Consent to Take Part in a Research Study

Q1.1.

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Research Study Consent

Title of research study: Data science workshops for biomedical and health professionals: Persona identification and workshop assessment[CD1]

Principal Investigator: Anne M Brown, PhD
Other study contact(s): Daniel Chen
Department: Research and Informatics, University Libraries
Telephone Number: 540-231-9231
Email Addresses: ambrown7@vt.edu, chend@vt.eu

Key Information: The following is a short summary of this study to help you decide whether or not to be a part of this study. More detailed information is listed later on in this form. This confidential survey is intended to identify key characteristics of learners attending a data science workshop geared towards medical and biomedical practitioners. Your responses are anonymous and will become part of summary data included in a report that is compiled by Daniel Chen from the Genetics, Bioinformatics, and Computational Biology (GBCB) PhD program at Virginia Tech.

Detailed Information: The following is more detailed information about this study in addition to the information listed above.

Who can I talk to?

If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at Anne Brown (ambrown7@vt.edu) or Daniel Chen (chend@vt.edu).

This research has been reviewed by the Virginia Tech Institutional Review Board (IRB). You may communicate with them at 540-231-3732 or irb@vt.edu if:
You have questions about your rights as a research subject Your questions, concerns, or complaints are not being answered by the research team You cannot reach the research team You want to talk to someone besides the research team to provide feedback about this research.

How many people will be studied?
We plan to include about 100-150 people in this research study.

What happens if I say yes, I want to be in this research?
Once you consent to participate in this study, you will follow a next arrow to the study survey and begin by answering some population-specific questions. One of these questions will include creating a unique identifier which will be used for future surveys in this study. This unique identifier will also be used to remove your data if you choose to not remain in the study. From there you will complete a pre-workshop questionnaire about your experiences with programming, data processing, project management, and what you plan to get out of the workshop. The survey should take approximately 10-15 minutes to complete. After the workshop you will be invited to participate in a post-workshop survey. You may opt to complete these surveys to provide feedback to the researchers about the materials without agreeing to participate in the research study.

Your responses in this survey will be used to assess the quality of a workshop and its learning materials. The researchers will then take feedback from the

workshop to improve materials for each type of learner. You have the option to provide feedback without having your data used in the final research analysis report.

This information will be used to improve the workshop materials and give instructors a sense of who is attending the workshop and what needs learners need before, during, and after the workshop ends. It will also help determine what kind of workshop is better suited for each kind of learner.

This consent form can be taken anywhere and survey can be taken anywhere. Workshops are planned for Fall 2020-Spring 20201. Your responses from this survey will be matched up with subsequent workshop assessments, if you do attend workshops, using your de-identified ID. You may choose to attend and register for the workshop even if you do not wish to partake in this study. You may also take the surveys to provide feedback without participating in the research study.

This study is designed that individuals will take a student self-assessment prior to workshop creation. As workshops are delivered, pre and post assessment will be performed on each workshop. A subsequent long-term survey may be issued. You are invited to participate in any of the surveys as well as the current student self-assessment.

Your de-identified (anonymous) survey responses will be shared on an open science platform such as the Open Science Framework (https://osf.io/), GitHub (https://github.com/), Zenodo (https://zenodo.org/), and/or VTechData (https://data.lib.vt.edu/).

[CD1]Anne: Do I need to change this title? I think the original title still works since all the surveys are for the same study
[CD2]Removing mention of personas since we're no longer going to create new

personas but still assessing learners of the workshop

Q1.2.

What happens if I say yes, but I change my mind later?

You can leave the research at any time, for any reason, and it will not be held against you.

You will still be allowed to participate in the workshop without completing this survey.
You may also take the survey to provide feedback and also opt-out of the research study.

If you decide to leave the research, you can discontinue filling out the survey by closing your web browser. Incomplete responses will not be used in the data analysis. If you already filled out the survey and wish to withdraw your participation, please let one of the study investigators know You can decide to pull all your information from analysis, or only withdraw from any follow-up surveys. If you decide to leave the research, no consequences will occur.

Is there any way being in this study could be bad for me? (Detailed Risks)

During the process of completing the survey, you will be asked questions about programming experience and your thoughts and attitudes surrounding the subject of statistics and data management. If there are any questions you would rather not answer or that you do not feel comfortable answering, you can move on to the next question.

There is minimal risk the by being a part of this study you could experience physical, psychological, privacy, legal, social, economic, or emotional distress given the subject of the survey.

This study is not meant to gather information about specific individuals, and the information you provide will be combined with that of other survey participants to gather information.

What happens to the information collected for the research?
We will make every effort to limit the use and disclosure of your personal information, including research study and medical records, only to people who have a need to review this information. We cannot promise complete confidentiality. Organizations that may inspect and copy your information include the IRB, Human Research Protection Program, and other authorized representatives of Virginia Tech.

