



Научный семинар  
Центра по изучению проблем народонаселения  
экономического факультета МГУ имени М. В. Ломоносова

**Александр АВДЕЕВ**

Сектор историко-демографических и библиографических исследований,  
Лаборатория экономики населения и демографии / Институт демографии  
Парижского университета Пантеон Сорбонна

Сообщение на тему:

# Методологические подходы к оценке смертности от COVID-19 во Франции

Москва, 25 мая 2020 года

# Предварительные замечания:

## меры распространения и присутствия болезни (патологии) в населении

- **Заболеваемость (incidence):**

число новых случаев выявленной патологии в течение определенного периода в данном населении

- Один из важнейших критериев оценки частоты и скорости распространения патологии (заболевания)

- **Распространенность (prevalence):**

мера присутствия определенной патологии в данном населении и в данный момент времени

$$\text{Распространенность} = \frac{\text{количество больных в момент } t}{\text{общая численность населения}}$$

- **Дополнительные характеристики заболевания**

- **длительность** болезни ( $\tau$ )

- при благополучном исходе ( $\tau_r$ )

- при летальном исходе ( $\tau_d$ )

$$\left. \begin{array}{l} \bullet \text{ длительность болезни } (\tau) \\ \bullet \text{ при благополучном исходе } (\tau_r) \\ \bullet \text{ при летальном исходе } (\tau_d) \end{array} \right\} \tau = \frac{\tau_r \cdot R + \tau_d \cdot D}{R + D}$$

$R$  – число выздоровевших в наблюдаемой когорте

$D$  – число умерших в наблюдаемой когорте

- **летальность** болезни ( $d$ )

доля умерших среди заболевших:  $d = \frac{D}{R + D}$

# Несколько важных определений:

- **Эндемия:**

постоянное присутствие болезни (патологии) в определенной географической зоне (заболеваемость и распространенность стабильны)

- **Эпидемия:**

возникновение и быстрое распространение заразного заболевания со значительными последствиями в ограниченном населении и на ограниченной территории,

- заболеваемость увеличивается по сравнению с обычным уровнем
- чаще всего вызывается инфекцией, передаваемой непосредственно или опосредованно
- не обязательно обусловлена наличием эндемии на данной территории (появление нового для данного населения патогена)
- эпидемия может:
  - развиться в пандемию
  - регрессировать в эндемию
  - исчезнуть

