Sharing Clinical Research Data: Benefits and Barriers from a Rare Disease Perspective

Beth A. Kozel, MD, PhD
Department of Pediatrics
Washington University School of Medicine
St. Louis, Missouri

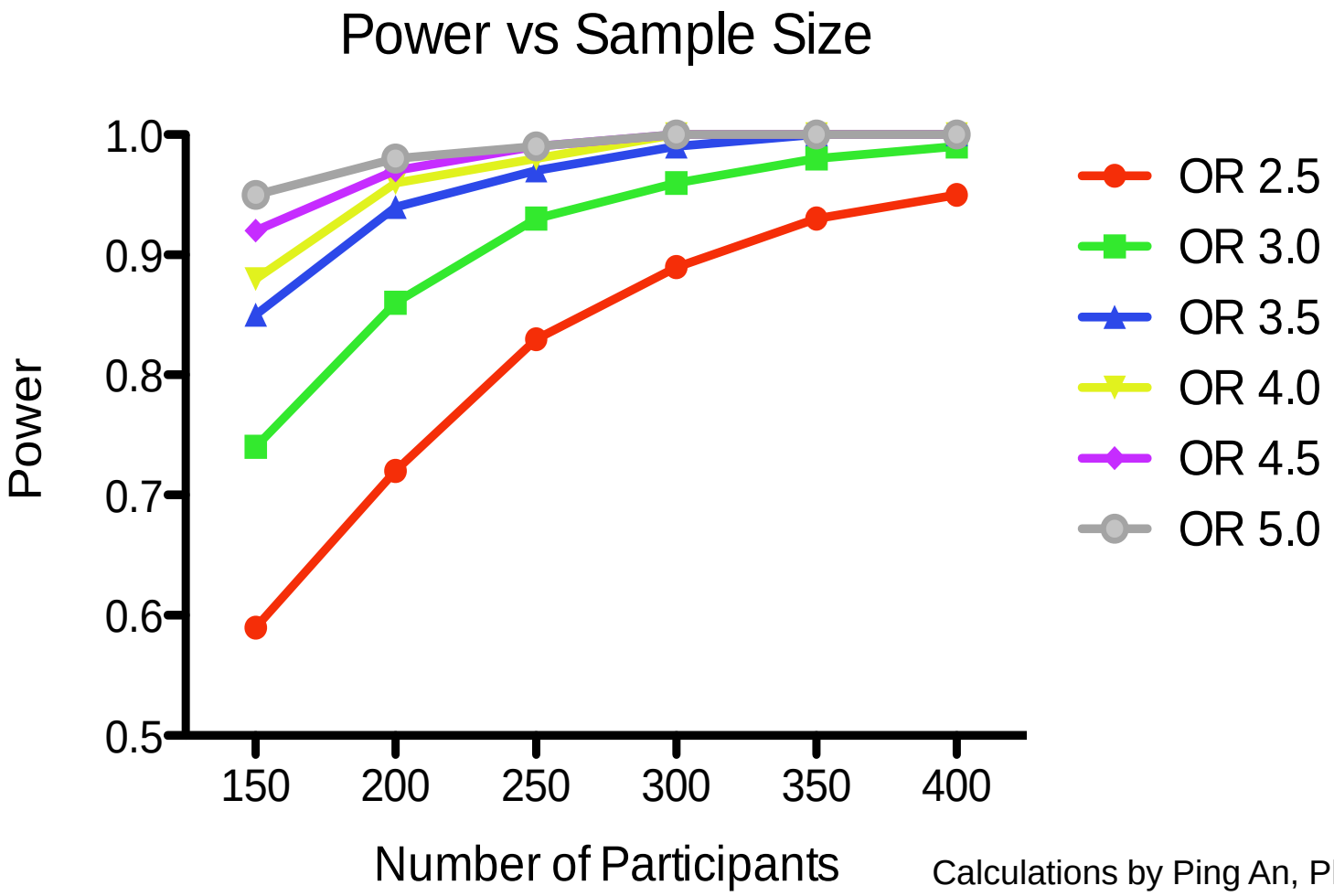
Beth A. Kozel, MD, PhD, has no financial or other conflicts of interest to disclose

