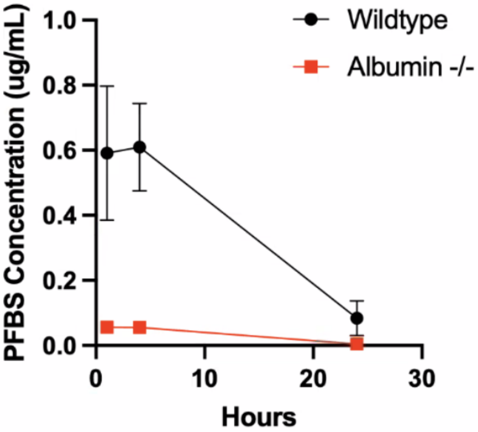
Experimental design and statistical analysis

Marie-Abele Bind

STEEP Trainees seminar

March 2025

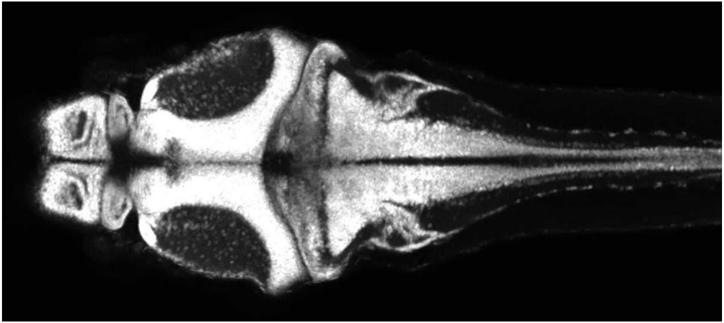
STEEP Project 3



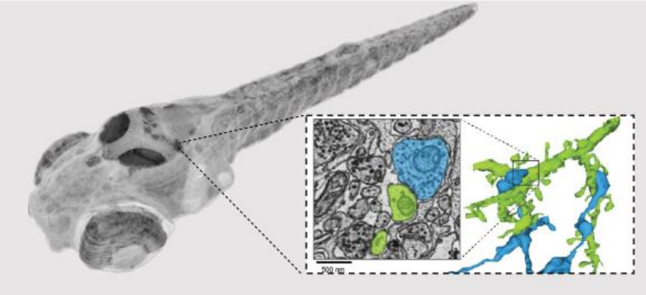
Florian : "You don't need statistics in neuroscience"

Florian : "You don't need [no stinking] statistics in neuroscience"

Florian's lab

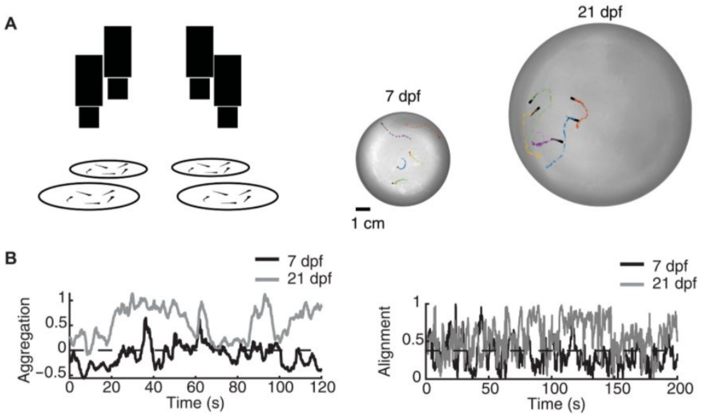


Florian's lab

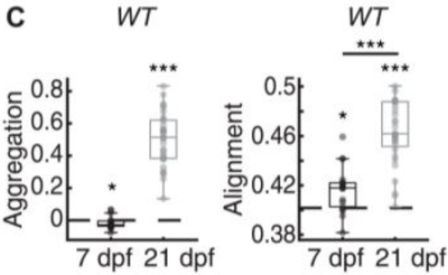


Florian's lab

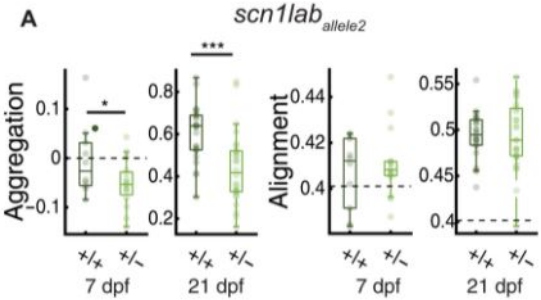
Neuroscience application



Neuroscience application



Neuroscience application



Paper submitted to Science Advances (April 2021)