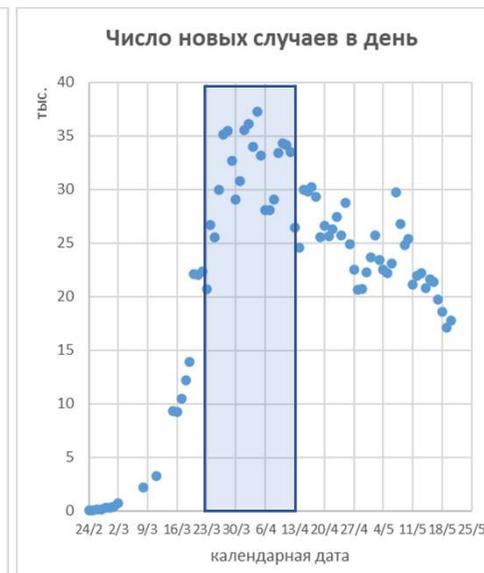
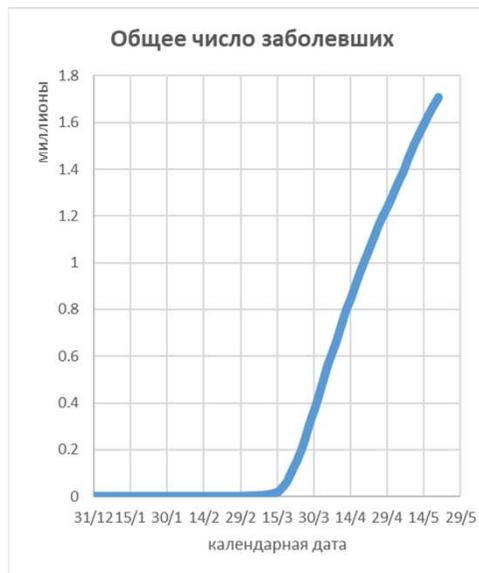
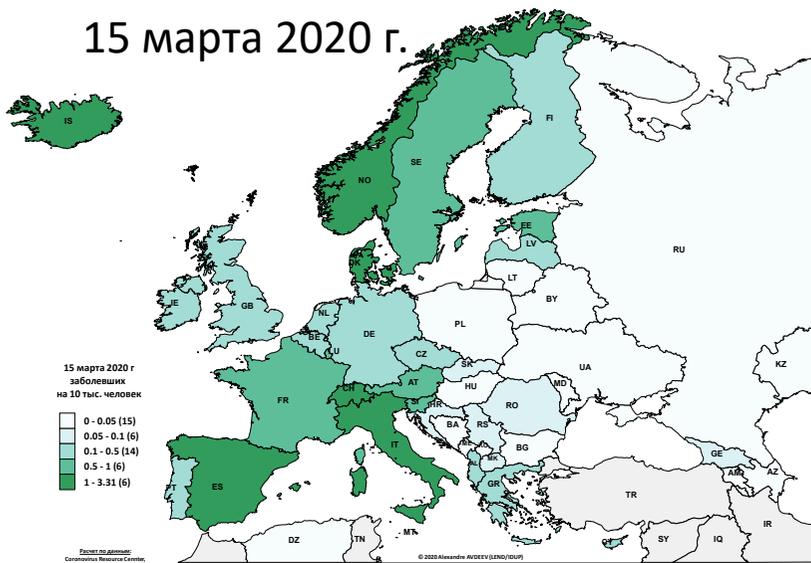
- **Пандемия:**

эпидемия, которая затрагивает большинство населения одного или нескольких континентов или всей планеты, возникает

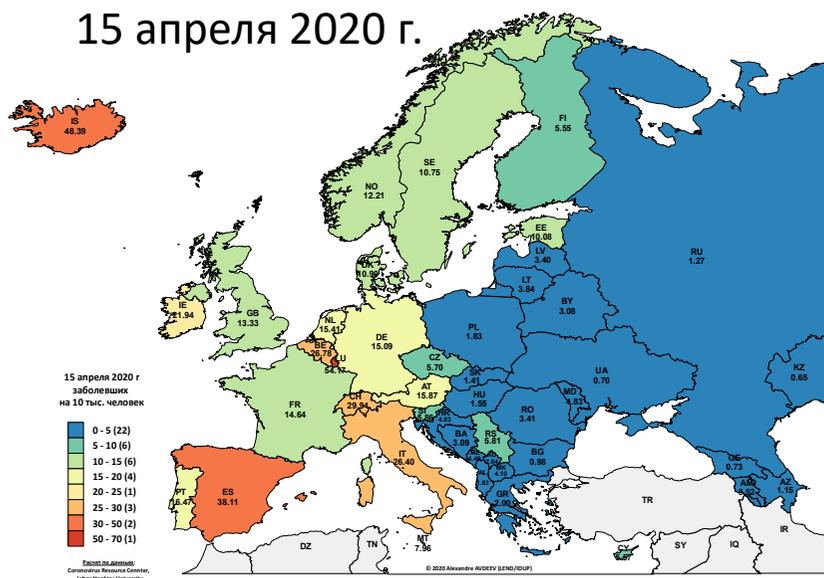
- либо в случае появления (воскрешения) **нового патогена (микроба, вируса), неизвестного иммунной системе человека (животного)**
- либо в случае **появления нового типа уже известного патогена** в результате значительной генетической мутации

