## Chapter 12: Reading Biomedical Research - Experiments

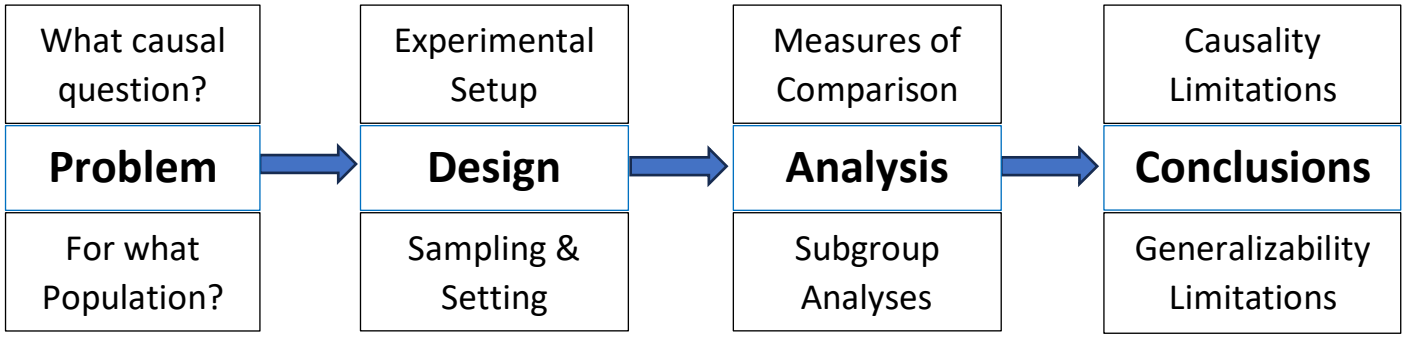
### Reading Experimental Studies

Start (and end) with the abstract!

- An **abstract** is about a 150 to 300-word _____ that will typically highlight 1) the motivation and aims of the paper, 2) a short description of the methods, 3) key results, and 4) a brief statement of implications or future work needed.
- Use the abstract to help you answer some of these key details up front.
- If reading through the paper for more detail, come back to the abstract periodically, or once finished with the paper, to orient yourself back to the paper's key points and contributions.

To make sense of the research articles we read, let's consider four parts of the research process. Furthermore, we'll consider both the causality argument and the generalizability argument put forward.

- **Problem:** What problem or gap in the research are the authors trying to address?
  - What causal relationship are they trying to understand? What variables are involved?
  - What population or setting does this problem or question apply to?
- **Design:** What process are the authors using to gather data to address this problem?
  - What kind of experiment and what experimental features are being incorporated? Have they cleanly isolated the proposed causal factor?
  - Who comprises the sample? Do they represent the population/setting of interest?
- **Analysis:** What are the analytical tests and methods that the authors used to make sense of their data?
  - What statistical measures were reported? What do we learn from the confidence intervals or p-values in determining evidence for a difference?
  - What subgroups were examined, and do the effects vary across different populations or conditions?
- **Conclusions:** What claim is being made by the authors? What uncertainty accompanies this claim?
  - What threats persist to the study's causality argument**?**
  - What threats persist to the study's generalizability argument
  - Are there any additional limitations or subjectivities in the design or analysis that we should make note of in understanding the strength of evidence and implications of this study?



## Abstract

**Importance.** Valganciclovir for 200 days is standard care for cytomegalovirus (CMV) prophylaxis in high-risk CMV-seronegative kidney transplant recipients who receive an organ from a CMV-seropositive donor, but its use is limited by myelosuppression.

**Objective.** To compare the efficacy and safety of letermovir with valganciclovir for prevention of CMV disease in CMV-seronegative kidney transplant recipients who receive an organ from a CMV-seropositive donor.

**Design, Setting, and Participants.** Randomized, double-masked, double-dummy, noninferiority, phase 3 trial in adult CMV-seronegative kidney transplant recipients who received an organ from a CMV-seropositive donor at 94 participating sites between May 2018 and April 2021 (final follow-up in April 2022).