Collective behavior emerges from genetically controlled simple behavioral motifs in zebrafish

Ariel C. Aspiras*[1], Roy Harpaz*[2,3], Sydney Chambule[1], Sierra Tseng[1], Florian Engert[2,3], Mark C. Fishman[1#§] & Armin Bahl[2,3,4§]

1 Harvard Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge 02138, USA.

2 Department of Molecular and Cellular Biology, Harvard University, Cambridge 02138, USA.

3 Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA.

4 Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz 78464, Germany.

Reviewer's feedback

- "The authors repeatedly and systematically misinterpret the p-values calculated for their statistical tests and consequently make a large number of false and misleading claims about the meaning of their results.

Reviewer's feedback

- The most charitable explanation for this is that the authors are simply ignorant of how to correctly interpret p-values despite their extensive shared expertise of quantitative data analysis."

Reviewer's feedback

- "The authors do not discuss effect sizes in any meaningful way. If the authors do not discuss effect sizes and what they mean biologically (some of which are very small and probably not meaningful) then they are abdicating their central role as scientists."

Reviewer's feedback

- "The reasoning for the choice of tests as well as the assumptions of each test and how these assumptions were validated for the data to which the tests were applied are not explained."

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.
- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.
- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.
- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., $W_i = 1$) or control (i.e., $W_i = 0$).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i = w)$ is a function of unit i and treatment w.

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.
- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i$=w) is a function of unit i and treatment w.
- For each unit i=1,..., N :
  - $Y_i(W_i$=1)=value of Y when i is exposed to the active treatment
  - $Y_i(W_i$=0)=value of Y when i is exposed to the control treatment.

Simple randomized experiment

- Consider a simple treatment-control experiment with a completely randomized assignment mechanism.
- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i$=w) is a function of unit i and treatment w.
- For each unit i=1,..., N :
  - $Y_i(W_i$=1)=value of Y when i is exposed to the active treatment
  - $Y_i(W_i$=0)=value of Y when i is exposed to the control treatment.
- Because of SUTVA, the science table simplifies to two columns and N rows.

Neuroscience experiment

- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.

Neuroscience experiment

- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.
- Consider N groups of 5 zebrafish, indexed by i, in an experiment hypothetically randomized to a binary genotype G, either mutant (i.e., $G_i = 1$) or wildtype (i.e., $G_i = 0$).

Neuroscience experiment

- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.
- Consider N groups of 5 zebrafish, indexed by i, in an experiment hypothetically randomized to a binary genotype G, either mutant (i.e., G_i=1) or wildtype (i.e., G_i=0).
- An outcome variable Y (e.g., aggregation, alignment at 7 dpf) is measured for mutant and wildtype zebrafish.

Neuroscience experiment

- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.
- Consider N groups of 5 zebrafish, indexed by i, in an experiment hypothetically randomized to a binary genotype G, either mutant (i.e., G_i=1) or wildtype (i.e., G_i=0).
- An outcome variable Y (e.g., aggregation, alignment at 7 dpf) is measured for mutant and wildtype zebrafish.
- No hidden mutation and no interference between groups of zebrafish (SUTVA). $Y_i(G_i$=g) is a function of unit i and genotype g.

Neuroscience experiment

- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.
- Consider N groups of 5 zebrafish, indexed by i, in an experiment hypothetically randomized to a binary genotype G, either mutant (i.e., G_i=1) or wildtype (i.e., G_i=0).
- An outcome variable Y (e.g., aggregation, alignment at 7 dpf) is measured for mutant and wildtype zebrafish.
- No hidden mutation and no interference between groups of zebrafish (SUTVA). $Y_i(G_i=g)$ is a function of unit i and genotype g.
- For each group of 5 larval zebrafish i=1,..., N :
  - $Y_i(G_i=1)$=value of Y when i is mutant
  - $Y_i(G_i=0)$=value of Y when i is wildtype.

