Mental Health Plays Significant Role in Physical Wellnesss in 1972 to 2022 US General Social Surveys*

Racial and Gender Unequity and Mental Illness Being Overlooked in the Medical Field

Mingjia Chen

April 14, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

*Code and data are available at: https://github.com/MjChen120/Mental_to_Physical_Health.git.

1 Introduction

Mental health is one essential aspect of the general health of an individual and yet it is often overlooked. Mental illnesses are not only "silent diseases" of one's mind, but could as well impact patients' physical well-being in general (Ohrnberger, Fichera, and Sutton 2017). This paper examines whether mental health-related and demographic factors impact the number of days an individual experiences physical unwellness in a month?

U.S. General Social Survey (GSS) data from 1972 to 2022 is used for this paper to examine whether the demographic and mental health-related factors as indicators of physical health conditions. The estimand is the difference in whether an individual could experience more occasions of feeling physically unwell if they have mental health unwellness versus not. This is considered in terms of those who were diagnosed as having depression by a doctor.

The analysis of GSS data, including gender, age, and mental and physical health variables, seeks to provide detailed insights into how mental health issues impact physical unwellness, contributing to discussions on intersectionality and overlooked mental health in the medical field.

This paper begins with Introduction framing the impact of mental health on one's well-being. The Data section details variables, methodology, and analysis cleaning processes. In the Model section, a model is proposed for determining how an individual's depression status and counted days of mental unwellness in the past 30 days impact their counted days of physical unwellness in the past 30 days. Results section analyzes how demographic and other mental health-related factors would affect physical health with model calculations, leading the Discussion section with gender intersectionality and mental health issues often being overlooked in the medical field. Limitations and Future Research sections are included for study limitations and future research directions. The Appendix provides further survey information.

2 Data

The paper uses data collected from the US General Social Survey (GSS) from NORIC at the University of Chicago ("General Social Survey" 2024). From the dataset, this paper focuses specifically on variables related to demographic backgrounds, mental health, and physical health, from 1992 to 2022. This longitudinal approach allows us to analyze whether an individual's demographic factors and mental health well-being could significantly contribute to their physical well-being.

2.1 Source Data

The data was downloaded and filtered for the selected variables from the selected data variables from GSS[1]. The data cleaning was performed based on value definitions as defined in the GSS codebooks (NORC 2018). The variable names are renamed to be more informative (Table 1).

Table 1: Source data retrieved from GSS

Variable	New Name	Description	Example
ID_	id	Response ID	1
YEAR	year	Year of the Data Recorded	1977
SEX	sex	Respondent's gender	Female/2
AGE	age	Respondent's age	25
HEALTH	health	Respondents' health condition	Good/2
PHYSHLTH	phys_days	Days of Respondents' physical health being not good	15
MNTLHLTH	ment_days	Days of Respondents' mental health being not good	20
DEPRESS	depress	Whether respondents have been told having depression	Yes/1

2.2 Data Cleaning

The data was cleaned by using the open source statistically programming language R (R Core Team 2024), with libraries `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016),

[1] https://gss.norc.org/documents/stata/GSS_stata.zip

dplyr (Wickham et al. 2022), readr (Wickham, Hester, and Bryan 2022), tibble (Muller and Wickham 2022), here (Müller 2020), kableExtra (Zhu 2021), janitor (Firke 2023), arrow (Richardson et al. 2024), and knitr (Xie 2014).

2.3 Survey Methodology

will be written

2.4 Demographic Variables

Table 2: Counted days of Physical and Mental unwellness of Population by Age and Gender Groups

ID	Gender	Age	Physical Unwellness	Mental Unwellness	Age Cohort
1	Female	25	0	0	20-39
2	Male	43	0	0	40-59
3	Female	30	0	2	20-39
4	Female	55	0	0	40-59
5	Male	37	0	30	20-39
6	Male	47	0	0	40-59

2.5 Mental and Physical Unwellness and Depression Diagnosis

Table 3: Counted days of Physical and Mental unwellness with Depression Diagnosis

ID	Physical Unwellness	Mental Unwellness	Depression Diagnosis
1	0	0	No
2	0	0	Yes
4	0	0	No
14	0	0	No
16	14	7	No
19	0	0	No

