Cervical Cancer in French regions in 2023

This document is describing cervical cancer disease in France in 2023. A departemental analysis of prevention plan of this cancer is not foreseen in first approach but at a regional granularity level. For more details about prevalence and to find all the datas for 2023, please see http://geodes.santepubliquefrance.fr. All other informations concerning the estimation of regional population on the 1st of January have been taken from https://statistiques-locales.insee.fr

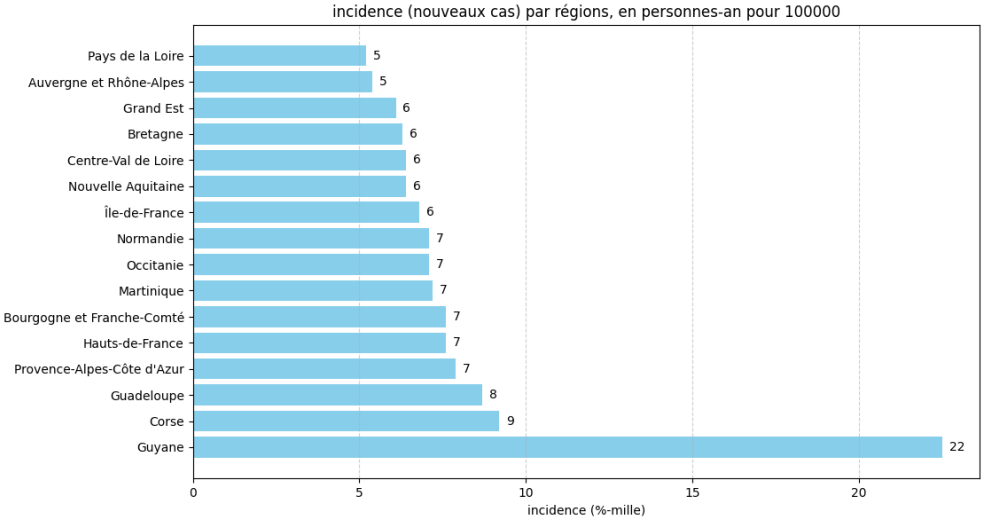
A first panel of the national strategy of prevention of cervical cancer among women consists in screening test every 2 years after the age of 25 years. Before jumping to the vaccination panel f the national prevention strategy, we start by analyzing the **correlations** figures between new cases of cervical cancer (incidence per 100000) and screening rate for each class of age in 16 regions of France

I) Regional Incidence of new cases (year 2022)

Here are some figures to start with:

1. Population (regional estimation of living persons on the 1st of Jan 2023)min: 301099, max: 12358932, mean: 4178717, std-deviation : 3279597
2. regional incidence (new cases of cervical cancers per 100000)min: 5.2 , max: , max: 22.5, mean: 7.969, std-deviation: 4.02
3. screening test (%) age class : 35 / 39 years oldmin: 37.1 % , max: 72.7%, mean: 62.03% , std-deviation: 10.61%