Neuroscience experiment

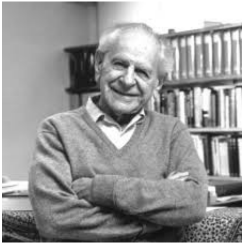
- Consider a simple mutant-wildtype zebrafish experiment with an hypothetical completely randomized assignment mechanism.
- Consider N groups of 5 zebrafish, indexed by i, in an experiment hypothetically randomized to a binary genotype G, either mutant (i.e., G_i=1) or wildtype (i.e., G_i=0).
- An outcome variable Y (e.g., aggregation, alignment at 7 dpf) is measured for mutant and wildtype zebrafish.
- No hidden mutation and no interference between groups of zebrafish (SUTVA). $Y_i(G_i=g)$ is a function of unit i and genotype g.
- For each group of 5 larval zebrafish i=1,..., N :
  - $Y_i(G_i=1)$=value of Y when i is mutant
  - $Y_i(G_i=0)$=value of Y when i is wildtype.
- Because of SUTVA, the science table simplifies to two columns and N rows.

Quote #1

"Science must begin with myths, and with the criticism of myths."

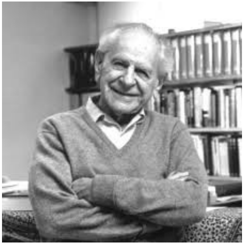
Quote #1

"Science must begin with myths, and with the criticism of myths."



Quote #1

"Science must begin with myths, and with the criticism of myths."
Popper, 1962



Quote #2

- "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise".

Quote #2

- "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise".

Quote #2

- "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise".

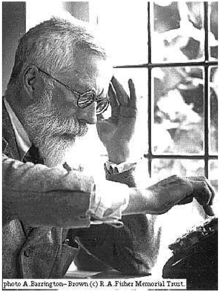
Tukey, 1962

Randomization-based inference

- Calculating a Fisher-exact p-value is superior to the current, more common, practice of calculating an approximating asymptotic (i.e., large sample) p-value.

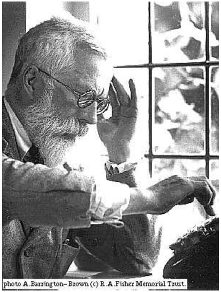
Randomization-based inference

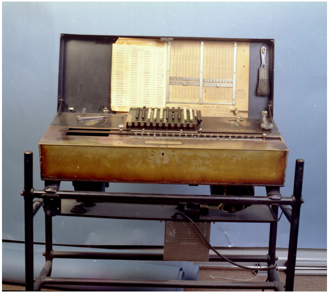
- Calculating a Fisher-exact p-value is superior to the current, more common, practice of calculating an approximating asymptotic (i.e., large sample) p-value.



Randomization-based inference

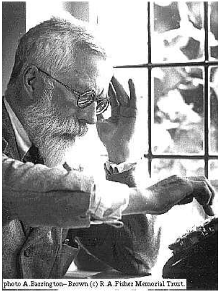
- Calculating a Fisher-exact p-value is superior to the current, more common, practice of calculating an approximating asymptotic (i.e., large sample) p-value.





Randomization-based inference

- Calculating a Fisher-exact p-value is superior to the current, more common, practice of calculating an approximating asymptotic (i.e., large sample) p-value.



Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.

  Example of myth : $\forall i \in \{1, ..., N\} \quad Y_i(G_i{=}1) = Y_i(G_i{=}0)$

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.

$$\underline{\text{Ex. of T}} : \quad T(G, Y^{obs}) = \frac{\frac{1}{N_t}\sum_{i:G_i=1} Y_i^{obs} - \frac{1}{N_c}\sum_{i:G_i=0} Y_i^{obs}}{\sqrt{\frac{s_t^2}{N_t} + \frac{s_c^2}{N_c}}}$$

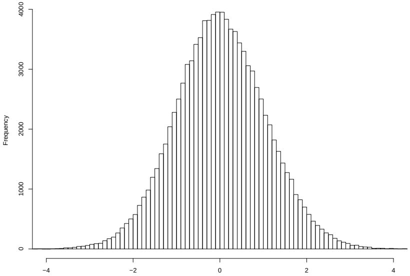
  - Statistic : function of the observed data.
  - Sensitive to expected departures from the null hypothesis.
  - Example of more extreme : $T(W, Y^{obs}) < T^{obs}$

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.

