Original and Replication Research with the NLSY79 and NLSY97 Kinship Links

Will Beasley, Michael Hunter, David Bard,
Mason Garrison, Patrick O'Keefe, & Joe Rodgers

BGA 2019

I'm Will, a member of Joe Rodgers' team that has been using the NLSY since the early 90s.

Two and Half Samples

NLSY79 42k familial relationships

- Gen1: 12k randomly drawn teenagers in 1979
- Gen2: 11k children of Gen1 moms

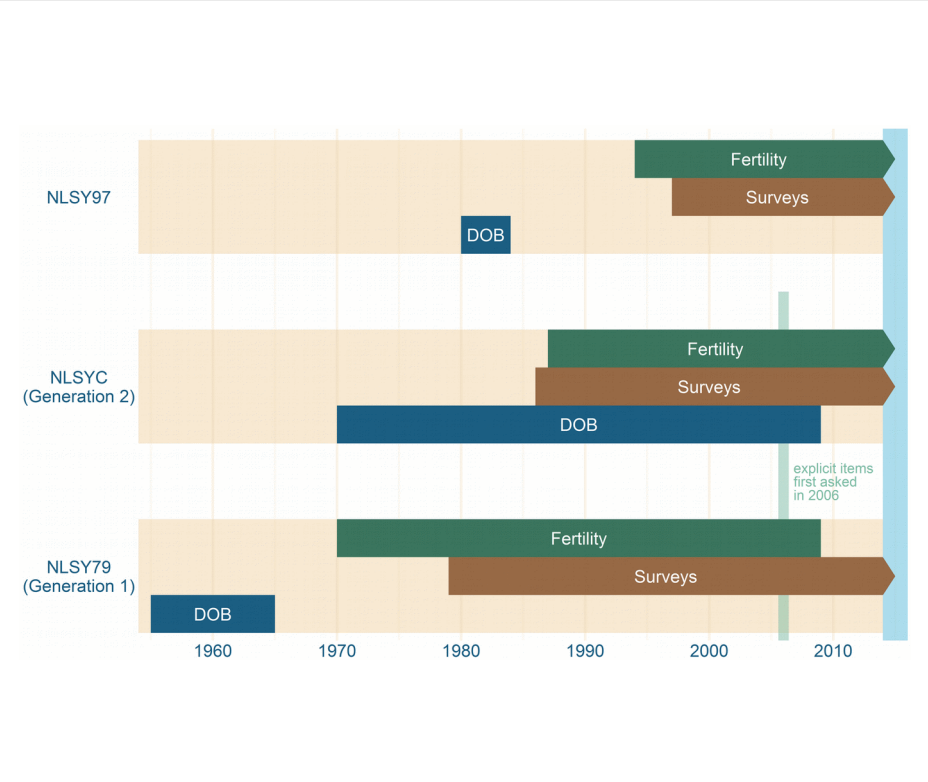
NLSY97 2.5k sibling relationships

- Gen1: 10k randomly drawn teenagers in 1997
  Only 1.8k have 1+ sibling

For our team's first 20 years, almost all our effort was towards creating the 42k kinship links within the two generations of the NLSY79.

Our current grant created the NLSY97 kinship links, which has 2,500 sibling relationships among 1,800 participants.

The 97 sample is completely independent of the 79. Different families.



Here's a quick glance of the samples. At the bottom you see the 79 Gen1s were born around 1960, and have been answering questions for forty years.

In the middle, the Gen2s have been answering questions almost that long.

At the top you see the 97s were born around 1980, and have been answering questions for twenty years

Two and Half Samples

NLSY79: 42k familial relationships

- Gen1: 12k randomly drawn teenagers in 1979
- Gen2: 11k children of Gen1 moms

NLSY97: 2.5k sibling relationships

- Gen1: 10k randomly drawn teenagers in 1997
  Only 1.8k have 1+ sibling

The 97 linking effort went smoother than I expected, but the 79 probably makes the most sense for beginning BGers. It has a larger volume of links, and also a huge variety of relationships. Including full sibs, half sibs, adoptive sibs… Parent-child. Aunt-Niece. Lots of cousins.

But the 97 items cover several topics the 79 doesn't, which might better support your research.

Also, you could use both 79 & 97 for an additional instance of replication.

