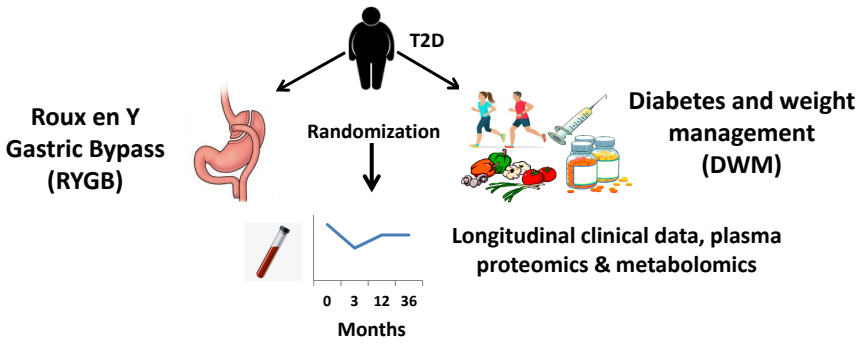
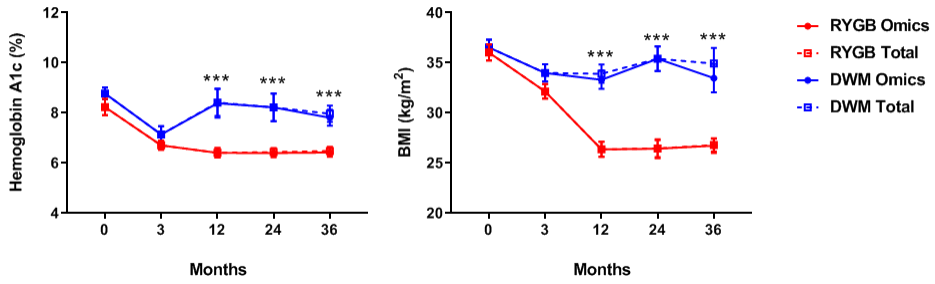
# *High-throughput mediation analysis (Hitman) of human plasma proteome and metabolome identifies mediators of post-surgical diabetes control*

Jonathan Dreyfuss, PhD
Instructor in Medicine
Co-Director
Bioinformatics & Biostatistics Core
Joslin Diabetes Center

# SLIMM-T2D trial



# HbA1c & BMI over time

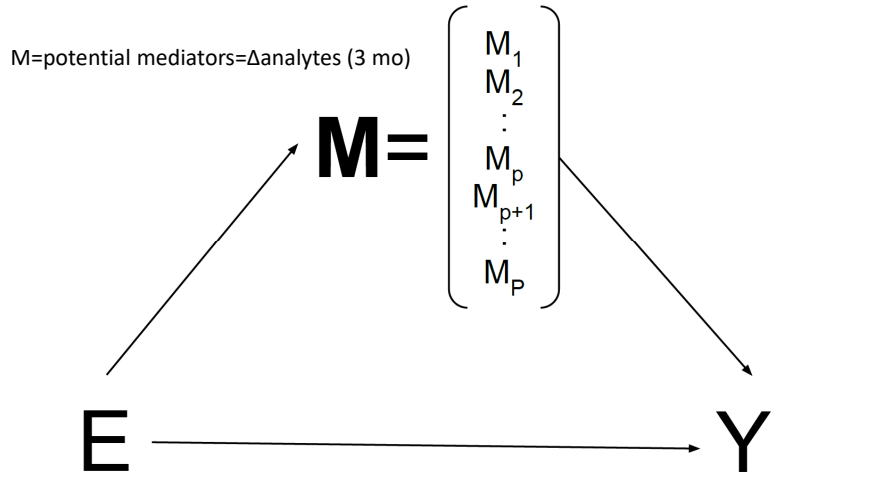



# Goal

- Identify if any analyte changes are part of the mechanism of RYGB, and so can be modulated non-surgically to improve diabetes control
- So want mediation analysis

# Causal model with all mediators

M=potential mediators=Δanalytes (3 mo)

