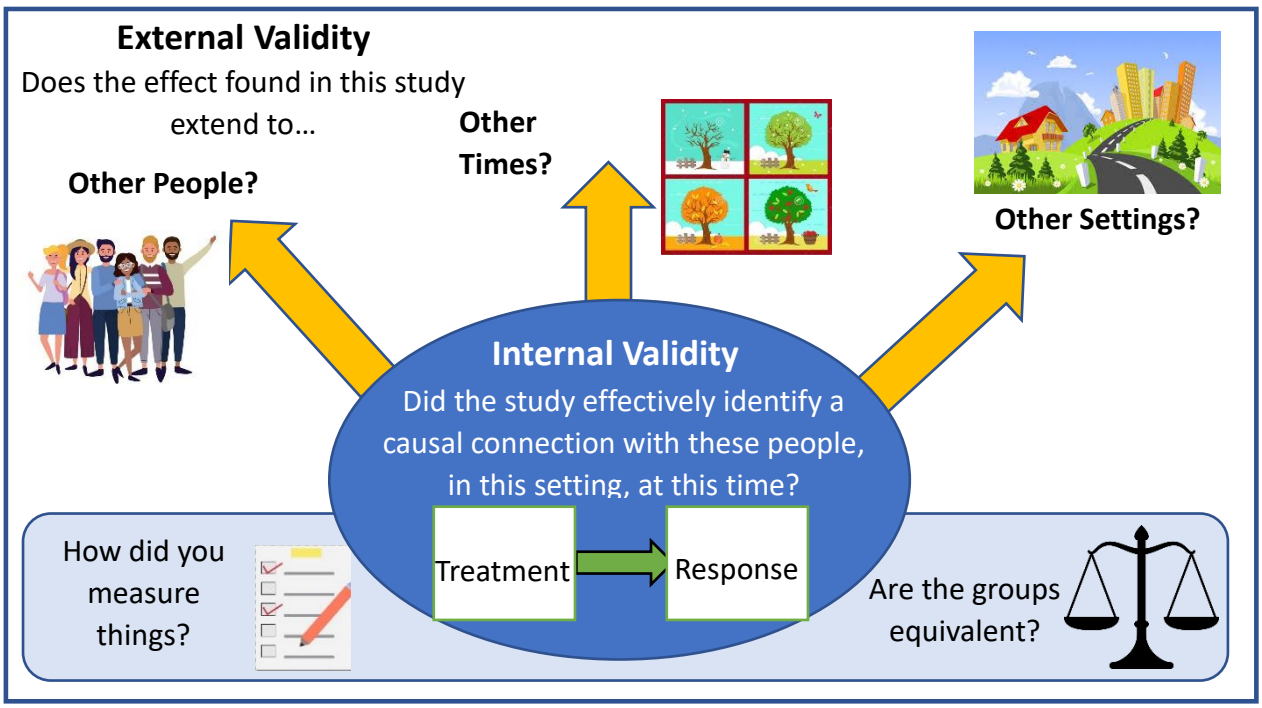
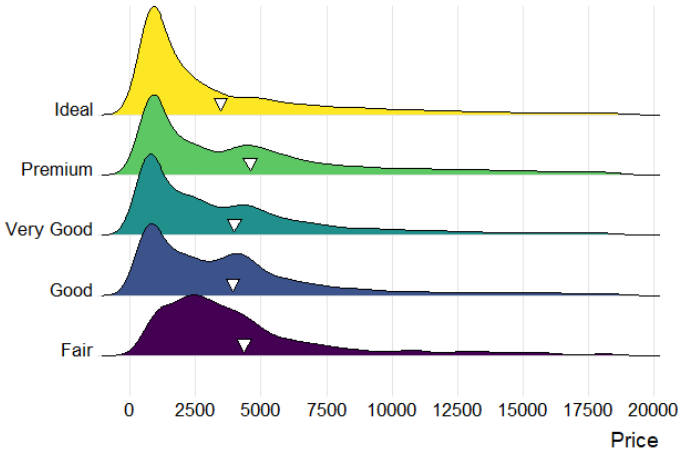
Closing Thoughts