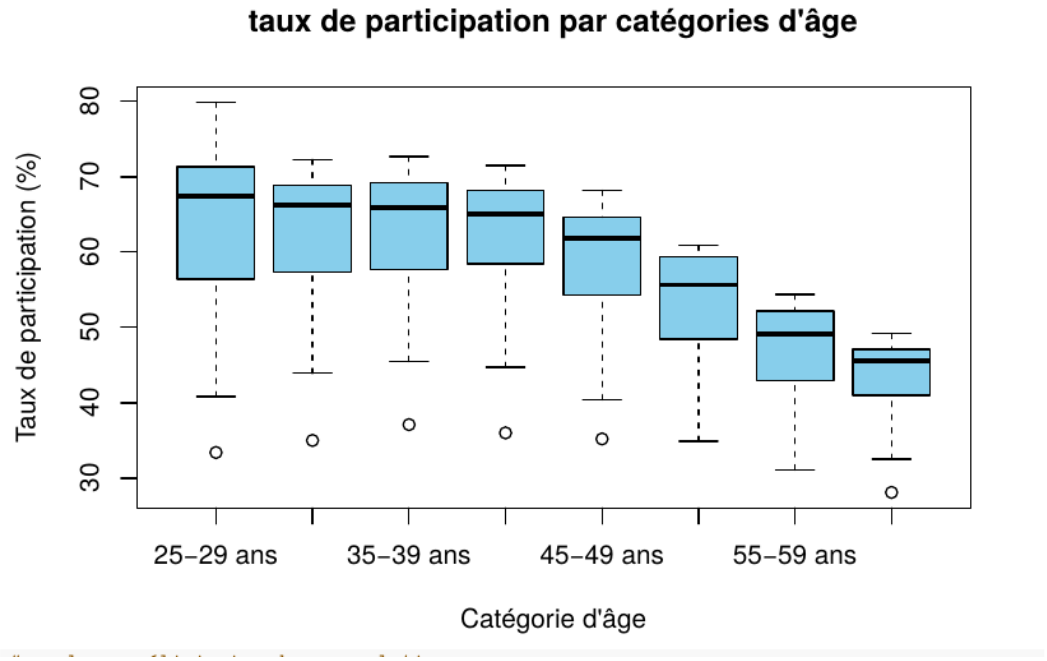
The following graph shows the outlier region “Guyane” at the bottom with 22 new cases of cervical cancer per 100.000 persons-year, not to be analysed yet.



Contact : erwan.simon@inserm.fr

II) Spliting cancer screening by ages

It appears on the following graph that a clear decrease in screening rate appear before the middle age of 50, known to be the age of median diagnostic of cervical cancer :



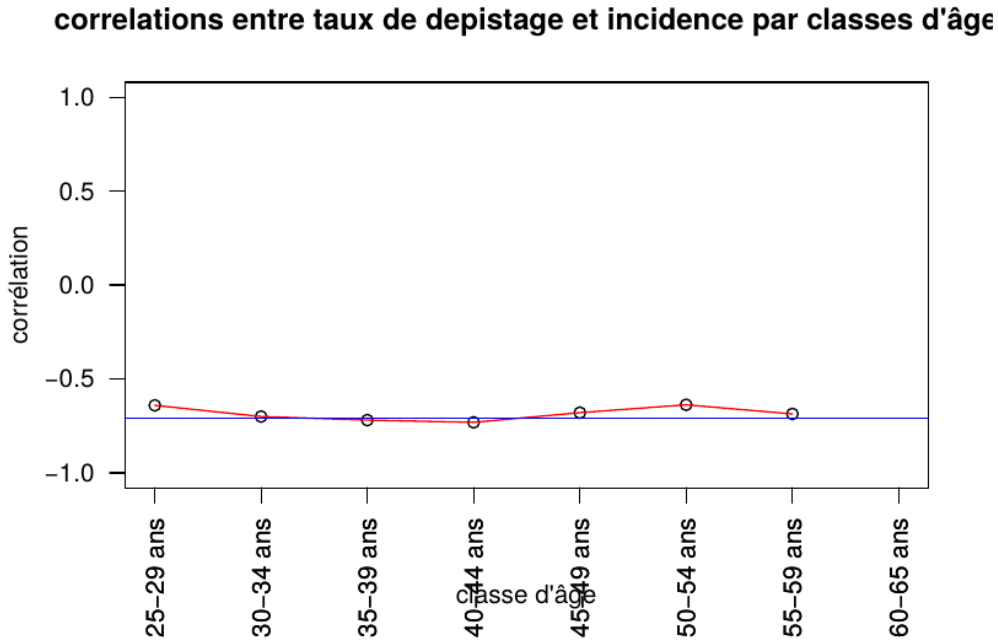
dramatic decrease of cancer screening rate after 50 years old !