# Пандемия Сов-19 в Европе: март-май 2020 года

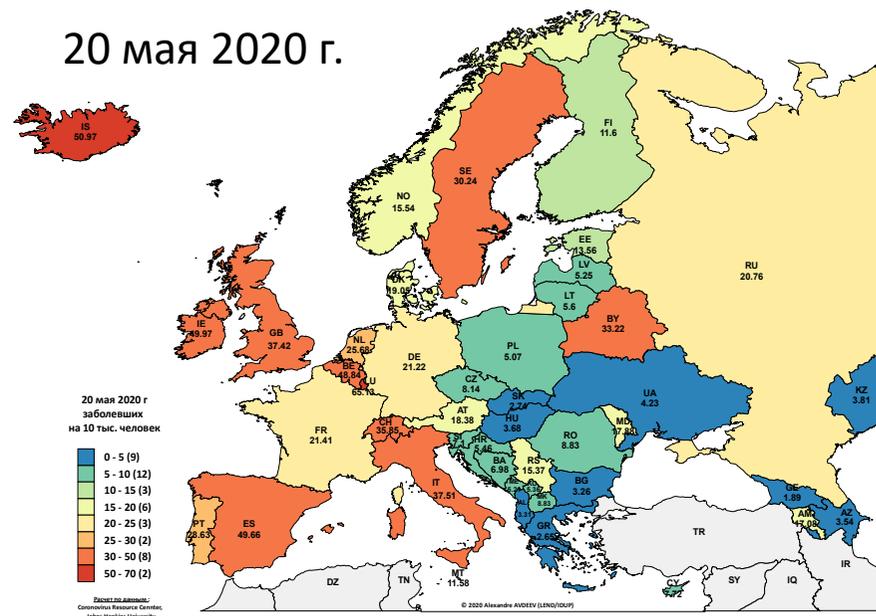
15 марта 2020 г.



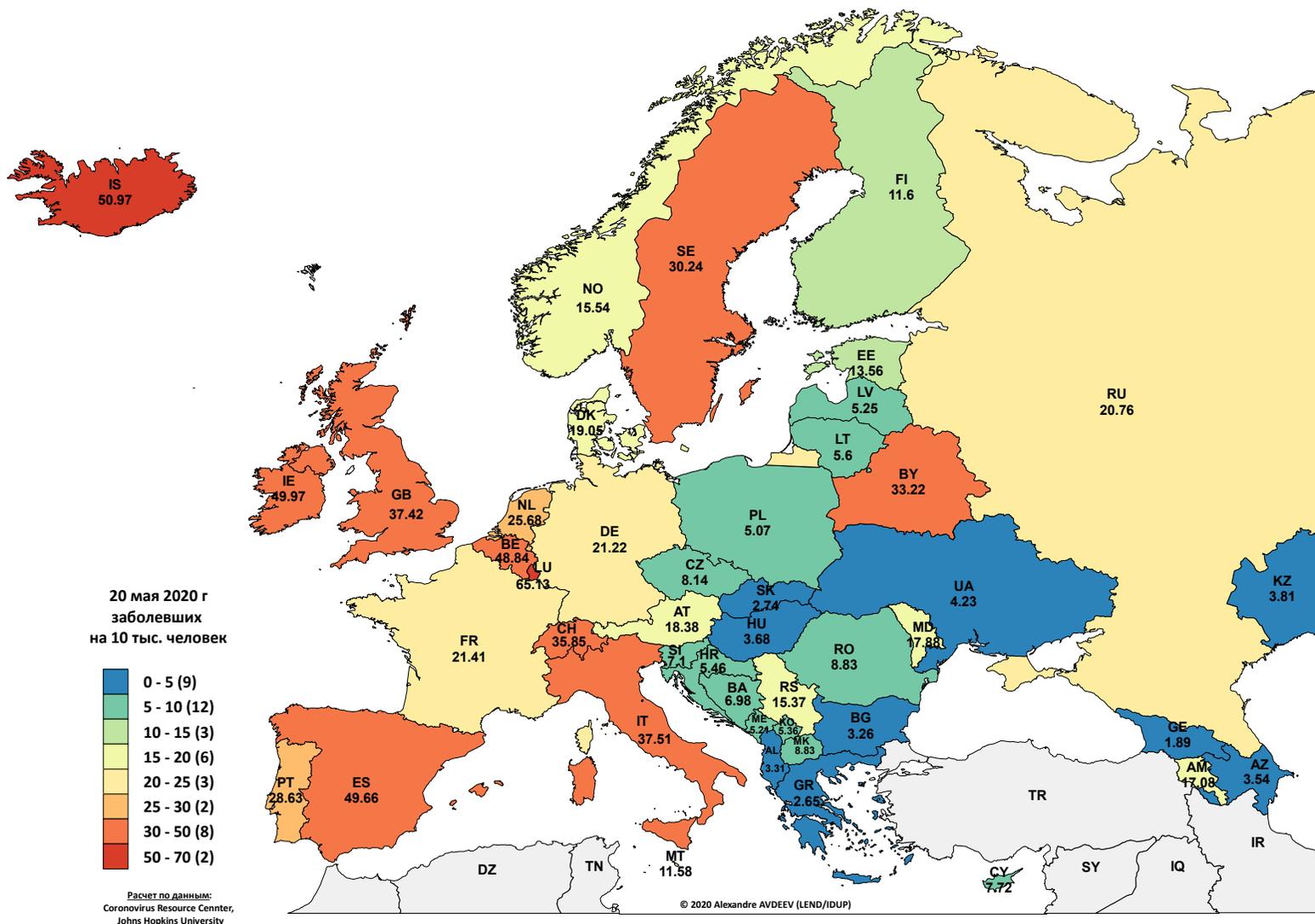
15 апреля 2020 г.



