# How to Talk about Your Research to Aliens Non-Statisticians Regina Nuzzo

Senior Advisor for Statistics Communication & Media Innovation

**American Statistical Association** 

**NISS Panel on Presenting Research** 

November 30, 2020

NISS © Nuzzo 2020



# Why bother trying to talk with aliens about your work?

# Because aliens often ...

• are the creatures signing the check

whave a platform to do things that affect the world

we are smart and curious but just don't know very much

4

To be an ideal communicator, overestimate the intelligence of your audience, and underestimate their knowledge.

– My adaption of a 1923 quote from American magazine editor Glenn Frank

"The ideal magazine article should be written as if the men and women who were to read it had just dropped from the planet Mars."

– more from Glenn Frank in 1923

6

# Try this easy template

A. What is the **SETTING**?

- B. What is the **PROBLEM**?
- C. SO WHAT?
- D. What are the **SOLUTIONS** I'm proposing?
- E. What are the **BENEFITS** and/or **NEXT STEPS**?

7

# And remember these few guidelines

- 1. Give even more context than you think you need
- 2. Replace specialized terms with simpler terms
- 3. Replace abstractions with concrete nouns, verbs, and examples

# Example: Stats abstract & my version for journalists

NISS © Nuzzo 2020

Intravenous busulfan is a standard component of the preparative regimen in allogeneic stem cell transplantation (allosct) for acute leukemia. Systemic busulfan exposure, characterized by the area under the plasma concentration curve, AUC, is strongly associated with clinical outcome. A high AUC is associated with severe toxicities, while a low AUC carries risks of disease recurrence and graft failure. An optimal AUC interval is determined for each patient by giving a preclinical dose. To determine if the optimal AUC interval varies with individual patient characteristics, we developed a method for determining covariate-specific, personalized AUC intervals. We used a Bayesian nonparametric survival regression model based on a dependent Dirichlet process and Gaussian process prior (DDP-GP) to analyze data from 151 allosct patients. The fitted model identified optimal AUC intervals that varied with age and whether the patient was in complete remission at transplant. Simulations showed that the DDP-GP model's performance compares favorably with several robust alternative models. An R package, DDPGPSurv, for general implementation of the DDP-GP survival regression model is provided.

# SETTING

Intravenous busulfan is a standard component of the preparative regimen in allogeneic stem cell transplantation (allosct) for acute leukemia.

# SETTING

Treating blood cancers like leukemia sometimes calls for the aggressive approach of transplanting stem cells from a healthy donor into a patient's bone marrow. To help prepare the patient's body for the transplant, a drug called busulfan is injected directly into the veins.

# PROBLEM

Systemic busulfan exposure, characterized by the area under the plasma concentration curve, AUC, is strongly associated with clinical outcome.

# PROBLEM

It's tricky to get its exact dosage correct, however

14

# SO WHAT

A high AUC is associated with severe toxicities, while a low AUC carries risks of disease recurrence and graft failure.

# SO WHAT

Too much of the drug can lead to toxicity or even death, while too little can make it easier for the cancer to return.

# SOLUTION

To determine if the optimal AUC interval varies with individual patient characteristics, we developed a method for determining covariate-specific, personalized AUC intervals. The fitted model identified optimal AUC intervals that varied with age and whether the patient was in complete remission at transplant.

# SOLUTION

In this presentation, the researchers will describe the new "precision medicine" statistical model that they created to determine the right dosage for any patient, which results in a method that be easily used by any transplant doctor.

# BENEFITS

Simulations showed that the DDP-GP model's performance compares favorably with several robust alternative models.

# BENEFITS

By switching from the current "one-size-fits-all" strategy to the new method, the researchers calculate that doctors can extend many patients' lives dramatically — by an average of 10 to 14 months, for example, for 40- to 60-year-olds in complete remission, which is an improvement of up to 290%.