Contact : erwan.simon@inserm.fr

Preliminary analysis of correlations

The following R code gives a correlation matrix between French regional screening rates for each class of age and incidence (new cases of cervical per 100000)

```
# analyse préliminaire des correlations
taux_depistage= donnees[,c("depistage_vingtaine","depistage_trentaineA","depistage_trentaineB","depista
correlations= cor(taux_depistage,donnees$`incidence(%)`)
```



La correlation la plus importante entre le taux de dépistage du cancer du col de l'utérus et le nombre de

highest correlation for 40-44 years old class (incidence & regional screening rate)

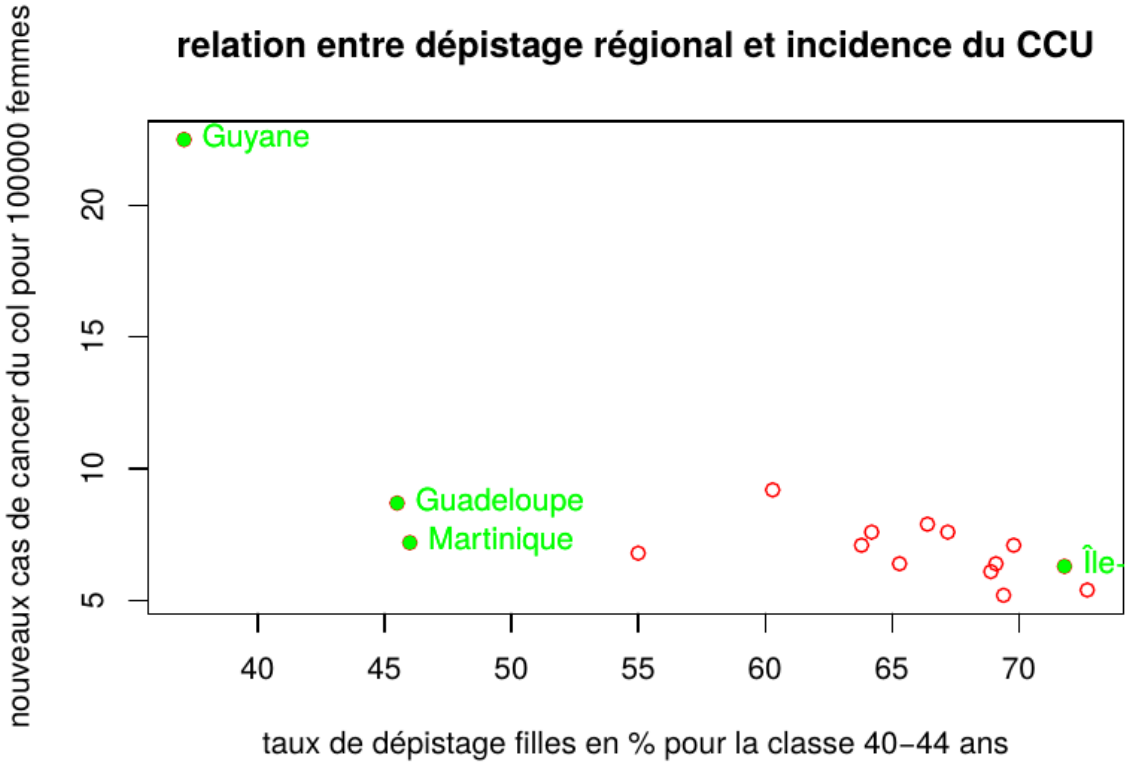
The highest correlation between the rate of cervical cancer screening and the number of new cases is found between the two intersections of the "red curve" and the "blue line" above; in absolute terms it is greater than 0.7.