**Interventions.** Participants were randomized in a 1:1 ratio (stratified by receipt of lymphocyte-depleting induction immunosuppression) to receive letermovir, 480 mg, orally daily (with acyclovir) or valganciclovir, 900 mg, orally daily (adjusted for kidney function) for up to 200 days after transplant, with matching placebos.

**Main Outcomes and Measures.** The primary outcome was CMV disease, confirmed by an independent masked adjudication committee, through posttransplant week 52 (prespecified noninferiority margin, 10%). CMV disease through week 28 and time to onset of CMV disease through week 52 were secondary outcomes. Exploratory outcomes included quantifiable CMV DNAemia and resistance. The rate of leukopenia or neutropenia through week 28 was a prespecified safety outcome.

**Results.** Among 601 participants randomized, 589 received at least 1 dose of the study drug (mean age, 49.6 years; 422 [71.6%] men). Letermovir (n = 289) was noninferior to valganciclovir (n = 297) for prevention of CMV disease through week 52 (10.4% vs 11.8% of participants with committee-confirmed CMV disease; stratum-adjusted difference −1.4% [95% CI, −6.5% to 3.8%]). No participants who received letermovir vs 5 participants (1.7%) who received valganciclovir developed CMV disease through week 28. Time to onset of CMV disease was comparable between the groups (hazard ratio, 0.90 [95% CI, 0.56-1.47]). Quantifiable CMV DNAemia was detected in 2.1% of participants in the letermovir group vs 8.8% in the valganciclovir group by week 28. Of participants evaluated for suspected CMV disease or CMV DNAemia, none (0/52) who received letermovir and 12.1% (8/66) who received valganciclovir had resistance-associated substitutions. The rate of leukopenia or neutropenia through week 28 was lower with letermovir vs valganciclovir (26% vs 64%; difference, −37.9% [95% CI, −45.1% to −30.3%]; $P < .001$). Fewer participants in the letermovir group than the valganciclovir group discontinued prophylaxis due to adverse events (4.1% vs 13.5%) or drug-related adverse events (2.7% vs 8.8%).

**Conclusion and Relevance**. Among adult CMV-seronegative kidney transplant recipients who received an organ from a CMV-seropositive donor, letermovir was noninferior to valganciclovir for prophylaxis of CMV disease over 52 weeks, with lower rates of leukopenia or neutropenia, supporting its use for this indication.