Williams syndrome

- Rare condition, 1:8000-12,000 prevalence
- Caused by deletion of 26-28 genes on chromosome 7
- Characteristic features include:
  - Cardiovascular anomalies
  - Characteristic neurocognitive profile
  - Failure to gain weight in infancy
  - Later predisposition for obesity and diabetes
  - Endocrine abnormalities
- Variability of disease severity in WS is likely attributable to differences in the deletion and in genetic background

Photos from http://www.williams-syndrome.org

Detecting polymorphisms responsible for differences in disease severity require large sample sizes



Calculations by Ping An, PhD,
Div. of Statistical Genomics, WUSM

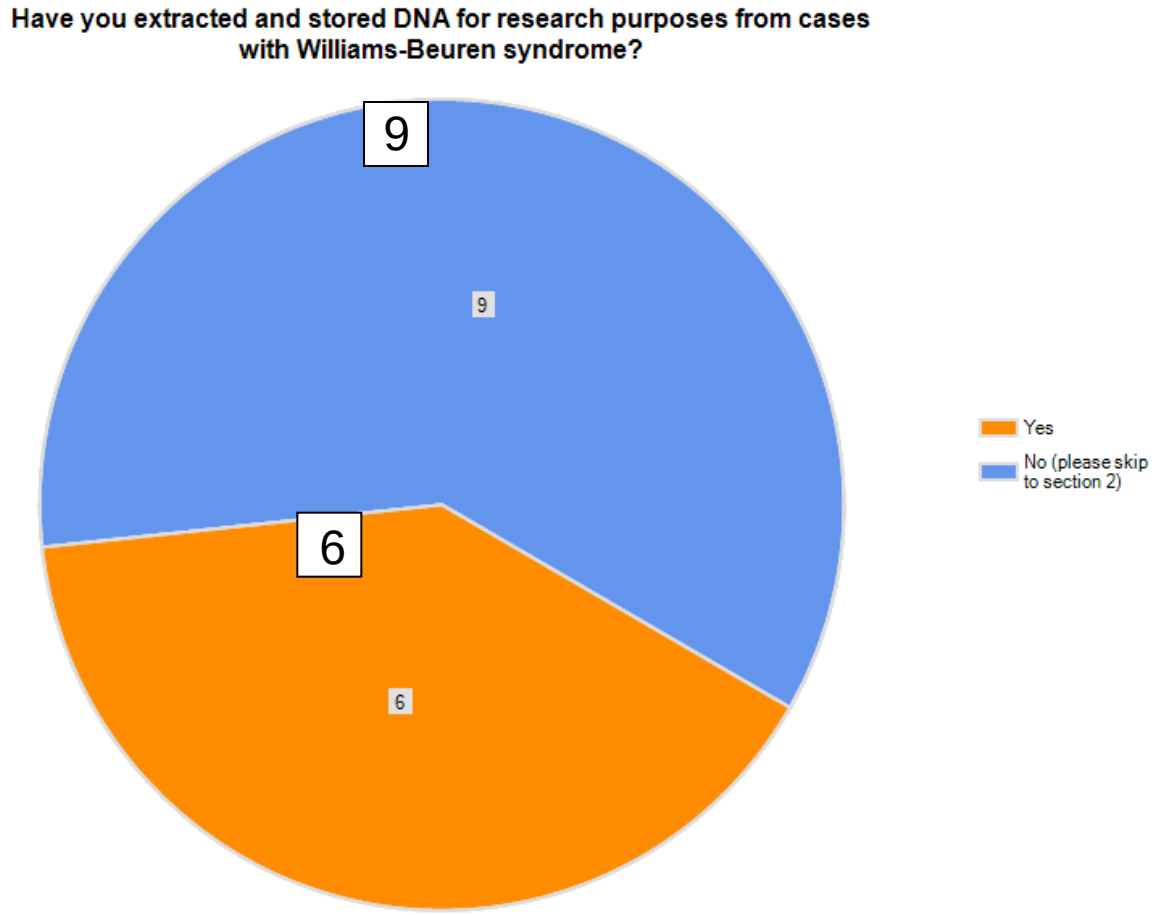
williamssyndrome ASSOCIATION

Dear Colleague September 7, 2010

The Williams Syndrome Association (WSA) serves the Williams syndrome (WS) community in many different ways. We provide support, education, and access to research for individuals with WS and their families. We also strongly support researchers by awarding grants, sponsoring professional and think-tank style meetings, and providing access to our membership.