20 мая 2020 г.



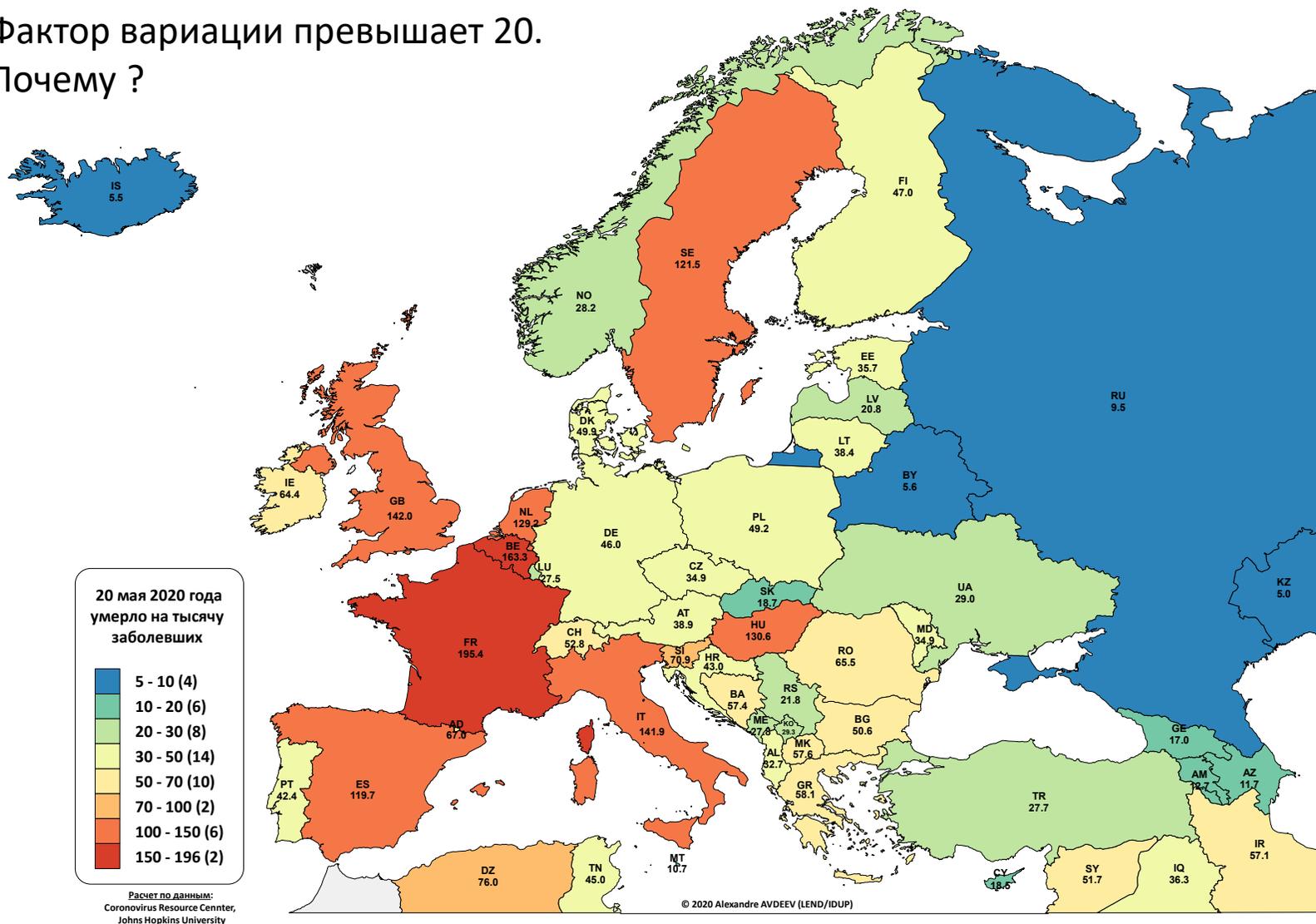
# Распространенность Сов-19 в странах Европы: число заболевших к 20 мая 2020 года общим итогом на 10 тыс. населения