2.6 Mental Unwellness and General Health Status

Table 4: Counted days of Mental unwellness with General Health Status

ID	Mental Unwellness	General Health
1	0	1
2	0	2
3	2	2
5	30	2
6	0	1
9	5	2

2.7 Non-response Rate

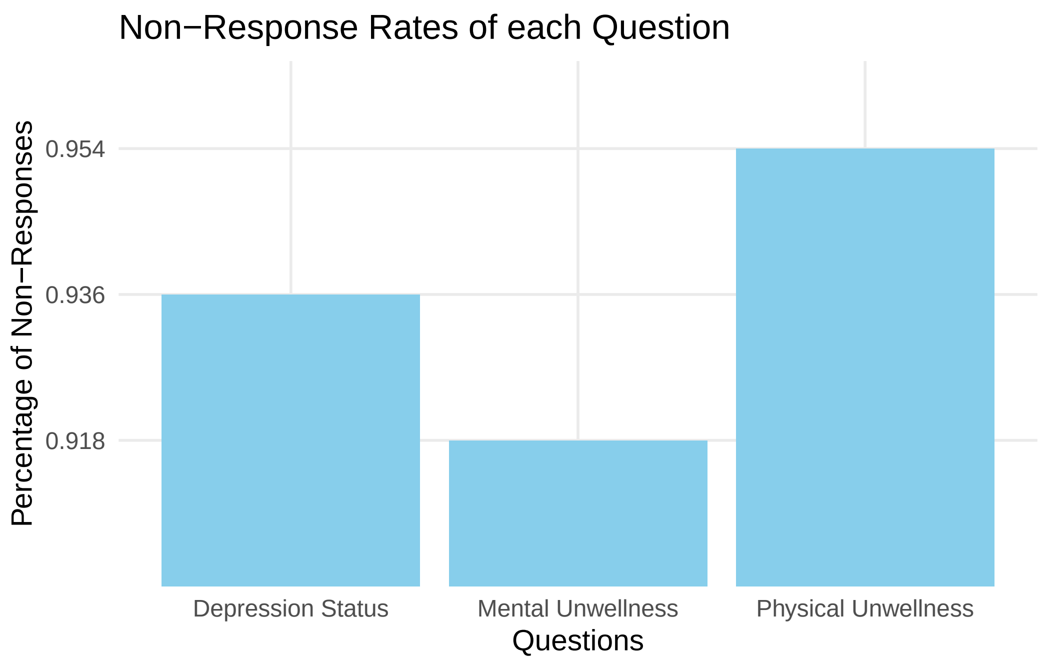


Figure 1: Nonresponse Rates

As shown in Figure 1, some of the questions in the dataset are not included resulting no responses collected in certain years due to the change in survey methodology. This fact requires extra attention when cleaning and handling since the number of available responses will decrease when additional variables are added. Hence, an additional sub set of the data that only includes counted days of mental and physical un-wellness in the past 30 days (Table 5) is also included for only speculating these two variables while keeping as many responses as possible.

Table 5: Counted days of Mental unwellness with General Health Status

ID	Physical Unwellness	Mental Unwellness
1	0	0
2	0	0
3	0	2
4	0	0
5	0	30
6	0	0

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of days that an individual feel physically unwell in the past 30 days. Then β_i is the number of days that this individual feel mentally unwell and γ_i indicates whether the individual is diagnosed as depressive by a doctor. y_i and β_i are measured in days.

$$
\begin{align}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) \tag{1}\\
\mu_i &= \alpha + \beta_i + \gamma_i \tag{2}\\
\alpha &\sim \text{Normal}(0, 2.5) \tag{3}\\
\beta &\sim \text{Normal}(0, 2.5) \tag{4}\\
\gamma &\sim \text{Normal}(0, 2.5) \tag{5}\\
\sigma &\sim \text{Exponential}(1) \tag{6}
\end{align}
$$

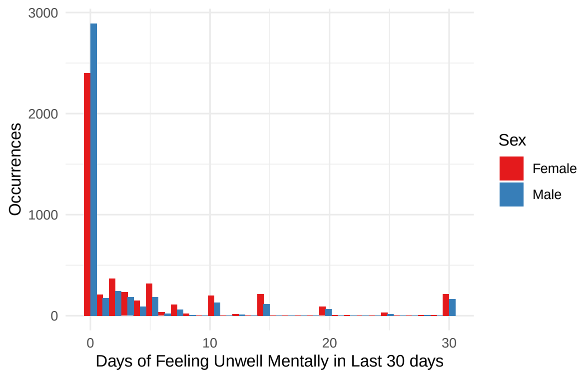
We run the model in R (R Core Team 2024) using the `rstanarm` package of Goodrich et al. (2024). We use the default priors from `rstanarm`.