**Letermovir vs Valganciclovir for Prophylaxis of Cytomegalovirus in High-Risk Kidney Transplant Recipients.**


**Problem**

*What are the response and explanatory variables in this study?*

*What problem / causal relationship are they trying to understand?*

*What population or setting does this problem or question apply to?*

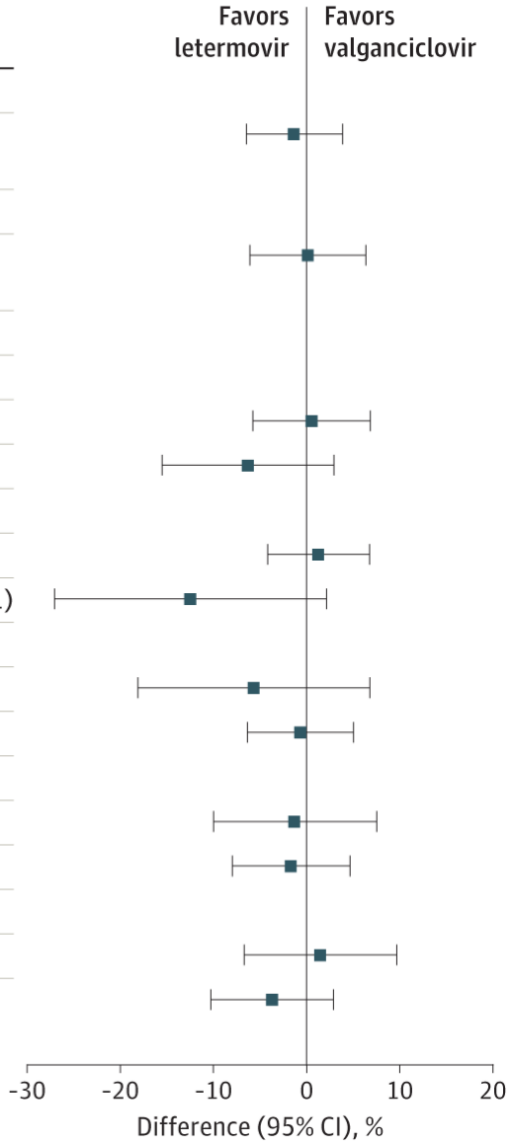
**Design**

*What kind of experiment and what experimental features are incorporated? Is it well designed?*

*Who comprises the sample? Any representation questions at this point?*

**Figure 2. Primary outcome of CMV Disesase with Letermovir vs Valganciclovir through Week 52.**

| | No./total No. (%) | | Difference (95% CI), %[b] |
|---|---|---|---|
| | Letermovir[a] | Valganciclovir | |
| Primary outcome | | | |
| CMV disease[c] | 30/289 (10.4) | 35/297 (11.8) | -1.4 (-6.5 to 3.8) (noninferior) |
| Sensitivity analysis | | | |
| Investigator-reported CMV disease[d] | 50/289 (17.3) | 51/297 (17.2) | 0.1 (-6.1 to 6.3) |
| Subgroup analysis | | | |
| Sex | | | |
| Men | 25/210 (11.9) | 24/209 (11.5) | 0.5 (-5.8 to 6.8) |
| Women | 5/79 (6.3) | 11/88 (12.5) | -6.3 (-15.5 to 2.9) |
| Age, y | | | |
| <65 | 26/242 (10.7) | 23/242 (9.5) | 1.2 (-4.2 to 6.7) |
| ≥65 | 4/47 (8.5) | 12/55 (21.8) | -12.5 (-27.0 to 2.1) |
| Race[e] | | | |
| Non-White | 2/37 (5.4) | 6/53 (11.3) | -5.7 (-18.1 to 6.7) |
| White | 28/250 (11.2) | 29/243 (11.9) | -0.7 (-6.4 to 5.0) |
| Region | | | |
| US | 15/118 (12.7) | 16/116 (13.8) | -1.3 (-10.0 to 7.5) |
| Non-US | 15/171 (8.8) | 19/181 (10.5) | -1.7 (-8.0 to 4.6) |
| Lymphocyte-depleting induction immunosuppression | | | |
| Use | 18/131 (13.7) | 17/138 (12.3) | 1.4 (-6.7 to 9.6) |
| Nonuse | 12/158 (7.6) | 18/159 (11.3) | -3.7 (-10.3 to 2.8) |




**Analysis Part 1**

*What statistical measure(s) was/were reported? What do we learn from the confidence intervals or p-values in determining evidence for a difference?*

*What subgroups were examined, and do the effects vary across different populations or conditions?*

**Table 2. Adverse Events through Week 28 in the Safety Population**

Table 2. Adverse Events Through Week 28 in the Safety Population[a]