# Летальность Covid-19 в странах Европы:

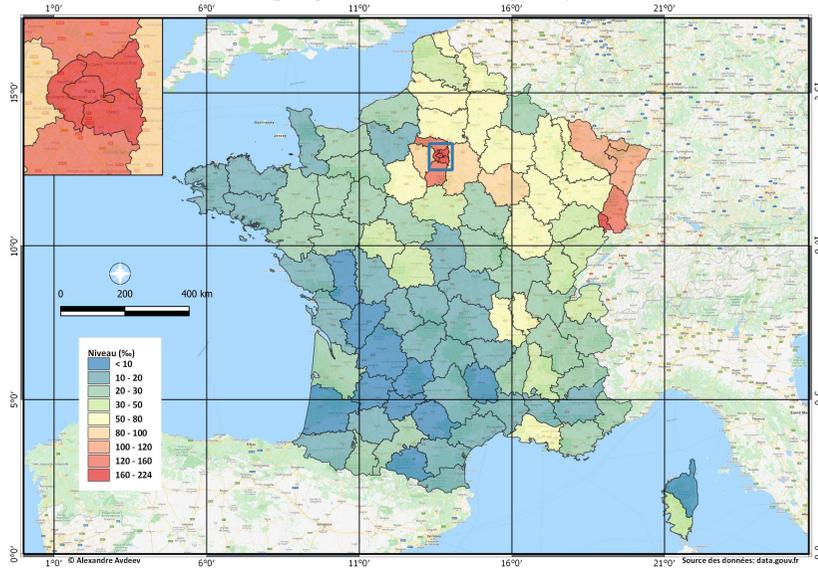
число умерших на 1000 заболевших общим итогом к 20 мая 2020 года

Фактор вариации превышает 20.  
Почему ?