Here is the development of the Student's T test for the correlation (screening rate vs incidence):

```
cor.test(donnees$`incidence(%)`,donnees$depistage_quarantaineA, conf.level = 0.95, method=c("pearson")

##
##  Pearson's product-moment correlation
##
## data:  donnees$`incidence(%)` and donnees$depistage_quarantaineA
## t = -4.0114, df = 14, p-value = 0.001287
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.9005333 -0.3694975
## sample estimates:
##       cor
## -0.7312633
```

Since the **correlation between cancer screening rate and the number of new cases, i.e incidence is highly significant** (~ 0. 001), we can now look at the *"satellite" regions*. These are made up of the ***"outlier" regions***, i.e. the singular points where the data do not fall within the range of measurements of participation in cervical cancer screening and do not lie within the 2 median quartiles, i.e. **French Guyana, Guadeloupe and Martinique**, as shown in the following graph, where "Ile de France" appears as an indication but not in the outliers.

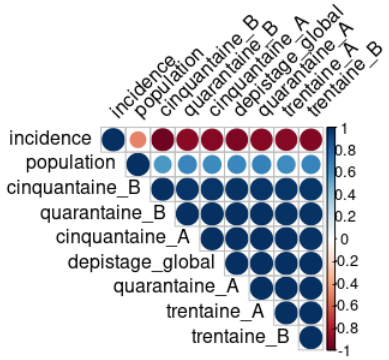


a weak screening rate of cervical cancer in "Antilles Françaises"

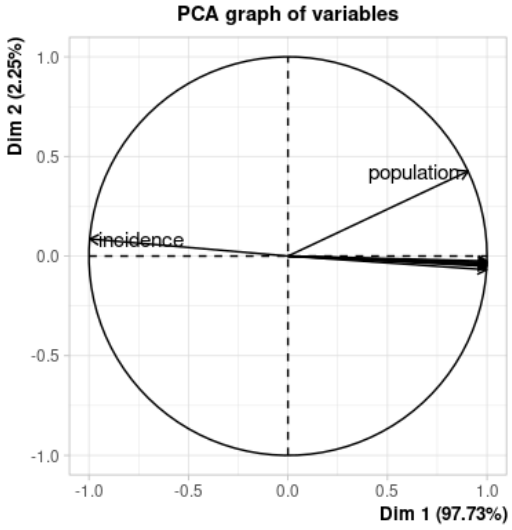
Contact : erwan.simon@inserm.fr

Multidimensional Analysis (Analyse en Composante Principale)

Including the population figures at a regional level is interesting to proceed to an ACP. As mentionned before, those demographic figures have been taken from Insee website. A first correlation plot is proposed here for our study:



Here is the graphical result of the normalized ACP:



We can see that incidence and overall screening are almost on the same axis of symmetry but opposite; there is therefore a strong correlation between them.

On the other hand: an angle close to 90° between “population” and “screening rate” means that no much correlation between these variables is to be expected.

Contact : erwan.simon@inserm.fr

ON THE OTHER HAND: we need to check further since there seems to be a correlation between the incidence (new cervical cancer cases) and the regional population!

Let's display the value of the correlation between incidence and regional population and its significance:

```
> cor.test(indicateurs$incidence, indicateurs$population)

        Pearson's product-moment correlation

data:  indicateurs$incidence and indicateurs$population
t = -0.99274, df = 3, p-value = 0.394
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.9588608  0.6859305
sample estimates:
       cor
-0.4972713
```

Despite the relatively high level of inverse correlation (-0.4), the p-value is greater than 0.05 and therefore this variable is not significant for a further logistic regression model. We might explore later a specific adult population from the Insee database to improve, or jump to the second cancer prevention national measure which is the vaccination against HR-HPV 16 and 18 with nonavalent Gardasil for instance. This is the end of this article on regional screening rates and incidence.

E.S, Nice le 08-02-2024

Index	libelle_region	population	incidence	depistage_global	trentaine_A	trentaine_B
1	Corse	351255	9.2	56.8	63.1	61.1
2	PACA	5160091	7.9	59.7	64.9	65.4
3	Bretagne	3429882	6.3	66.4	75.2	74.1
4	Guadeloupe	375845	8.7	48	52.2	52.9
5	Guyane	301099	22.5	30	33.7	34.3

Contact : erwan.simon@inserm.fr