



Cervical Cancer Screening

Help prevent cervical cancer by identifying at-risk populations

\$100,000 · 40 teams · a year ago

[Overview](#)[Data](#)[Kernels](#)[Discussion](#)[Leaderboard](#)[More](#)[Submit Predictions](#)

Overview

Description

Evaluation

Prizes

Educational Resources

Timeline

This is a Masters competition. You must be a [Kaggle Master](#) to participate.

Cervical cancer is the third most common cancer in women worldwide, affecting over 500,000 women and resulting in approximately 275,000 deaths every year. After reading these statistics, you may be surprised to hear that cervical cancer is potentially preventable and curable.

Cervical cancer can be prevented through early administration of the [HPV vaccine](#) and regular [pap smear](#) screenings, which indicate the presence of precancerous cells. It is also sometimes curable by the removal of the early-stage cancerous tissue that is identified through pap smears. Screening and early treatment can lead to potential cures in about 95% of women at risk for cervical cancer.

Most women in the US have access to cervical cancer screening, yet 4,000 women die every year from cervical cancer in the US and it is estimated that 30% of US women do not receive regular pap screenings. We know little about who these women are and why they are not getting screened. Prior research suggests that lower screening rates are associated with low income, low education, lack of interaction with the healthcare system, and lack of health insurance. But research also shows that even in women with access to healthcare fail to get this preventive test, indicating that barriers like lack of education and not being comfortable with the procedure are influencing their behavior ([Patient Survey](#)).

There are many patient advocacy programs on the importance of pap smears in cervical cancer prevention. However, these widespread programs may not be reaching or effectively speaking to the most vulnerable populations. If one could better identify these women, education campaigns could target them with content that speaks directly to their unique risk factors. Identifying predictors of not receiving pap smears will provide important information to stakeholders in cervical cancer prevention who run awareness programs.



With this Masters competition, Genentech is asking you to join their mission to help prevent cervical cancer. Given a dataset of de-identified health records, your challenge is to predict which women will not be screened for cervical cancer on the recommended schedule. Identifying at-risk populations will make education and other intervention efforts more effective, ideally ultimately reducing the number of women who die from this disease.

About Genentech

Founded more than 35 years ago, Genentech is a leading biotechnology company that discovers, develops, manufactures and commercializes medicines to treat patients with serious or life-threatening medical conditions. The company, a member of the Roche Group, has headquarters in South San Francisco, California. For additional information about the company, please visit <http://www.gene.com>.

Acknowledgements

The dataset for this competition is provided by [Symphony Health Solutions](#).



@Jotform.Show(35)

Leaderboard >

- 1 Michael & Giulio
- 2 :-D , :-) , :-P , 8-)
- 3 n_m
- 4 :- (, :- [, 8- { , : ^ (, 8- [
- 5 Medrr
- 6 better luck in 2016 !
- 7 AERO
- 8 SDDDS

Kernels >

Pawel & Maks insights

8 votes · a year ago

Plot test3

2 votes · a year ago

Last week

1 vote · a year ago

Age and Screeners

1 vote · a year ago

PlotTest5

1 vote · a year ago

38 discussion topics >

Finding scripts from this com...

1 reply · 10 months ago

Announcement: Insights Awa...

7 replies · a year ago

Award Event in San Francisco

10 replies · a year ago

More Insights

1 reply · a year ago

Jiwei & Xueer's insight

1 reply · a year ago

Launch

a year ago

Close

a year ago

a year ago
Rules Acceptance
Deadline

40
Teams

63
Competitors

Points **This competition did not award standard [ranking points](#)**

Tiers **This competition counted towards [tiers](#)**