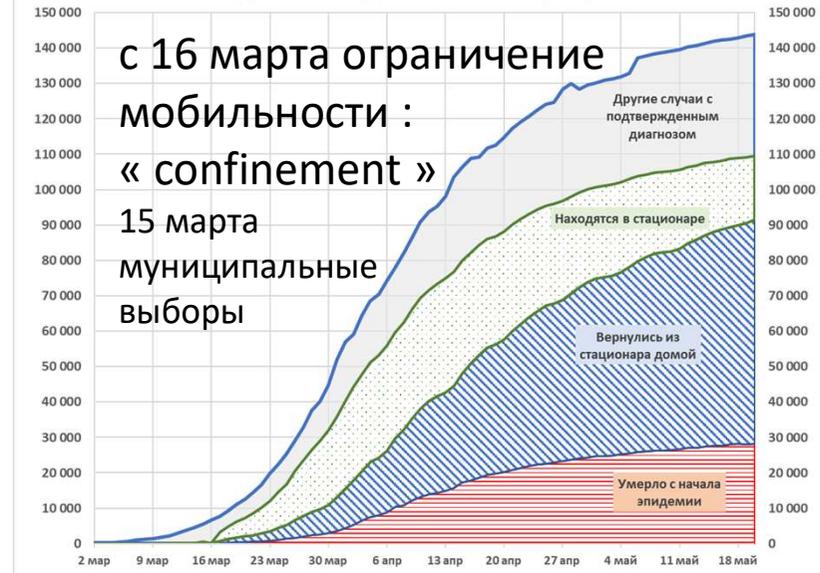


# Эпидемия Covid-19 во Франции

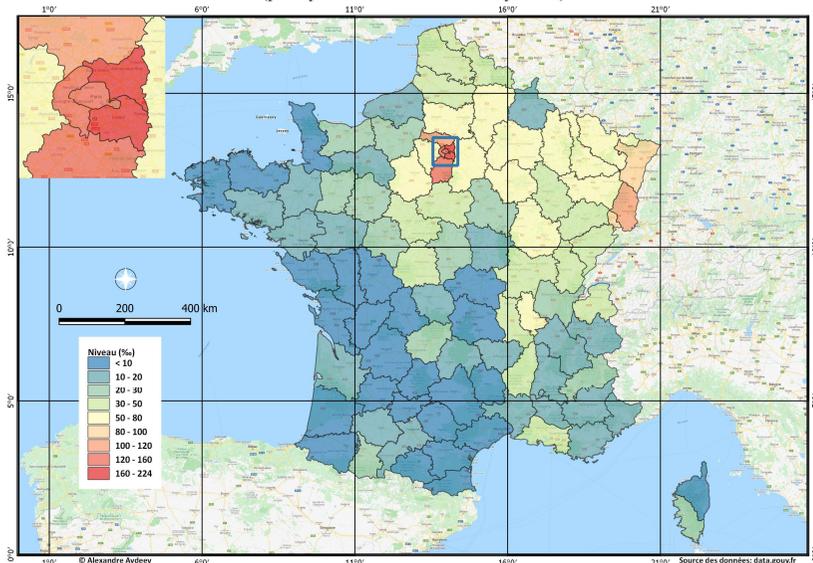
Nombre de personnes hospitalisées à cause de Covid-19 au 14 avril 2020 pour 1000 habitants (par départements de la France métropolitaine)