Looking at Statistics with a Critical Eye

- Investigating the Investigation
  - "There are lies, damned lies, and statistics!" Many of the most egregious deceptions are made *using* data.
  - The idea of backing up claims with statistics and studies is (rightly) associated with being more reliable and trustworthy, but the mere presence of data does not guarantee sound claims.
  - An important part of evaluating the data-based claims we see in the news, and even in scholarly journals, is to look at studies holistically.
- **Design Matters**
  - **Experiments/Clinical Trials** are often better for determining causality
    - They are usually more intensive to coordinate and costly to carry out. They may also not always be possible, appropriate, or ethical in every situation.
    - But when possible, experiments provide controlled intervention and *typically* provide a stronger case for a causal connection.
  - **Observational studies** can indicate associations, but make causal inference more difficult.
    - Observational data can often be collected quite easily and at mass scale in many situations.
    - But what is gained in quantity and ease might be lost in quality without the same controlled intervention.
    - If wanting to make more causal inferences, collect data from confounders and stratify/control for confounders appropriately in analysis.
  - And be sure to judge the **Generalizability** of the claims by looking at the data source.

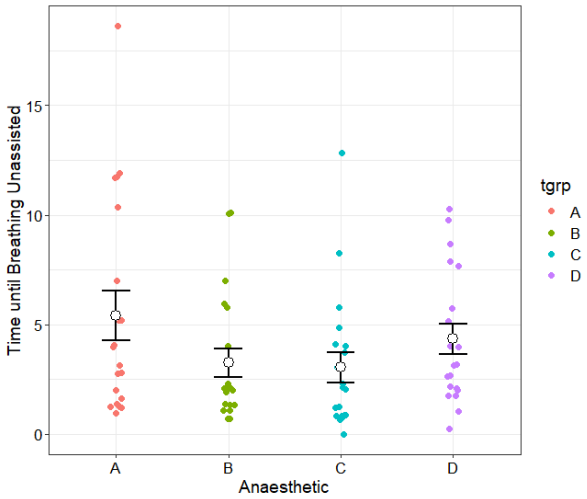


- **Look at your Data**
  - If doing analysis of your own, explore your data to see what's going on.
  - Create summary measures and compare groups. Visualize variables and variable relationships.
  - Where is there variation in the data? What variables explain that variation? What is surprising or unusual?
  - Visualization is a way to investigate the data to find explanations.
  - If analyzing others' data presentations, consider if they have shown you a sufficient amount of the data to know the full story—what do you want to see?



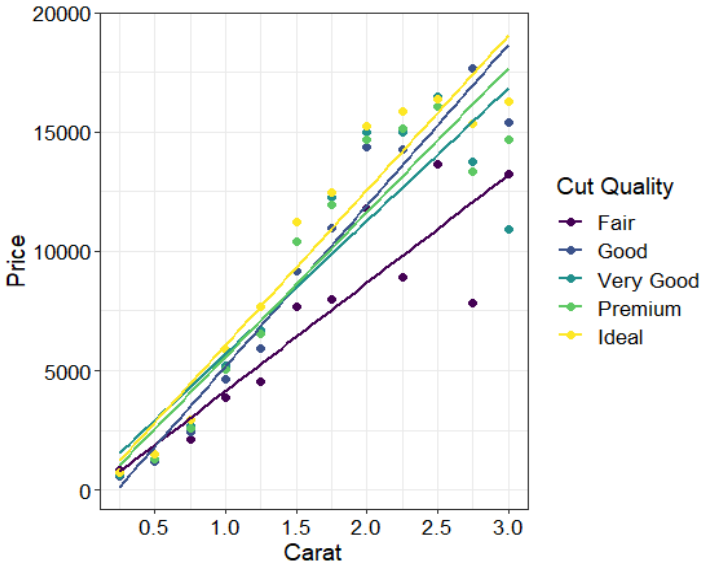
- **Test Claims**
  - *Is there an association?*
    - While not required for all analysis, it's a good idea to determine if any differences or relationship you see can or cannot be explained as sampling variability.
    - P-values provide more objective credibility by making a probabilistic argument for your claims.