3.1.1 Model justification

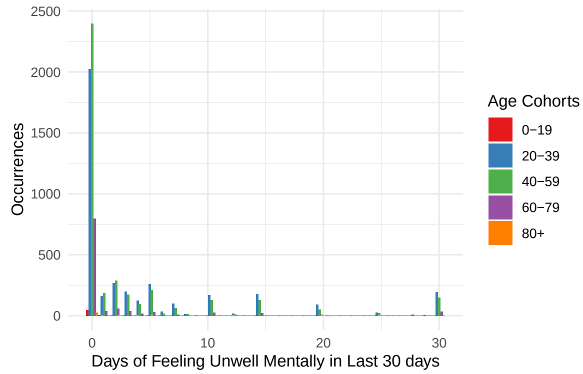
will be written

4 Results

Our results are summarized in Table 6.

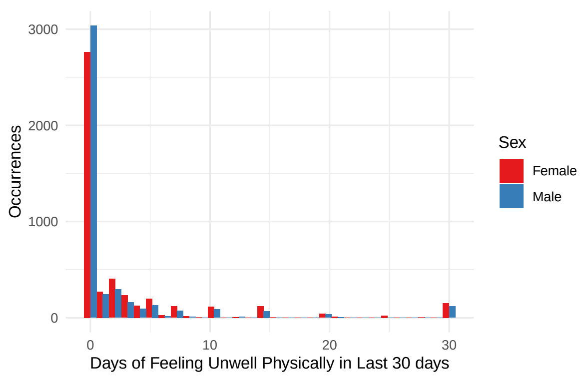


(a) Gender

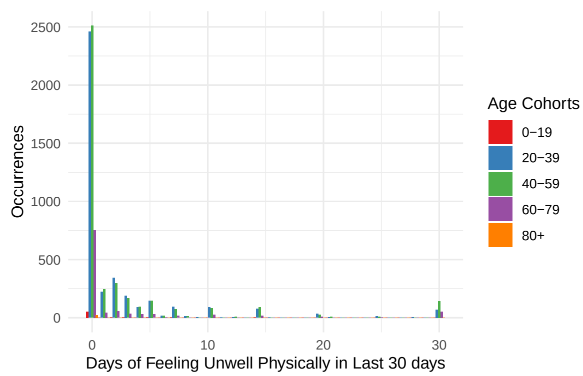


(b) Age cohorts

Figure 2: Female Participants from 20 to 59 Years Old are More likely to Experience Mental Unwellness



(a) Gender



(b) Age cohorts

Figure 3: Female Participants from 20 to 59 Years Old are More likely to Experience Physical Unwellness

Include some research about female having trouble due to medical racism etc. and possible reason why these age cohorts are more likely to experience both physical and mental unwellness.