| Adverse event | No. (%) Letermovir (n = 292) | No. (%) Valganciclovir (n = 297) | Difference (95% CI), %[b] |
|---|---|---|---|
| Adverse event summary | | | |
| ≥1 adverse event | 271 (92.8) | 276 (92.9) | −0.1 (−4.4 to 4.2) |
| Serious adverse events[c] | 106 (36.3) | 113 (38.0) | −1.7 (−9.5 to 6.1) |
| Drug-related adverse events[d] | 58 (19.9) | 104 (35.0) | −15.2 (−22.2 to −8.0) |
| Serious drug-related adverse events[c,d] | 4 (1.4) | 15 (5.1) | −3.7 (−7.0 to −0.9) |
| Death | 2 (0.7) | 1 (0.3) | 0.3 (−1.3 to 2.2) |
| Discontinued due to adverse events | 12 (4.1) | 40 (13.5) | −9.4 (−14.1 to −4.9) |
| Discontinued due to serious adverse events[c] | 6 (2.1) | 14 (4.7) | −2.7 (−5.9 to 0.3) |
| Discontinued due to drug-related adverse events[d] | 8 (2.7) | 26 (8.8) | −6.0 (−10.1 to −2.4) |
| Discontinued due to serious drug-related adverse events[c,d] | 2 (0.7) | 7 (2.4) | −1.7 (−4.2 to 0.4) |
| Adverse events in ≥10% of participants | | | |
| Diarrhea | 92 (31.5) | 85 (28.6) | 2.9 (−4.5 to 10.3) |
| Tremor | 53 (18.2) | 52 (17.5) | 0.6 (−5.6 to 6.9) |
| Urinary tract infection | 41 (14.0) | 42 (14.1) | 0.1 (−5.8 to 5.6) |
| Peripheral edema | 39 (13.4) | 38 (12.8) | 0.6 (−4.9 to 6.1) |
| Hypomagnesemia | 37 (12.7) | 39 (13.1) | −0.5 (−5.9 to 5.0) |
| Leukopenia | 33 (11.3) | 110 (37.0) | −25.7 (−32.3 to −19.1) |
| Hypertension | 33 (11.3) | 36 (12.1) | −0.8 (−6.1 to 4.5) |
| Increased creatinine | 30 (10.3) | 41 (13.8) | −3.5 (−8.9 to 1.8) |
| Hypophosphatemia | 30 (10.3) | 35 (11.8) | −1.5 (−6.7 to 3.6) |
| Hyperkalemia | 27 (9.2) | 32 (10.8) | −1.5 (−6.5 to 3.4) |
| Nausea | 25 (8.6) | 33 (11.1) | −2.5 (−7.5 to 2.3) |
| Fatigue | 18 (6.2) | 32 (10.8) | −4.6 (−9.3 to −0.1) |
| Neutropenia | 8 (2.7) | 49 (16.5) | −13.8 (−18.7 to −9.3) |

**Analysis Part 2!**

*What do we learn from the confidence intervals or p-values in determining evidence for a difference in terms of safety outcomes?*

**Conclusions**

**Summary** on Reading Tables and Figures in an Experimental Design Research Paper

- Experimental studies often have a table comparing **group demographics**. While randomization of larger groups is highly reliable for creating balanced groups, this can provide an additional check!
- What kinds of measures did the researchers use to compare these groups?
    - **Relative risks** or **odds ratios** (one risk divided by the other) are generally used when comparing risks with a categorical response variable.
    - **Absolute risk reductions** (one risk minus the other) might be used with a categorical response variable as well.
    - **Hazard ratios** are used when making time-to-event comparisons. They will often accompany **Kaplan-Meier plots**.
    - **Two mean** or **two median** differences are often made with a numeric response variable.
- In addition to reporting general group differences, researchers will often share a **subgroup analysis.**
    - These are *often* reported with a **Forest Plot** to quickly show which subgroups saw differences as a result of the causal factor or not!
    - Keep in mind that the ***uncertainty*** in these comparisons will be higher (sometimes ***much higher!)*** because the **subgroups** will have much **smaller samples sizes** than the whole groups!
- Also be on the lookout for tables/figures reporting **adverse effect comparisons**. Even if a treatment is effective at treating the primary response variable, it's good to also check whether it increases other risk factors.
- Finally, check the end of the paper for **limitations** section! This will likely cover any weaknesses to the causality or generalizability arguments, and perhaps a few other important limitations.

**Reflection Questions**

---

**12.1.** What is an abstract?

**12.2.** What kinds of things should we consider when evaluating a study's causality argument? What about when evaluating a generalizability argument?

**12.3.** Why might the confidence intervals for comparisons of subgroups often be much wider than the confidence intervals for the full group comparisons?