Your de-identified (anonymous) survey responses will be shared on an open science platform such as the Open Science Framework (https://osf.io/), GitHub (https://github.com/), Zenodo (https://zenodo.org/), and/or VTechData (https://data.lib.vt.edu/).

Your data could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

The results of this research study may be presented in summary form at conferences, in presentations, reports to the sponsor, academic papers, and as part of a thesis/dissertation.

Can I be removed from the research without my OK?
The person in charge of the research study or the sponsor can remove you from the research study without your approval. Possible reasons for removal include incomplete responses.

What else do I need to know?
Any expenses accrued for seeking or receiving medical or mental health

treatment will be your responsibility and not that of the research project, research team, or Virginia Tech.

Are you at least 18 years of age?
Yes. I am 18 years of age or older.
No. I am not at least 18 years of age.

Do you agree to participate in the research study?

Yes. I have read the consent form and this response will serve as my consent to participate in the research study
No. I do not want to participate in the research study.

Would you like to provide pre-workshop feedback? Your responses would not be used for the research study.
Yes. I would like to provide feedback about the workshop and its learning materials
No. I do not want to provide feedback.

Q1.3. Are you at least 18 years of age?

- ◯ Yes. I am 18 years of age or older.
- ◯ No. I am not at least 18 years of age.

Q1.4.
Do you agree to participate in the research study?

- ◯ Yes. I have read the consent form and this response will serve as my consent to participate in the research study.
- ◯ No. I do not want to participate in the research study.

Demographics

Q2.1.
Hello:

Thank you for participating in the "Data Science for Medical and Biomedical Practitioners" workshop.

This is the learner's **long-term** workshop and/or self-learning survey.

This survey helps determine workshop material efficacy for individuals to learn data science in the biomedical and health fields.

These survey results will allow us to create materials specific to needs of the group.

Q2.2.
Please create a unique identifier. This unique identifier will be used to like your survey responses but keep your personal information anonymous.

To create an identifier type in:

- Number of siblings (as numeric) +
- First two letters of the city you were born in (lowercase) +
- First three letters of your current street (lowercase).

E.g., (Sherlock Homes has **1** brother, was born in **Po**rsmouth, and lives on **Bac**ker Street - **1pobac**)

Q2.3. Please select the first date of your workshop

- [ ] Tuesday, October 20, 2020
- [ ] Wednesday, December 9, 2020
- [ ] Tuesday, February 2, 2021
- [ ] Monday, May 17, 2021
- [ ] Tuesday, June 29, 2021 - In Person
- [ ] Tuesday, June 29, 2021 - Virtual
- [ ] I went through the online materials on my own
- [ ] Other

Q2.4. What is your current occupation/career stage (select all that apply).

- [ ] DO/MD
- [ ] DVM
- [ ] RN/PA
- [ ] PhD
- [ ] Academic
- [ ] Analyst
- [ ] Student (Masters e.g., MPH)
- [ ] Student (MD/DO)
- [ ] Student (Nurse, PA)
- [ ] Student (Graduate)
- [ ] Student (Undergraduate)
- [ ] iTHRIV Scholar
- [ ] Other, please describe

Behaviors and Confidence

Q3.1. Which of the following behaviors have you adopted as a result of completing the workshop / going through the materials.

- [ ] Improving data management and project organization
- [ ] Developing a data management and analysis plan
- [ ] Transforming step-by-step workflows into scripts
- [ ] Using programming languages like R or Python to automate repetitive tasks
- [ ] Reusing code
- [ ] Sharing code or data publicly
- [ ] None
- [ ] Other

Q3.2. Before the workshop, how often did you use programming languages?

- ○ I had not been using tools like these
- ○ Less than once a per half-year
- ○ Several times per half-year
- ○ Monthly
- ○ Weekly
- ○ Daily

Q3.3. Since taking the workshop, how often did you use programming languages?

- ○ I had not been using tools like these
- ○ Less than once a in the last 6 months
- ○ Several times in the last 6 months
- ○ Monthly
- ○ Weekly
- ○ Daily

Q3.4. How would you rate your change in confidence in the tools that were covered during your workshop compared to before the workshop?

- ○ I'm more confident now
- ○ I'm equally confident now
- ○ I'm less confident now

Workshop Framing and Motivation

Q4.1. Please rate your level of agreement with the following statements

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Agree	Strongly Agree
I believe having access to the original, raw data is important to be able to repeat an analysis.	○	○	○	○	○	○	○
I can write a small program, script, or macro to address a problem in my own work.	○	○	○	○	○	○	○
I know how to search for answers to my technical questions online.	○	○	○	○	○	○	○
While working on a programming project, if I got stuck, I can find ways of overcoming the problem.	○	○	○	○	○	○	○
I am confident in my ability to make use of	○	○	○	○	○	○	○

programming language (like R or Python) can make my analyses easier to reproduce.	○	○	○	○	○	○	○
Using a programming language (like R or Python) can make me more efficient at working with data.	○	○	○	○	○	○	○

Q4.2. Please rate your level of agreement with your ability to do the following tasks

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Agree	Strongly Agree
Name the features of a tidy/clean dataset	○	○	○	○	○	○	○
Transform data for analysis	○	○	○	○	○	○	○
Identify when spreadsheets are useful	○	○	○	○	○	○	○
Assess when a task should not be done in a spreadsheet software	○	○	○	○	○	○	○
Break down data processing into smaller individual (and more manageable) steps	○	○	○	○	○	○	○
Construct a plot and table for exploratory data analysis	○	○	○	○	○	○	○
Build a data processing pipeline that can be used in multiple programs	○	○	○	○	○	○	○

statistical analysis of the data

Summative assessment

Q5.1.