  - *How much association?*
    - In addition to determining if there is any association, try to estimate the size of that association.
    - This can be accomplished with confidence intervals of the appropriate parameter (estimate the mean difference, or estimate the slope), as well as with effect size measures like Cohen's d or r^2



  - *Assess the appropriateness of methods*
    - Different statistical methods are appropriate for different data contexts and questions.
    - Most methods may also not produce reliable results if certain assumptions are not met. Address how well the data fit the assumptions made, and discuss limitations and possible reliability issues due to assumption violations.

- **Model Relationships**
  - If you have a numeric predictor, or multiple predictor variables, consider building a model to predict and better understand your response variable of interest.
  - Models can be advantageous for predicting new observations.
  - Models can also be helpful as a way to better understand the specific effect of one predictor by controlling for other related predictors.



- **Make Meaning**
  - This is the most important part—statistical work must also seek to understand why things relate the way they do.
    - Why does this variable explain that variable?
    - Why does it even matter?
  - Coming to conclusions also relies on a holistic view of all information shared in the article (including references to related research).

    - How do the statistical results support the research question(s) posed?
    - What design and validity weaknesses should we note? What information was left out? What further research or replication may be needed?
    - Contextually, what change might these results prompt?

Investigation: These questions come from an article recently published in *The Lancet*: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)32466-1/fulltext

Background: "Older adults (aged ≥70 years) are at increased risk of severe disease and death if they develop COVID-19 and are therefore a priority for immunization should an efficacious vaccine be developed. Immunogenicity of vaccines is often worse in older adults as a result of immunosenescence. We have reported the immunogenicity of a novel chimpanzee adenovirus-vectored vaccine, ChAdOx1 nCoV-19, in young adults, and now describe the safety and immunogenicity of this vaccine in a wider range of participants, including adults aged 70 years and older."

Question: What is this study principally investigating?

This article reports on the phase 2 component of a single-blind, randomized control trial in testing the efficacy of a possible vaccine for COVID-19.

Question: What does it mean that the study was "single blind"?

Healthy adults aged 18 years and older were enrolled at two UK clinical research facilities, in an age-escalation manner, into 18–55 years, 56–69 years, and 70 years and older immunogenicity subgroups. Participants were eligible if they did not have severe or uncontrolled medical comorbidities or a high frailty score (if aged ≥65 years). Participants 70 years and older were recruited from the NIHR Southampton Clinical Research Facility, University Hospital Southampton NHS Foundation Trust. All other participants were recruited at the Oxford Vaccine Centre, Centre for Clinical Vaccinology and Tropical Medicine, University of Oxford. Among the analysed population, 280 (50%) of 552 participants were female. 524 (95%) of 552 participants identified as white, and 540 (98%) were non-smokers. A large proportion of health-care workers who were predominantly female were enrolled in the 18–55 years and 56–69 years age groups.

Question: Comment on the external validity of this study, specifically focusing on participant selection.

Participants were first recruited to a low-dose cohort, and within each age group (18-55, 56-69, and 70+ years old), participants were randomly assigned to receive either intramuscular ChAdOx1 nCoV-19 ($2{\cdot}2\times10^{10}$ virus particles) or a control vaccine (MenACWY), using block randomization and stratified by age and study site.

Question: What type of design was used in this study?

A. Randomized Controlled Experiment
B. Case-Control Study
C. Cross-over Design
D. Cohort Study

More participants were then recruited to be part of a standard-dose cohort ($3{\cdot}5–6{\cdot}5\times10^{10}$ virus particles of ChAdOx1 nCoV-19 or the control vaccine, MenACWY), and the same randomization procedures were followed…In addition to either receiving a low dose or standard dose, participants either received their vaccine as either a single-dose or two-dose regimen (28 days apart).

Question: Keeping all of the previous information in mind, what seems to be a ***primary*** goal of this clinical trial?