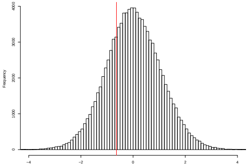
Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



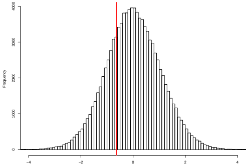
Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



Stochastic proof by contradiction

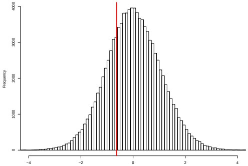
- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



  3. Locate the observed value of the test statistic, T^{obs}, in the null randomization distribution constructed in Step 2.

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate scalar test statistic, T, and define more extreme.
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



  3. Locate the observed value of the test statistic, T^{obs}, in the null randomization distribution constructed in Step 2.

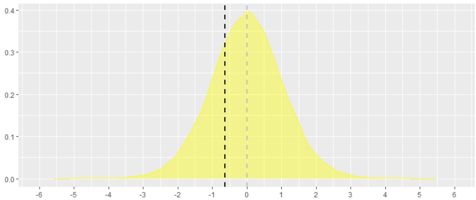
The Fisher-exact p-value corresponds to the proportion of values of the test statistic that are as extreme or more extreme than the observed value of that test statistic.

Beyond hypothesis testing

- For the alignment experiment, we could have reported a counternull value, which was introduced by Rosenthal and Rubin (1994) as a nonnull value of an effect (e.g., difference in alignment between *scn1lab*$^{+/+}$ and *scn1lab*$^{+/-}$) that is supported by exactly the same amount of evidence as the null value of the effect.

Counternull sets in randomized experiment

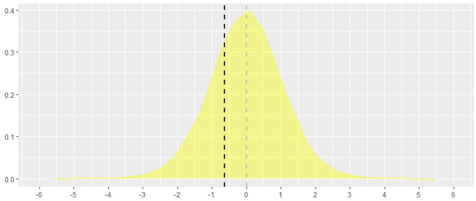
- Here, we can equate the randomization-based null and counternull p-values.



Counternull sets in randomized experiment

- Here, we can equate the randomization-based null and counternull p-values.
  - H_{Null} states that for each unit i,

$$Y_i(W_i = 1) - Y_i(W_i = 0) = 0$$



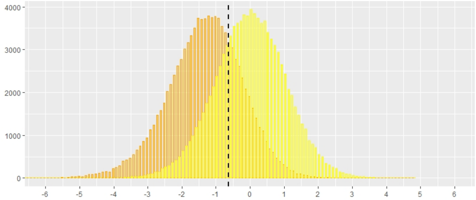
Counternull sets in randomized experiment

- Here, we can equate the randomization-based null and counternull p-values.
  - H_{Null} states that for each unit i,

$$Y_i(W_i = 1) - Y_i(W_i = 0) = 0$$

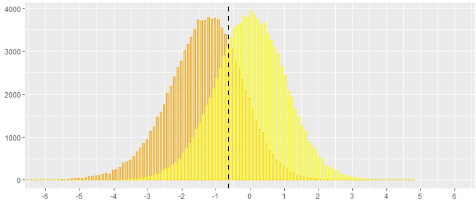
  - $H_{Counternull}$ states that for each unit i,

$$Y_i(W_i = 1) - Y_i(W_i = 0) = a \quad (a \neq 0).$$



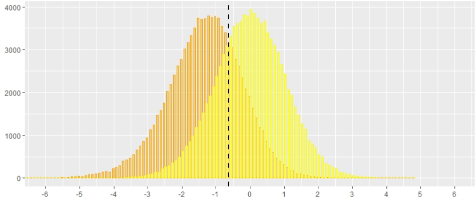
Counternull sets in randomized experiment

- Here, we can equate the randomization-based null and counternull p-values.
  - H_{Null} states that for each unit i,
  $$Y_i(W_i = 1) - Y_i(W_i = 0) = 0$$
  - $H_{Counternull}$ states that for each unit i,
  $$Y_i(W_i = 1) - Y_i(W_i = 0) = a \quad (a \neq 0).$$
- Equate proportion of values of the statistic that are as "unusual or more unusual" than T_{obs}.