4.1 Depression Playing an Role

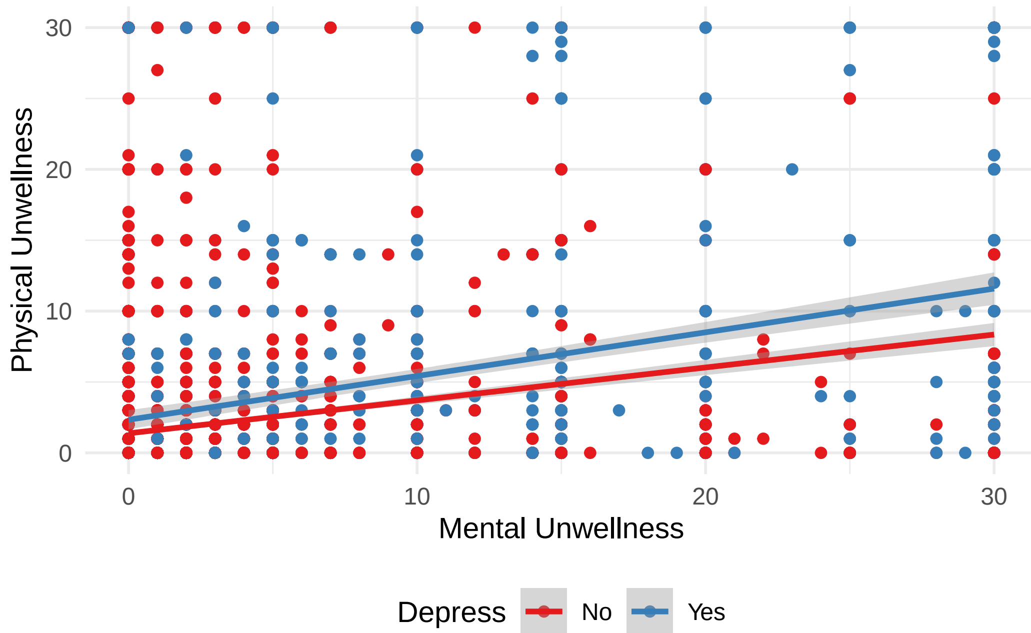


Figure 4: Counts of Days Feeling Unwell in Last 30 Days

4.2 Mental Health and Health in General

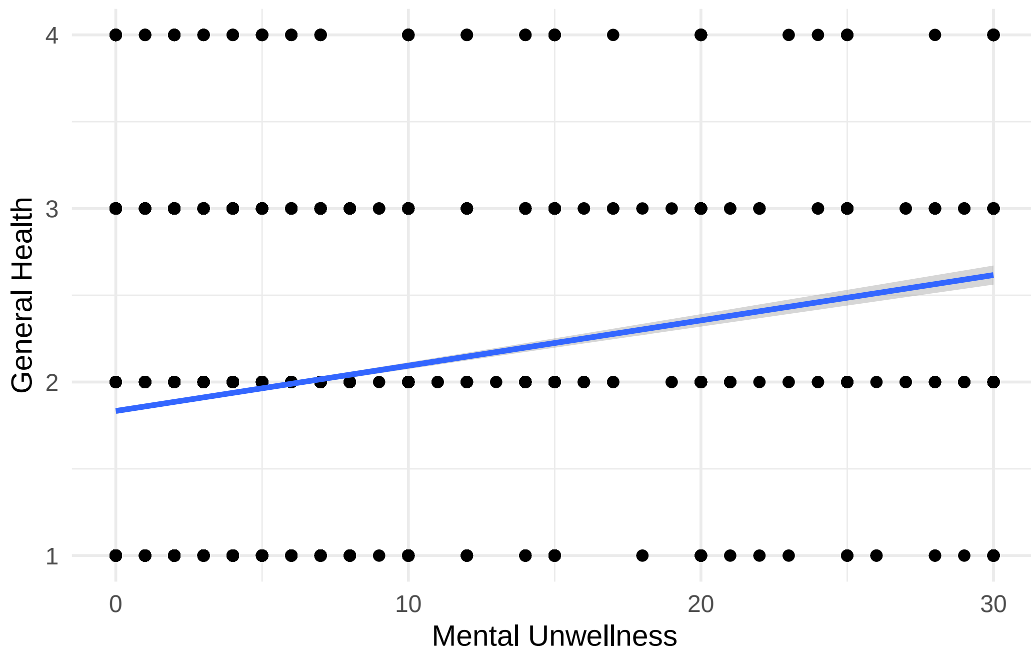


Figure 5: Counts of Days Feeling Unwell in Last 30 Days

Table 6: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	4.15
	(0.43)
ment_days	0.27
	(0.01)
depress	−1.43
	(0.22)
Num.Obs.	4527
R2	0.132
R2 Adj.	0.130
Log.Lik.	−14 351.092
ELPD	−14 358.6
ELPD s.e.	118.8
LOOIC	28 717.1
LOOIC s.e.	237.7
WAIC	28 717.1
RMSE	5.76