A. To see if the experimental vaccine is more effective than the control vaccine
B. To see if participants at one study site develop more antibodies than participants at the other study site
C. To compare the effectiveness of different dosages within different age brackets
D. To see if older patients are more likely to die than younger patients

Between May 30 and Aug 8, 2020, 560 participants were enrolled: 160 aged 18–55 years (100 assigned to ChAdOx1 nCoV-19, 60 assigned to MenACWY), 160 aged 56–69 years (120 assigned to ChAdOx1 nCoV-19: 40 assigned to MenACWY), and 240 aged 70 years and older (200 assigned to ChAdOx1 nCoV-19: 40 assigned to MenACWY). <u>One participant (in the 18–55 years low-dose group) received the incorrect vaccine after randomisation and was excluded from analysis. Seven participants randomly assigned to receive two doses of vaccine chose not to continue with the boost dose and were excluded from further analyses. Three participants were excluded from immunology analyses due to incorrectly labelled samples (either incorrect participant identification numbers or incorrect timepoints noted on the label,or both).</u>

Question: What validity threat would be related to the underlined information above? How concerning is this validity threat?

Using a multiplex immunoassay that detected total IgG against RBD and trimeric spike protein, we observed that participants who received the prime vaccination of standard-dose ChAdOx1 nCoV-19 had similar anti-spike antibody titres by day 28 after their prime vaccination as those who received a low dose (p=0.12)

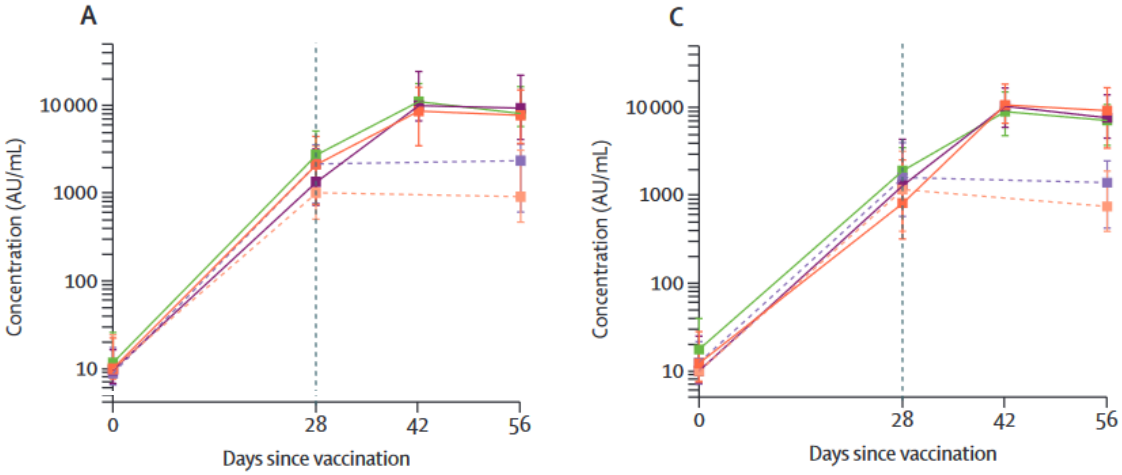
Question: What groups are being compared here, and what does the p-value communicate about this comparison?

At both dose levels, and for all dose groups combined, anti-spike IgG responses at day 28 decreased with increasing age (low-dose groups: 18–55 years, median 6439 arbitrary units [AU]/mL [IQR 4338–10640], n=49;56–69 years, 4553 AU/mL [2657–12462], n=60; ≥70 years, 3565 AU/mL [1507–6345], n=93; p=0.0037; standard-dose groups: 18–55 years, median 9807 AU/mL [IQR 5847–17220], n=43; 56–69 years, 5496 AU/mL [2548–12061], n=55; ≥70 years, 4156 [2122–12595], n=97; p=0.0044)

Question: What groups are being compared here, and what do the p-values communicate about these comparisons?

By 28 days after the boost vaccination, similar antibody titres were seen across all two-dose groups, regardless of age or vaccine dose (e.g., standard-dose groups: 18–55 years, median 20713 AU/mL [IQR 13898–33550], n=39; 56–69 years, 16170 AU/mL [10233–40353], n=26; and ≥70 years, 17561 AU/mL [9705–37796], n=47; p=0.68), and were higher than for those who did not receive a boost vaccination

Question: What additional information is provided by this test result?



The left image represents immune response for participants who received the standard dose, and the right represents participants who received the low dose. The dashed lines represent the one-dose groups and the solid lines represent the two-dose groups.

Question: How do these graphs represent and support the statistical findings reported