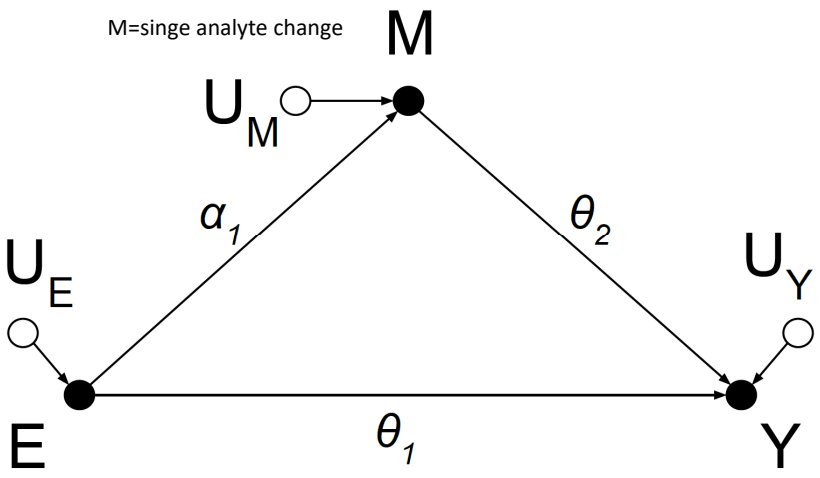



E=exposure=RYGB

Y=outcome=ΔHbA1c (1 yr)

# Causal model per mediator

M=singe analyte change



## Linear structural equation model (SEM)

- Linear SEM per mediator:
  - $E=\nu_0+U_E$ (1)
  - $M=\alpha_0+\alpha_1 E+U_M$ (2)
  - $Y=\theta_0+\theta_1 E+\theta_2 M+U_Y$ (3)
- Indirect or mediation effect = $\alpha_1\theta_2$
- Direct effect = $\theta_1$

## Mediation approaches

- Joint significance (JS) method tests $\alpha_1=0$ & $\theta_2=0$, so $p=\max(p_1, p_2)$
- Product method tests $\alpha_1\theta_2=0$
- Difference in coefficient method considers the model $Y=\lambda_0+\lambda_1 E+U_Y$ and tests $\lambda_1-\theta_1=0$
- Bootstrap method tests $\alpha_1\theta_2=0$ by comparing product to bootstrapped products
- Potential outcome-based method tests difference of potential outcomes between having Δanalyte from RYGB vs Δanalyte from DWM. Tests with bootstraps or other sampling.

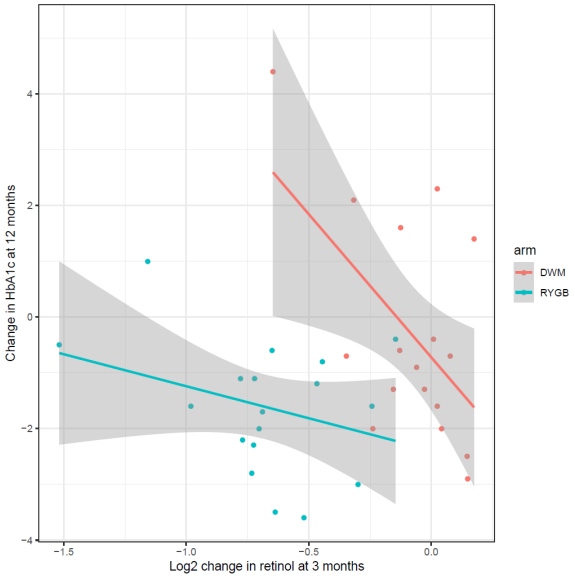
## Two comparisons chose JS test

- MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods*. 2002;7(1):83–104.
- Barfield R, Shen J, Just AC, et al. Testing for the indirect effect under the null for genome-wide mediation analyses. *Genet Epidemiol*. 2017;41(8):824–833.