Counternull sets in randomized experiment

- Here, we can equate the randomization-based null and counternull p-values.
  - H_{Null} states that for each unit i,
$$Y_i(W_i = 1) - Y_i(W_i = 0) = 0$$
  - $H_{Counternull}$ states that for each unit i,
$$Y_i(W_i = 1) - Y_i(W_i = 0) = a \quad (a \neq 0).$$
- Equate proportion of values of the statistic that are as "unusual or more unusual" than T_{obs}.
- Counternull set =[-0.00762456 ; -0.00762451]



Advantage of reporting counternull values

- The counternull approach helps avoid misinterpretations when testing a null hypothesis.

Advantage of reporting counternull values

- First, the counternull set is associated with a p-value ≈ 0.27. My colleague did not reject the null and subsequently was tempted to accept the null hypothesis. In this situation, reporting the counternull set forced a discussion on accepting the null value, but also the counternull set, which did not contain substantial values!

Advantage of reporting counternull values

- Now consider a counternull value of an effect that is associated with a low p-value, but with low magnitude. In this situation, the counternull value is also worth reporting, because, if small, it would indicate that my colleague's intervention is not necessarily biologically relevant.

Happy ending


ScienceAdvances

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE ADVANCES > VOL. 7, NO. 41 > COLLECTIVE BEHAVIOR EMERGES FROM GENETICALLY CONTROLLED SIMPLE BEHAVIORAL MOTIFS IN ZEBRAFISH

RESEARCH ARTICLE | NEUROSCIENCE

Collective behavior emerges from genetically controlled simple behavioral motifs in zebrafish

ROY HARPAZ, ARIEL C. ASPIRAS, SYDNEY CHAMBULE, SIERRA TSENG, MARIE-ABÈLE BIND, FLORIAN ENGERT, MARK C. FISHMAN, AND ARMIN BAHL Authors Info & Affiliations

SCIENCE ADVANCES • 6 Oct 2021 • Vol 7, Issue 41 • DOI: 10.1126/sciadv.abi7460

PDF
Help

Discussion

- The counternull hypothesis does not have to be about constant and additive effects.

  An alternative counternull hypothesis ($H_{Counternull}$) is :

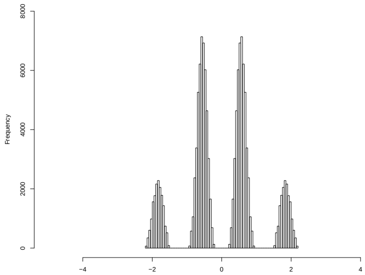
  $Y_i(W_i = 1) - Y_i(W_i = 0) = 0$ for all male units ($a \neq 0$)
  $Y_i(W_i = 1) - Y_i(W_i = 0) = a$ for all female units.

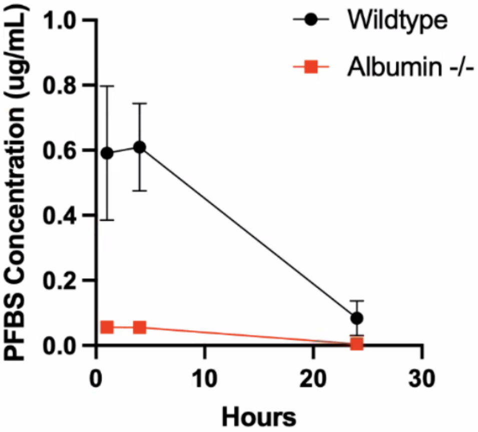
Discussion

- The counternull hypothesis does not have to be about constant and additive effects.

  An alternative counternull hypothesis ($H_{Counternull}$) is :

  $Y_i(W_i = 1) - Y_i(W_i = 0) = 0$ for all male units ($a \neq 0$)
  $Y_i(W_i = 1) - Y_i(W_i = 0) = a$ for all female units.




PFBS Concentration (ug/mL)
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
Hours
Wildtype
Albumin -/-


thank you