4.3 Model Results

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

A Appendix

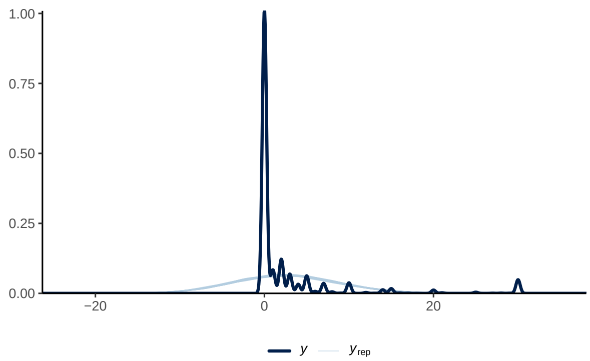
A.1 Additional data details

B Model details

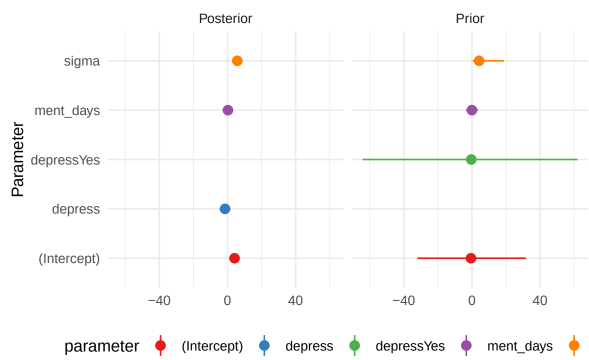
B.1 Posterior predictive check

In Figure 6a we implement a posterior predictive check. This shows...

In Figure 6b we compare the posterior with the prior. This shows...



(a) Posterior prediction check



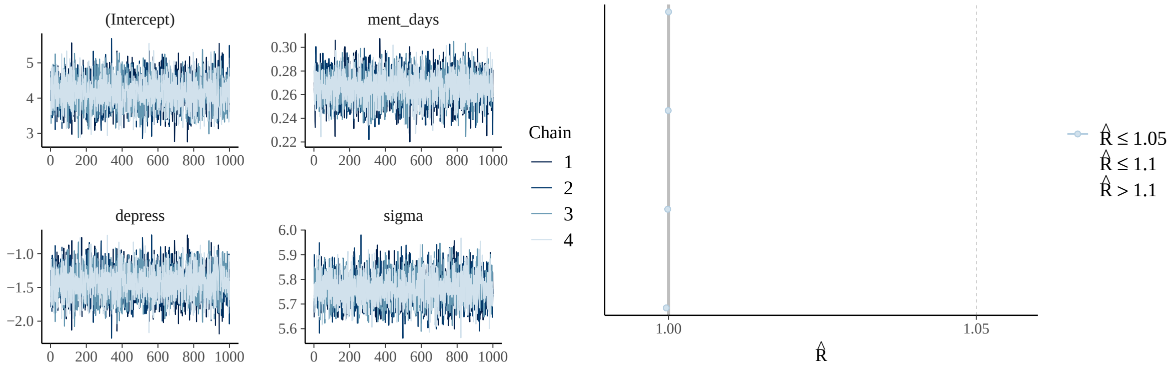
(b) Comparing the posterior with the prior

Figure 6: Examining how the model fits, and is affected by, the data

B.2 Diagnostics

Figure 7a is a trace plot. It shows... This suggests...

Figure 7b is a Rhat plot. It shows... This suggests...



(a) Trace plot

(b) Rhat plot

Figure 7: Checking the convergence of the MCMC algorithm

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

"General Social Survey." 2024. *General Social Survey*. NORC. https://gss.norc.org/get-the-data/stata.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Muller, Kirill, and Hadley Wickham. 2022. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

NORC. 2018. *1972-2018 GSS Cros-Section Codebook*. https://gss.norc.org/Documents/codebook/GSS_Codebook.pdf.

Ohrnberger, Julius, Eleonora Fichera, and Matt Sutton. 2017. "The Relationship Between Physical and Mental Health: A Mediation Analysis." *Social Science &Amp; Medicine* 195 (December): 42–49. https://doi.org/10.1016/j.socscimed.2017.11.008.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Toronto, Canada: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal* *of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain Francois, Lionel Henry, and Kirill Muller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.