Moral Concerns and the Willingness to Donate to a Research Biobank

Research biobanks are increasing in number and importance, with great potential for advancing knowledge of human health, disease, and treatment. Recruitment of donors is vital to their success and relies largely on blanket consent, in which donors give one-time permission for any future research uses of their coded specimens. This approach to consent has been endorsed recently in proposed changes to federal regulations.²

Previous studies suggest that donors may have moral, religious, and cultural concerns about the use to which their specimens are put, which may affect their willingness to give blanket consent.³,⁴ These earlier studies, however, used convenience samples unrepresentative of the US population.

Methods | The institutional review boards at the University of Michigan and Michigan State University approved this study as exempt. Between June 18, 2014, and June 30, 2014, we used the GfK KnowledgePanel (a probability-based online panel of adults aged 18 years or older, designed to represent the civilian, noninstitutionalized US population) to field a survey examining associations between moral concerns and the willingness to donate to a biobank.

Respondents read an introductory description of a fictional biobank and then used a 6-point scale—from strongly agree to strongly disagree—to indicate their willingness to donate, first using blanket consent and then “even if” their samples might be used in each of 7 potential research scenarios presenting moral concerns. We then gave respondents short descriptions of the benefits and consequences of 5 methods of gaining consent and asked them to indicate which were the acceptable, best, and worst options.

All analyses were weighted to correct for the stratified sampling designs and other sources of survey errors including nonresponse and noncoverage. We used conditional logistic regression to compare willingness to consent with blanket consent vs other scenarios. Analyses were done using Stata version 13.1 (StataCorp); all tests were 2-sided, with a threshold of P = .05.

Results | After excluding 39 surveys with nonresponses to at least half of the substantive survey questions, our final analysis included 1599 participants, resulting in a response rate of 60.2% (1599 of 2654 participants). Respondents were older (51 years vs 45 years for nonrespondents), were more commonly white (82% vs 75%), and had higher levels of education and household income (eTable in the Supplement).

Table 1. Willingness to Give Blanket Consent at Baseline and for 7 Potential Research Scenarios Raising Moral Concerns

<table>
<thead>
<tr>
<th>Scenario Description</th>
<th>Totala</th>
<th>Agreedb</th>
<th>% (95% CI)c</th>
<th>P Valuee</th>
</tr>
</thead>
<tbody>
<tr>
<td>At baseline: &quot;I would donate tissue samples and medical information to the biobank, so that it can use them for any research study that it allows, without further consent from me.&quot;</td>
<td>1593</td>
<td>1122</td>
<td>68.0 (65.5-70.5)</td>
<td></td>
</tr>
<tr>
<td>Under research scenario: &quot;I would donate tissue samples and medical information to the biobank, so that the biobank can use them for any research study that it allows, without further consent from me even if researchers might use donations to...&quot;</td>
<td>1598</td>
<td>790</td>
<td>49.5 (46.9-52.1)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>...develop more safe and effective abortion methods.</td>
<td>1592</td>
<td>1066</td>
<td>64.2 (61.6-66.8)</td>
<td>.007</td>
</tr>
<tr>
<td>...develop kidney stem cells. They would then try to grow these cells in a pig embryo that would grow into an adult pig with human kidneys. The goal would be to grow kidneys or other organs that could be transplanted into people.</td>
<td>1591</td>
<td>912</td>
<td>55.2 (52.6-57.8)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>...develop patents and earn profits for commercial companies. Most new drugs used to treat or prevent disease come from commercial companies.</td>
<td>1591</td>
<td>1151</td>
<td>70.1 (67.6-72.6)</td>
<td>.17</td>
</tr>
<tr>
<td>...develop stem cells that have the donor’s genetic code. These could be kept alive for many years. Scientists might use those stem cells to create many different kinds of tissues and organs for use in medical research.</td>
<td>1590</td>
<td>918</td>
<td>56.6 (53.9-59.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>...create vaccines against new biological weapons. The government might need to develop biological weapons of its own when it does this research.</td>
<td>1591</td>
<td>1042</td>
<td>64.0 (61.5-66.6)</td>
<td>.005</td>
</tr>
<tr>
<td>...understand the evolution of different ethnic groups, and where they come from. What they learn might conflict with some religious or cultural beliefs.</td>
<td>1591</td>
<td>946</td>
<td>58.1 (55.5-60.7)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>...discover genes that make some people more violent. This could lead to ways to reduce violent behavior. But if these genes are found to be more common among some racial and ethnic groups, this might increase prejudice.</td>
<td>1591</td>
<td>1122</td>
<td>68.0 (65.5-70.5)</td>
<td></td>
</tr>
</tbody>
</table>

*a Excluded those who refused to respond to each question.

b Selected 4, 5, or 6 on a 6-point scale (1 = strongly disagree and 6 = strongly agree).

c Percentages accounted for poststratification weights.

d From comparisons between willingness to consent under each scenario vs willingness to first give blanket consent, using conditional logistic regression with survey weights. Each conditional logistic regression model used paired binary willingness responses (under each scenario and under blanket consent) from each participant as the dependent variable, and the P value was from testing for the significance of the parameter estimate of the indicator for the scenario (vs blanket consent).

e Descriptions of scenarios as presented to respondents.
the least preferred option. But blanket consent, the option currently in widespread use, was not far behind. This suggests that an adequate approach for dealing with donors’ moral concerns may lie between these 2 extremes.

