MADA Final Project

Trends in PrEP Use in the US, 2012-2022

Nicole Luisi

2023-02-03

# 1. PART 1

**Overview**:

The purpose of this analysis is to measure the public health impact of pre-exposure prophylaxis (PrEP) in the US over the past 10 years. Previous research by [Sullivan et al. (2018)](https://doi.org/10.1016/j.annepidem.2018.06.009) evaluated trends in the prevalence of PrEP use from 2012-2017. Outcomes include period prevalence of PrEP use and PrEP-to-need ratio (PnR) by year. This analysis will extend the previous work to include data through the year 2022; subgroups will include region, state, age group, gender, and race/ethnicity *(race/ethnicity PrEP data were not available at the time of the previous analysis and were not included)*. (Patrick S. Sullivan et al., 2018)

**Data Source**:

*Note: New diagnoses data are currently available through 2020; PrEP data through 2021. If 2021-2022 data are not available by the time the analysis needs to be completed, the analysis will be temporarily restricted to 2012-2020.*

* HIV New Diagnoses: [AIDSVu](https://aidsvu.org/) (Patrick Sean Sullivan et al., 2020)
* PrEP and PnR: IQVIA (Quintiles IMS Holdings)
* Annual Population Estimates: National Historical Geographic Information System (NHGIS), American Community Surveys (ACS), US Census Bureau

**Analysis**:

* N/% of annual PrEP users by age, gender, race/ethnicity, and region
* Annual prevalence of PrEP use per 100,000 population by age, gender, race/ethnicity, region, and state
* Estimated annual percent change (EAPC) for prevalence
  + Calculated using [Joinpoint Regression Software (R-Callable Command-Line Version 4.9.0.0)](https://surveillance.cancer.gov/help/joinpoint/tech-help/citation)
* Prep-to-need ratio by age, gender, race/ethnicity, and region
* Other modeling TBD based on course modules

The structure below is one possible setup for a data analysis project (including the course project). For a manuscript, adjust as needed. You don’t need to have exactly these sections, but the content covering those sections should be addressed.

This uses MS Word as output format. [See here](https://quarto.org/docs/output-formats/ms-word.html) for more information. You can switch to other formats, like html or pdf. See [the Quarto documentation](https://quarto.org/) for other formats.

# 2. Summary/Abstract

*Write a summary of your project.*

# 3. Introduction

## 3.1 General Background Information

*Provide enough background on your topic that others can understand the why and how of your analysis*

## 3.2 Description of data and data source

*Describe what the data is, what it contains, where it is from, etc. Eventually this might be part of a methods section.*

## 3.3 Questions/Hypotheses to be addressed

*State the research questions you plan to answer with this analysis.*

To cite other work (important everywhere, but likely happens first in introduction), make sure your references are in the bibtex file specified in the YAML header above (here dataanalysis\_template\_references.bib) and have the right bibtex key. Then you can include like this:

Examples of reproducible research projects can for instance be found in []

# 4. Methods

*Describe your methods. That should describe the data, the cleaning processes, and the analysis approaches. You might want to provide a shorter description here and all the details in the supplement.*

## 4.1 Data aquisition

*As applicable, explain where and how you got the data. If you directly import the data from an online source, you can combine this section with the next.*

## 4.2 Data import and cleaning

*Write code that reads in the file and cleans it so it’s ready for analysis. Since this will be fairly long code for most datasets, it might be a good idea to have it in one or several R scripts. If that is the case, explain here briefly what kind of cleaning/processing you do, and provide more details and well documented code somewhere (e.g. as supplement in a paper). All materials, including files that contain code, should be commented well so everyone can follow along.*

## 4.3 Statistical analysis

*Explain anything related to your statistical analyses.*

# 5. Results

## 5.1 Exploratory/Descriptive analysis

*Use a combination of text/tables/figures to explore and describe your data. Show the most important descriptive results here. Additional ones should go in the supplement. Even more can be in the R and Quarto files that are part of your project.*

**?@tbl-summarytable** shows a summary of the data.

Note the loading of the data providing a **relative** path using the ../../ notation. (Two dots means a folder up). You never want to specify an **absolute** path like C:\ahandel\myproject\results\ because if you share this with someone, it won’t work for them since they don’t have that path. You can also use the here R package to create paths. See examples of that below.

**?(caption)**

## 5.2 Basic statistical analysis

*To get some further insight into your data, if reasonable you could compute simple statistics (e.g. simple models with 1 predictor) to look for associations between your outcome(s) and each individual predictor variable. Though note that unless you pre-specified the outcome and main exposure, any “p<0.05 means statistical significance” interpretation is not valid.*

**?@fig-result** shows a scatterplot figure produced by one of the R scripts.

## 5.3 Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You want the code to produce figures and data ready for display as tables, and save those. Then you load them here.*

Example **?@tbl-resulttable2** shows a summary of a linear model fit.

**?(caption)**

# 6. Discussion

## 6.1 Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## 6.2 Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## 6.3 Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

This paper [] discusses types of analyses.

These papers [] are good examples of papers published using a fully reproducible setup similar to the one shown in this template.

Note that this cited reference will show up at the end of the document, the reference formatting is determined by the CSL file specified in the YAML header. Many more style files for almost any journal [are available](https://www.zotero.org/styles). You also specify the location of your bibtex reference file in the YAML. You can call your reference file anything you like, I just used the generic word references.bib but giving it a more descriptive name is probably better.

# 7. References

Sullivan, Patrick S., Giler, R. M., Mouhanna, F., Pembleton, E. S., Guest, J. L., Jones, J., … Siegler, A. J. (2018). Trends in the use of oral emtricitabine/tenofovir disoproxil fumarate for pre-exposure prophylaxis against HIV infection, united states, 2012–2017. *Annals of Epidemiology*, *28*(12), 833–840. https://doi.org/<https://doi.org/10.1016/j.annepidem.2018.06.009>

Sullivan, Patrick Sean, Woodyatt, C., Koski, C., Pembleton, E., McGuinness, P., Taussig, J., … Sanchez, T. H. (2020). A data visualization and dissemination resource to support HIV prevention and care at the local level: Analysis and uses of the AIDSVu public data resource. *J Med Internet Res*, *22*(10), e23173. <https://doi.org/10.2196/23173>