Cytomegalovirus (CMV) is a common virus that normally does not cause any problems in the body. However, it can be of concern for those who are pregnant or immunocompromised.

Suppose you have the following Cytomegalovirus dataset [1] of CMV reactivation among patients after Allogenetic Hematopoietic Stem Cell Transplant (HSCT) in an excel sheet (first 10 rows shown below):

	A	B	C	D	E	F
1	**ID**	**age**	**prior.radiation**	**aKIRs**	**donor_negative**	**donor_positive**
2	1	61	0	1	recipient_positive	
3	2	62	1	5	recipient_negative	
4	3	63	0	3		recipient_positive
5	4	33	1	2	recipient_positive	
6	5	54	0	6		recipient_positive
7	6	55	0	2		recipient_positive
8	7	67	0	1		recipient_positive
9	8	51	0	2		recipient_positive
10	9	44	1	2		recipient_positive
11	10	59	0	4	recipient_negative	

It contains a patient's:

- ID age
- priror.radiation: whether or not patient had prior radiation treatment (0 = no, 1 = yes)
- aKIRs: Number of donor activating killer immunoglobulin-link receptors
- donor_negative: the recipient's CMV status when the donor was CMV negative
- donor_positive: the recipient's CMV status when the donor was CMV positive

It is believed that the donor activating KIR genotype is a contributing factor for CMV reactivation after myeloablastive allogenetic HSCT. You want to do

Assuming this is the version of the data you need for the tidying, plotting, and modeling:

	A	B	C	D	E	F
1	ID	age	prior.radiation	aKIRs	donor_status	recipient_status
2	1	61	0	1	donor_negative	recipient_positive
3	2	62	1	5	donor_negative	recipient_negative
4	3	63	0	3	donor_positive	recipient_positive
5	4	33	1	2	donor_negative	recipient_positive
6	5	54	0	6	donor_positive	recipient_positive
7	6	55	0	2	donor_positive	recipient_positive
8	7	67	0	1	donor_positive	recipient_positive
9	8	51	0	2	donor_positive	recipient_positive
10	9	44	1	2	donor_positive	recipient_positive
11	10	59	0	4	donor_negative	recipient_negative

How would you rate your ability to accomplish the following tasks:

[1]: Sobecks et al. “Cytomegalovirus Reactivation After Matched Sibling DonorReduced-Intensity Conditioning Allogeneic HematopoieticStem Cell Transplant Correlates With Donor KillerImmunoglobulin-like Receptor Genotype”. Exp Clin Transplant2011; 1: 7-13.

	I wouldn’t know where to start	I could struggle through, but not confident I could do it	I could struggle through by trial and error with a lot of web searches	I could do it quickly with little or no use of external help
Load the excel sheet into R	○	○	○	○
Filter the data for individuals over the age of 65 (in R)	○	○	○	○
Save filtered dataset (in R) as an Excel file to send to a colleague	○	○	○	○

Fit a model (e.g., logistic regression) to see which variables are associated with patient CMV reactivation (in R)	○	○	○	○

Impact

Q6.1.

The statements below reflect ways in which completing the workshop may have impacted you.

Please indicate your level of agreement with these statements.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I have used skills I learned at the workshop to advance my career.	○	○	○	○	○
I have been motivated to seek more knowledge about the tools I learned at the workshop.	○	○	○	○	○
I have made my analysis more reproducible as a result of completing the workshop.	○	○	○	○	○
I have improved my coding practices as a result of completing the Workshop	○	○	○	○	○
My research productivity has improved as a result of completing the workshop	○	○	○	○	○
I have gained confidence in working with data as a result of completing the workshop.	○	○	○	○	○

Q6.2. Did you go back to the online materials after the workshop?

- [ ] I went back to the code I wrote for reference
- [ ] I went back to the code the instructor posted for reference
- [ ] I went back to the video recording for the workshop
- [ ] I went back to the online written materials for reference
- [ ] I did not go back to any of the workshop materials
- [ ] Other

Q6.3. Why did you not go back to a particular workshop resource?

Q6.4. Please tell us the most important way you were impacted as a result of the workshop.

Q6.5.
Please provide any outcomes as a result of attending this workshop.

Q6.6. If you would like to make additional comments about the workshop experience, or ways you've used the tools you learned in the workshop please comment below.

Block 6

Powered by Qualtrics