Limitations include a response rate of 60%, with respondents and nonrespondents differing on some characteristics that may introduce bias. Because respondents may be more in favor of research, the association between moral concerns and decreased willingness to donate may be a conservative estimate. Also, respondents’ views were based on brief scenarios rather than on detailed understanding of the issues. Deliberative engagement with citizens may deepen understanding of public opinion regarding biobank policy.

Tom Tomlinson, PhD
Raymond De Vries, PhD
Kerry Ryan, MA
Hyungjin Myra Kim, ScD
Nicole Lehpamer, MA
Scott Y. H. Kim, MD, PhD

Author Affiliations: Center for Ethics and Humanities in the Life Sciences, Michigan State University, East Lansing (Tomlinson); Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (De Vries, Ryan); Center for Statistical Consultation and Research, University of Michigan, Ann Arbor (H. M. Kim); Department of Sociology, Michigan State University, East Lansing (Lehpamer); Department of Bioethics, National Institutes of Health, Bethesda, Maryland (S. Y. H. Kim).

Corresponding Author: Tom Tomlinson, PhD, Center for Ethics and Humanities in the Life Sciences, Michigan State University, 965 Fee Rd, East Lansing, MI 48824 (tom.tomlinson@ht.msu.edu).

Author Contributions: Dr Tomlinson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Tomlinson, De Vries, Ryan, H. Kim, S. Kim. Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: All authors. Critical revision of the manuscript for important intellectual content: Tomlinson, De Vries, S. Kim.

Statistical analysis: H. Kim.

Obtained funding: Tomlinson, S. Kim.

Administrative, technical, or material support: De Vries, Ryan, Lehpamer.

Study supervision: Tomlinson, De Vries.
COMMENT & RESPONSE

Heterogeneity in Meta-analysis of FDG-PET Studies to Diagnose Lung Cancer

To the Editor Dr Deppen and colleagues conducted a large meta-analysis that showed the limitations of lung cancer diagnosis using fluodeoxyglucose F 18 combined with positron emission tomography (FDG-PET) in areas with endemic infectious lung disease. Although the sensitivity and specificity of FDG-PET diagnosis was heterogeneous across the included studies, thereby compromising interpretation of the pooled results, the relevance of presenting an F statistic to underscore and interpret the extent of the heterogeneity should be questioned. The large reported F values may be an artifact of their chosen measure (proportions) and may not solely reflect important clinical or contextual sources of heterogeneity. The F statistic is perhaps the most popular method to assess the extent of statistical heterogeneity within meta-analyses, mostly due to its uncritical promotion within the Cochrane Collaboration. We also modified the traditional forest plot (Figure 2 in article), accordingly; F statistics were removed and confidence intervals for combined and overall effects were derived from the random-effects logistic model (instead of the traditional pooled model). To highlight study heterogeneity, we displayed prediction intervals (Figure 3) from our model.

In Reply Dr Mills and colleagues wish to de-emphasize F statistics. We agree that F statistics can be problematic and that the naive pooled model (ie, the model assuming no study-to-study heterogeneity) is often inappropriate for meta-analyses. This is why, in our study, F statistics were only reported for completeness and never factored into our analysis. First, Mills and colleagues assume that F statistics played a prominent role in our analysis. They did not. They were reported in accordance with PRISMA guidelines but served no other function. We quantified the extent of study-to-study heterogeneity with a component of variance from the best-fit random-effects logistic regression model.

When the F statistic was first reported, Higgins et al suggested cutoffs to describe heterogeneity qualitatively (eg, >75% is high). Their benchmark was based on comparative measures (eg, odds ratios), and these behave differently than proportions. More specifically, given equal sample size, the variance of a proportion is smaller than the variance of an odds ratio; therefore, the I2 statistic will tend to be larger for proportions.

Edward J. Mills, PhD, MSc, MST
Jeroen P. Jansen, PhD
Steve Kanters, MSc

Conflict of Interest Disclosures: The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.
Funding/Support: The work was supported by grant 1 R01 HG007172-01A1 from the National Human Genome Research Institute.
Role of the Funders/Sponsors: The National Human Genome Research Institute had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.
Disclaimer: The ideas and opinions expressed by Dr S. Kim in this article are his own; they do not represent any position or policy of the National Institutes of Health, the Department of Health and Human Services, or the US government.

5. Wendler D. One-time general consent for research on biological samples. BMJ. 2006;332(7540):544-547.

In Reply Dr Mills and colleagues wish to deemphasize F statistics. We agree that F statistics can be problematic and that the naive pooled model (ie, the model assuming no study-to-study heterogeneity) is often inappropriate for meta-analyses. This is why, in our study, F statistics were only reported for completeness and never factored into our analysis.

First, Mills and colleagues assume that F statistics played a prominent role in our analysis. They did not. They were reported in accordance with PRISMA guidelines but served no other function. We quantified the extent of study-to-study heterogeneity with a component of variance from the best-fit random-effects logistic regression model.

We also modified the traditional forest plot (Figure 2 in article), accordingly; F statistics were removed and confidence intervals for combined and overall effects were derived from the random-effects logistic model (instead of the traditional pooled model). To highlight study heterogeneity, we displayed prediction intervals (Figure 2) from our model.

In this context, the difference between prediction and confidence interval length is due to study-to-study heterogeneity. The prediction interval for specificity was wider than that for sensitivity. Even though this is suggested by Mills and colleagues’ eyeball test, the formal testing approach we used, based on model selection, is likely to perform better.

Second, in our case, an overreliance on F statistics would have led to the same conclusions. The F statistics for sensitivity and specificity were both large, indicating the pooled model was not supported by the data. Thus, the pooled model