Recruitment, Retention and Collection of Data with a Focus on Small or Hard to Reach Populations

Tracy Onega, PhD, MA, MS
Geisel School of Medicine at Dartmouth & the Norris Cotton Cancer Center
CENTER FOR SCIENTIFIC ACCURACY OF Clichés

NEEDLE IN A HAYSTACK DIVISION

For the first time, scientists were close to determining how difficult it is to actually find a needle in a haystack.
Framing the Issues

Small

Hard-to-reach

Three distinct populations?

With distinct issues and approaches?
Improving Research in Small/HTR Populations

........a tale of two tasks.......
Commonalities in Reaching Populations

- Mixture (albeit varying) of ‘boots on the ground’ with remote reach
- Increasingly relying on technology
- Always predicated on knowledge of population
- Data collection / measurement objective(s)
- Must work across phases: recruitment, retention, etc.
Distinctions in Reaching Populations

- Settings – urban, rural, specific venues, distributed
- Sampling frame – individuals, providers, communities
- Sampling strategy – snowball, RDS, IFWS, etc.
- Technology v. human components variably effective
- Barriers vary: linguistic, cultural, technological, geographic, etc.
- Heterogeneous criteria for “small and/or hard-to-reach”
Model of Components

Recruitment → Retention → Data Collection → Intervention

Informing and improving methods
informing outcomes/effectiveness

Recruitment → Retention → Data Collection → Intervention
Innovative Strategies

**Relationship-based**

Dr. Sanders Thompson
- Research fellows training program (CRFT)
- Mirroring community

Dr. Scutchfield
- CHNA
- Hub and spoke model

**Technology-based**

Dr. Mooney
- Social media, EHRs, wearables, remote sensors, video, telecommunications
For Dr. Mooney

- You have a nice example of combining person-based and technology-based methods. Are there populations and/or settings when person precedes technology or vice-versa for best effectiveness?

- You incorporate data collection across phases. This seems crucial, but what unique challenges does this pose?
Posing Key Questions

For Dr. Scutchfield

You gave excellent examples of community partnerships. These efforts seem broadly targeted, which can maximize ‘reach’. Can you comment on whether ‘casting the wide net’ misses some populations of interest, and how you would know.

The ‘hub and spoke’ model, such as with the Markey Cancer Center and Cancer Coalitions seems to work well. What are its best applications and limitations in terms of reaching small/HTR populations?
Posing Key Questions

For Dr. Sanders Thompson

You gave a wonderful example of matching the right media for your ‘market’ and knowing your audience

▪ Is this all done a priori, iteratively, .............
▪ What are the implications for cost/feasibility and ‘getting it right’ as well as potentially alienation populations/individuals if you don’t tailor correctly and how do you balance that?

For populations for which you can’t “go where they are” and/or mirror the audience – what then?

▪ Virtual v actual “going where they are”
▪ Can something similar be adapted to online communities – technology/social media-savvy embedded individuals?
Posing Key Questions

For All Speakers

• What are we doing about populations we can’t reach? Do we know who and/or where they are?

• Is there a comprehensive compilation of small and hard-to-reach populations, such that we can track/address:
  ▪ Which have been reached and how?
  ▪ Which haven’t?
  ▪ For which do we have evidence – or even information – on how to recruit, retain, and intervene?
Where Do We Go From Here?

- What technologies actually work, and according to what factors (age, race, geography, etc.)

- Tall order to determine effective strategies specific to populations, data needs, AND by phase. What intentionality should we as researchers bring to this? (need-based prioritization, low hanging fruit, piecemeal, etc.)
Additional Questions

- What existing data & resources can we leverage:
  - Web content mining
  - Existing geospatial or governmental resources
  - Online communities

- What data can/should we generate to inform best strategies?

- How can we best leverage/maximize what we learn?