BG with the NLSY

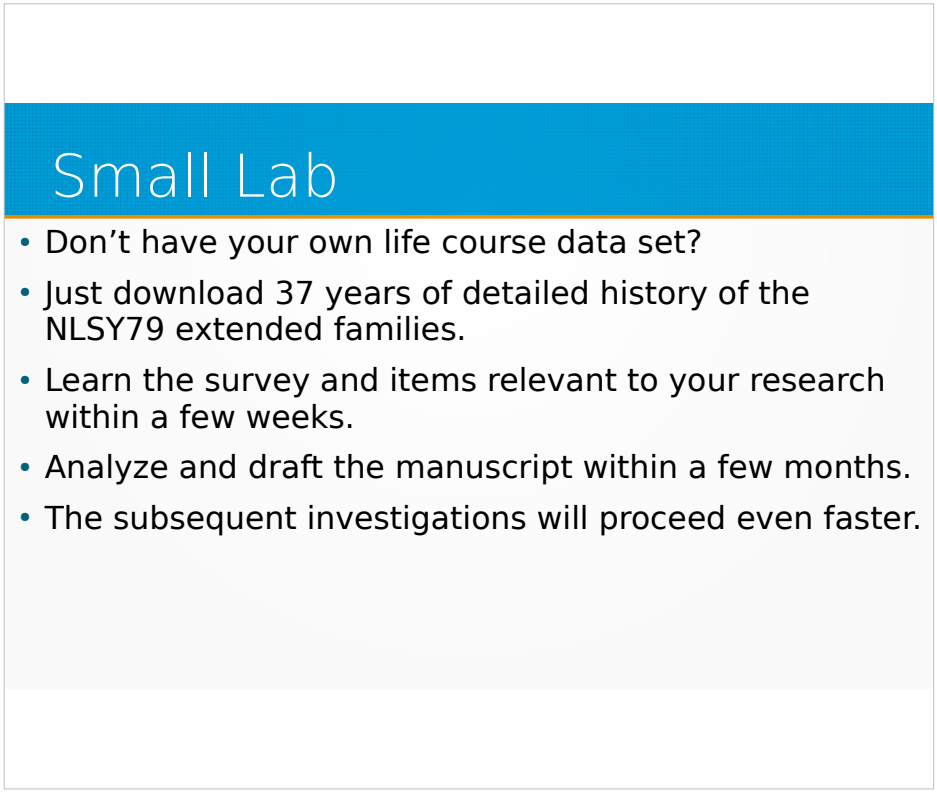
Can benefit 3+ types of BG labs:

1. small labs
2. maturing labs
3. distinguished labs

BG is a small part of my professional life. Most of my time I'm running the Clinical Data Warehouse of our Med School.

And the NLSY provides a lot of the same benefits to a variety of labs: small, medium, & big.

But better than a CDW, the NLSY is a nationally representative sample.



Here are the benefits if you’re a small lab. You probably don’t posses the time or funding to establish your own life course dataset. Yet you can immediately download 37 years of professionally-developed & groomed items.

It will take a few weeks to learn the survey, and a few more months to analyze and write about your findings.

It’s very realistic to submit a manuscript within a semester.

Maturing Labs

- Planning which topics and instruments to introduce in your expanding twin study?
- Run similar analyses on this nationally-representative sample while writing your grant proposal.
- The results can justify which topics & items are likely to be (un)fruitful with your sample.

And the Small Lab benefits apply to you too.

If you have a growing twin study, the 40 years of NLSY items can inform which topics and instruments make the cut of your upcoming waves. Estimate population trajectories. Estimate b/n & w/ n subject variability. Estimate power.

Use it for your internal planning and also the justification in your funding proposals.

For our clinical data warehouse, at least 25% of the projects aren't research themselves, but feasibility assessments incorporated in funding proposals.

Distinguished Lab

- Already manage your own productive twin study?
- Leverage the thousands of economic, cognitive, education, and health variables
- Replicate and further support your substantive findings.

The Small & Maturing lab apply to you too.

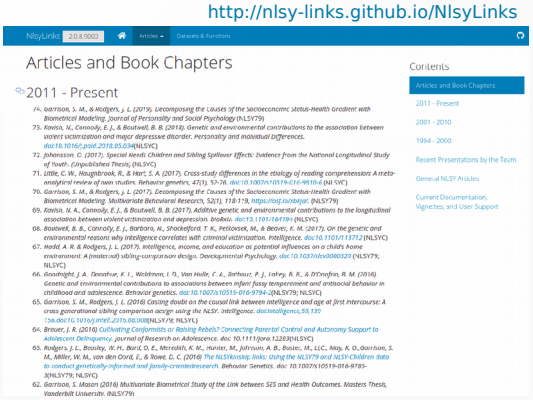
Finally, no matter how huge or established or molecular your twin study is, your conclusions can always be strengthened by replication in an independent sample. Especially a nationally representative random sample.

Like the small labs, you can replicate your findings within a semester.

Not Just Hypothetical Productivity

- 70+ articles across ~8 teams have been published using our links.
- Only the surface has been scratched of the thousands of NLSY items.

http://nlsy-links.github.io/NlsyLinks



I'm confident because we've seen so many publications using our kinship links from all ages and outlets. At least 10 in this list of 70 are theses and dissertations.

And the topics investigated are nowhere close to being exhausted.

The kinship links are stable, but we still add to the 79 Gen2s as more information becomes available each survey wave.

I'd like to continue to make NLSY research quicker to start and easier to build. Some of our future plans focus on the NLSY elements to apply to all investigations, not just BG.

We've always been funded to support other researchers using the links I started a Google Groups recently to move some of that public, but we'll still happily answer emails.