy was allowed to brag about the fact that Anna moved in reaction to his words. Results suggest that there is nothing to brag about. While absence of evidence is different from evidence of absence, our trial had very high power to detect a very small difference (half a movement more on average every 30 seconds, or 1 more movement every minute) and so, even if there was any difference, our trial suggest it to be so small to be hardly of any relevance and surely undetectable by mommies and daddies around the world without conducting proper randomised experiments.

Our pre-specified subgroup analyses suggested that there might be an interaction between treatment and pregnancy month. However, as we described in the results section, this might partially be caused by hiccups experienced in the control arm early on and in the active arm later. The proper way to confirm this result would be to repeat the study for anther pregnancy but limited to the eighth month. Future work will determine whether the first and last author will have the strength to embark in such an endeavor.

One important caveat is that our trial might only apply to Anna, and not to every child around the world. In particular, after designing this stupid experiment, when seeing the results the father realised that if the baby turned up being deaf, this would have haunted him for the rest of his days. Luckily, hearing test at birth was fine, but this does not exclude that Anna might be a bit of a sociopath like her father, and hence less responsive than most babies.

Randomisation has been very useful to answer this question, but it could be considered unethical to solve other problems that are now affecting the life of the mother and the father, like understanding why Anna cries. No ethics committee could possibly accept the idea of randomising Anna to receiving no milk when shouting, although to be honest even accepting to let her hear her father singing Leonard Cohen’s Halleluja for 30 seconds was borderline acceptable.

The trial was open-label, because it was considered impossible to perform placebo talking given the inability of the father to act as a ventriloquist. Because of this, we could not rule out the possibility that Anna might be a superior being, messing with her parents and doctoring the results of the trial.

Finally, both the active and control arms involved the father placing his hands on the mother’s belly to feel the movements. This was a necessary feature of the design, because the mother could not feel them as well. However, this could change the results in case the light pressure coming from the father’s hands annoyed Anna so much not to care about whether someone was talking or not.

*4.1 Conclusion*

Anna most likely did not care at all about her father talking to her while busy swimming in her personal thermal bath in the womb. Future research will explore whether this changes later on in pregnancy.

Acknowledgements

The authors would like to thank Andrew Copas, Tim Morris and Michele Scalseggi for reviewing the trial protocol and giving very valuable suggestions, that the authors decided to completely ignore out of lazyness. They would also like to thank Stefano Spada for providing code to produce the forest plot and reviewing the protocol. The father would additionally like to thank the mother who, on top of bringing around his daughter in her belly for 9 months and feeding her for an indefinite period afterwards, accepted to take part in this pointless experiment in the name of science. Finally the mother and the father would like to thank the second author for making their lives a bit messier, but a lot happier.

References

1. Checked in on Google, seems to be true, [DOI:16092020](https://www.google.it/search?source=hp&ei=icppX7jUMOeflwTE2KzIAg&q=when+fetal+auditory+system+develops&oq=when+fetal+auditory+system+develops&gs_lcp=CgZwc3ktYWIQAzIICCEQFhAdEB46CAgAELEDEIMBOgUIABCxAzoLCC4QsQMQxwEQowI6AggAOggILhCxAxCDAToCCC46CAguEMcBEKMCOggILhDHARCvAToFCC4QsQM6BggAEBYQHjoICAAQhgMQiwNQxgVYg1Jgh1NoAHAAeACAAbEBiAGPH5IBBTE0LjIwmAEAoAEBqgEHZ3dzLXdpergBAg&sclient=psy-ab&ved=0ahUKEwj4vPT1v_zrAhXnz4UKHUQsCykQ4dUDCA0&uact=5).
2. Tremors (1990), Ron Underwood. [Wiki.](https://en.wikipedia.org/wiki/Tremors_(film))
3. Randomised Controlled Trials during peacetime. The Upper Quartile. [Link](https://theupperquartile.wixsite.com/matteo/post/randomised-controlled-trials-in-peacetime).
4. Protocol of ANNA: ANte-Natal Auditory system assessment. [Link](https://github.com/Matteo21Q/ANNA/blob/master/ANNA%20-%20Protocol.docx).
5. Duan, N, Kravitz, RL, Schmid, CH. Single-patient (n-of- 1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. J Clin Epidemiol 2013; 66: S21–S28.