## Drawbacks of joint significance test

- Not designed for omics datasets, so doesn't exploit empirical Bayes modelling of analyte's variance, e.g. with the linear modeling R package *limma*
- Does not penalize *inconsistent* mediators, where $\alpha_1$ & $\theta_2$ are significantly different from 0, but direction of mediation is not consistent with direction of total effect of E on Y

# Inconsistent mediator: retinol




- RYGB → ↓HbA1c
- RYGB → ↓ retinol
- ↓retinol → ↑HbA1c
- Retinol mediation is in **wrong** direction

# Hitman: limma

- Limma models analyte abundance as dependent variable
- Can test $\alpha_1$ in $M=\alpha_0+\alpha_1 E+U_M$ (eqn 2)
- Can't test $\theta_2$ in $Y=\theta_0+\theta_1 E+\theta_2 M+U_Y$ (eqn 3)
- Instead, regress $Y=\lambda_0+\lambda_1 E+\varepsilon_Y$
- Then test $\theta_2=0$ with limma in $u_M=\varepsilon_Y*\theta_2+\varepsilon$

## Hitman: direction

- Considers direction of effect, allowing “one sided” testing
  - If consistent, $p_{Hitman} = p_{JSL}/2$
  - Else, $p_{Hitman} = 1 - p_{JSL}/2$
  - So, retinol $p_{Hitman} = 0.97$
- $p_{JSL}$ is joint significance implemented with limma

## Simulation settings

- Simulations followed those of Barfield et al. & Mackinnon et al.
- We simulated all combinations of $\alpha_1$, $\theta_2$ in (0, 0.14, 0.39), which correspond to “zero”, “small,” and “medium” effect size
- We tested in what proportion of 10,000 simulations with 50 samples the mediator achieved a p-value  0.05

# Simulation results

| α1 | θ2 | Hitman | joint | mediate |
|---|---|---|---|---|
| 0 | 0 | 0.0057 | 0.0022 | 0.002 |
| 0 | 0.14 | 0.0124 | 0.0062 | 0.005 |
| 0.14 | 0 | 0.0157 | 0.008 | 0.007 |
| 0 | 0.39 | 0.0399 | 0.0364 | 0.03 |
| 0.39 | 0 | 0.0477 | 0.0375 | 0.028 |
| 0.14 | 0.14 | 0.0629 | 0.0232 | 0.015 |
| 0.14 | 0.39 | 0.206 | 0.1084 | 0.103 |
| 0.39 | 0.14 | 0.2067 | 0.1112 | 0.127 |
| 0.39 | 0.39 | 0.6824 | 0.5389 | 0.574 |

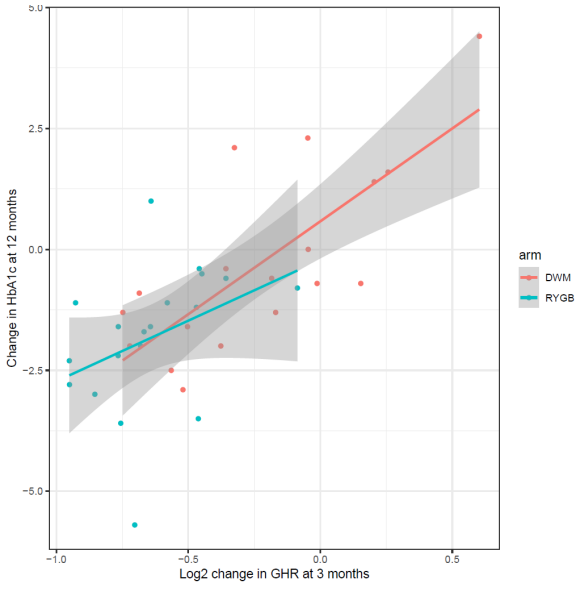
# Few mediation methods for high-throughput

- They try to remove confounding by decomposing measured analytes into latent, independent mediators
- But latent variables are hard to interpret & confounders often unmeasured

# SLIMM-T2D Analyte mediators

- Top protein mediator
  - Growth Hormone Receptor (GHR)
- Top metabolite mediator
  - Proline-hydroxyproline
- These have stronger p-value than body mass index (BMI)
- These are consistent with previous findings, but more experiments needed to confirm their role as mediators

# Consistent mediator: GHR



## Hitman limitations

- There are possible confounders that affect both $U_M$ and $U_Y$, so Hitman, like other mediation analyses, is exploratory
- Based on linear SEM, so assumes Y (perhaps after being transformed) can be modeled by linear regression model with additive E and M

## Conclusion

- Hitman improves on best-in-class joint significance method for high-throughput mediation analysis
- Hitman identifies robust mediators of gastric bypass, which can be further experimentally assayed and possibly serve as future non-surgical therapeutic targets

Thanks!