- 1. **SETTING** Treating blood cancers like leukemia sometimes calls for the aggressive approach of transplanting stem cells from a healthy donor into a patient's bone marrow. To help prepare the patient's body for the transplant, a drug called busulfan is injected directly into the veins.
- 2. **PROBLEM** It's tricky to get its exact dosage correct, however.
- 3. SO WHAT Too much of the drug can lead to toxicity or even death, while too little can make it easier for the cancer to return.
- 4. **SOLUTION** In this presentation, the researchers will describe the new "precision medicine" statistical model that they created to determine the right dosage for any patient, resulting in a method that be easily used by any transplant doctor.
- 5. **BENEFITS** By switching from the current "one-size-fits-all" strategy to the new method, the researchers calculate that doctors can extend many patients' lives dramatically – by an average of 10 to 14 months, for example, for 40- to 60-year-olds in complete remission, which is an improvement of up to 290%.



# Benefits Solution

So What

NISS © Nuzzo 2020

# Problem



# PROBLEM

# And a game to get you into the mood for talking to aliens

Write using only the ten hundred most common words in the English language (using Randall Munroe's Simple Writer xkcd.com/simplewriter)



# SIMPLE WRITER

WRITE LIKE UP GOER FIVE AND THING EXPLAINER

# PUT WORDS HERE

A p-value is the probability of obtaining test results at least as extreme as the results actually observed, under the assumption that the null hypothesis is correct



# SIMPLE WRITER WRITE LIKE UP GOER FIVE AND THING EXPLAINER

# PUT WORDS HERE

A p-value is a number that tells you how surprising another group of numbers would be if everything happened just by chance.

NISS © Nuzzo 2020

# Thank you! Regina Nuzzo regina@amstat.org

# Bonus sides

# for people who

managed to

read all the way

down this far

NISS © Nuzzo 2020



28



# SIMPLE WRITER

WRITE LIKE UP GOER FIVE AND THING EXPLAINER

# PUT WORDS HERE

To help people whose blood is sick, doctors sometimes need to take special tiny water-bags from a person who's not sick and put them straight into the sick person's bones. Doctors help prepare the sick person's body by putting a special drink into the sick person's blood. Doctors need to be careful with the special blood drink, though. Too much of it can kill the sick person, but too little of it can make it easier for the person's blood to get sick again later. Today I will tell you about a new idea that my friends and I had that will help doctors figure out exactly how much of the special blood drink to give the sick person. Our idea will make it easy for any doctor to help a person with sick blood, even if the doctor doesn't know much about numbers. My friends and I were able to figure out that if doctors use our new idea, they can help people with very sick blood live much longer than they would have without our idea -- some people will live between 10 and 14 months longer than normal, which is a very nice thing for these sick people and their families!

NISS © Nuzzo 2020

# Another stats abstract

We propose a generative model and an inference scheme for epidemic processes on dynamic, adaptive contact networks. Network evolution is formulated as a link-Markovian process, which is then coupled to an individual-level stochastic SIR model, in order to describe the interplay between epidemic dynamics on a network and network link changes. A Markov chain Monte Carlo framework is developed for likelihood-based inference from partial epidemic observations, with a novel data augmentation algorithm specifically designed to deal with missing individual recovery times under the dynamic network setting. Through a series of simulation experiments, we demonstrate the validity and flexibility of the model as well as the efficacy and efficiency of the data augmentation inference scheme. The model is also applied to a recent real-world dataset on influenza-like-illness transmission with highresolution social contact tracking records.

# For a general audience

- 1. Traditional infectious disease models assume that people mix randomly in a population, and that everyone is equally likely to come into contact everyone else.
- 2. In reality, however, we each have our own network of contacts that we're more likely to mix with,
- 3. which is why contact tracing is so important during pandemics.
- 4. Fan Bu at Duke University and colleagues set out to develop a new method combining both approaches and also improving on past models. Their method can account for how our networks evolve during an epidemic as our behavior changes, and how this in turn affects disease spread. The model can also handle real-world situations with only partial data and where uncertainty is important.
- 5. In this presentation, Bu will use real data from a 2013 flu transmission to show how the new method can incorporate high-tech contact tracing data to improve modeling and forecasting.