We believe the WSA can also drive research progress by promoting collaboration among researchers. A long term goal could be the creation of a consortium containing biological samples and clinical data potentially sharable among qualified investigators (akin to that created by the Alzheimer community http://www.nytimes.com/2010/08/13/health/research/13alzheimer.html).

- Survey sent to 30 individuals/ groups known to be active in WS research
- 15 survey responses were received



Barriers to Clinical Data Sharing

- Maintenance of genetic confidentiality and return of results
  - Genetic signatures identifiable in public databases
  - Expectations for return of results
  - Re-consenting of minors
  - IRB and regulatory limitations on science conducted on decades old genetic material

Barriers to Clinical Data Sharing

- Assignment of academic credit for scientific discoveries in rare diseases
- Scientific “clout” may be associated with an individual’s access to rare samples

Barriers to Clinical Data Sharing

- Samples are limiting (cheek swabs)
- Journals not interested in publishing small cohorts of rare diseases
- Single sample may belong to multiple cohorts

What can be done?
Regulatory groups

- IRBs: Permit patients and families to become active partners in judgment re: genetic confidentiality
  - Use social media to engage patients and families in real time re: status of samples
  - Permit longitudinal data acquisition of phenotype over extended periods of time
  - Inform patients and families about plans for safe data sharing

Sharing may yield great rewards but assigning credit has become increasingly complicated

ARTICLE

doi:10.1038/nature11327

Subgroup-specific structural variation across 1,000 medulloblastoma genomes

A list of authors and their affiliations appears at the end of the paper

Medulloblastoma, the most common malignant paediatric brain tumour, is currently treated with nonspecific cytotoxic therapies including surgery, whole-brain radiation, and aggressive chemotherapy. As medulloblastoma exhibits marked intertumoural heterogeneity, with at least four distinct molecular variants, previous attempts to identify targets for therapy have been underpowered because of small samples sizes. Here we report somatic copy number aberrations (SCNAs) in 1,087 unique medulloblastomas. SCNAs are common in medulloblastoma, and are predominantly subgroup-enriched. The most common region of focal copy number gain is a tandem duplication of *SNCAIP*, a gene associated with Parkinson's disease, which is exquisitely restricted to Group 4α. Recurrent translocations of *PVT1*, including *PVT1-MYC* and *PVT1-NDRG1*, that arise through chromothripsis are restricted to Group 3. Numerous targetable SCNAs, including recurrent events targeting TGF-β signalling in Group 3, and NF-κB signalling in Group 4, suggest future avenues for rational, targeted therapy.

What can be done?
Academia

- New mechanisms to assign credit for collaborative works
- Changes to laboratory structure and academic promotion based on “team science” model

What can be done?
Funding organizations

- Consideration for establishing central biobanks for rare diseases
  - Expand support for science of biobanks (e.g., changes associated with sample storage)
  - Expense of well run biobanks are too large for many small family groups to finance.
- When samples are limiting (non-renewable), hosting of down stream data (sequence data, expression data) may be preferable to storing and distributing samples

What can be done?
Family advocacy groups

- Educate members about the pros and cons of data sharing
- Consider advocating for safe data sharing as a group
  - Ask members to look for/ask about sharing statements in consent forms.
  - When data sharing is not allowed, ask why.
- Support central databases with GUID when funding allows.

What can be done?
Journals

- Continue to require primary data from genomic studies to be deposited in protected but accessible sites online.
- Consider mechanisms to connect authors of underpowered research *vs* allowing publication of lower powered studies that can later be re-studied by meta analysis

Summary

- Development of personalized medicine strategies for rare diseases requires large data sets
- Acquisition of large numbers of rare samples will require coordinated efforts of multiple groups
- Changes in practice will likely be needed from all players in the scientific arena