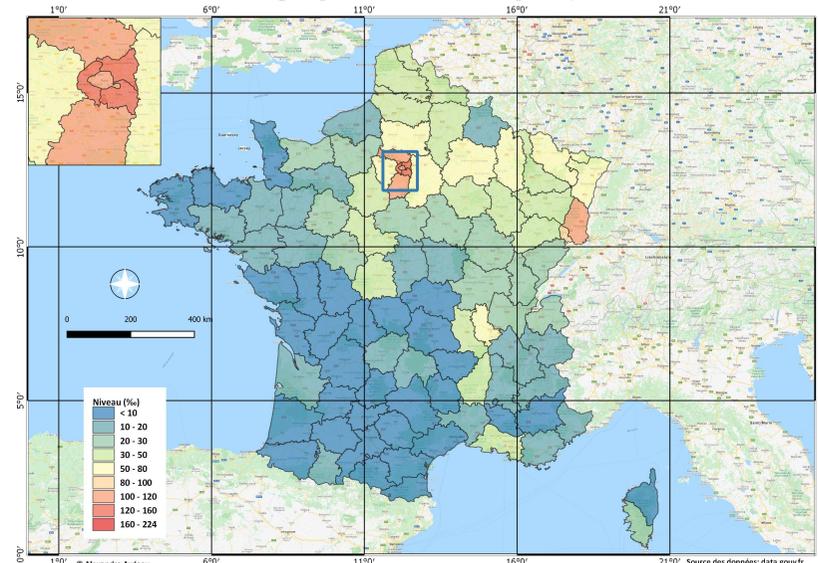
Динамика эпидемии Covid-19 во Франции



Nombre de personnes hospitalisées à cause de Covid-19 au 3 mai 2020 pour 1000 habitants (par départements de la France métropolitaine)

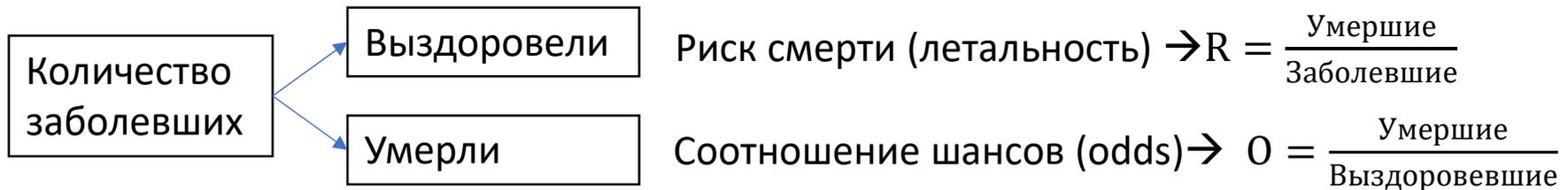


Nombre de personnes hospitalisées à cause de Covid-19 au 12 mai 2020 pour 1000 habitants (par départements de la France métropolitaine)



## Методы оценки смертности (дожития) когорты

**Риск смерти** – вероятность того, что болезнь завершится смертью пациента



Очевидно, что  $R = \frac{O}{1+O}$ , и что  $O = \frac{1-R}{R}$

Если  $t$  – интервал времени от момента  $T$  до  $T+t$ ;

$D_t$  – число умерших в интервале  $t$  (от данной причины)

$P(T)$  – численность заболевших в момент  $T$

$P(t)$  – численность населения, подверженного риску смерти в период  $t$

То можно рассчитать т.н. **общий коэффициент смертности (exposure rate)**

$$m_t = \frac{D_t}{\bar{P}_{(T, T+t)} \cdot t}$$

или оценку **средней** плотности вероятности умереть (силы смертности и т. П.), которая при *соблюдении условий стабильности динамики и однородности населения*, позволяет оценить риск для интервала  $t$

$$R(t) = 1 - \exp(-m_t \cdot t)$$

# Источники данных:

1. Оперативные (ежедневные) данные больничных стационаров (Etablissement hospitaliers) с 18 марта по 19 мая 2020 года (по полу, возрасту и административным территориям - департаментам) <https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>
  - A. Число лиц, госпитализированных в течение суток с диагнозом (в связи с) Covid-19 (новые случаи, incidences)
  - B. Число лиц, находящихся на стационарном лечении в связи с Covid-19 инфекцией
  - C. Число умерших в течение суток (из числа госпитализированных) в связи с Covid-19
  - D. Число лиц, выздоровевших (выписанных домой) в течение суток (из числа госпитализированных) в связи с Covid-19
  - E. Число лиц, выздоровевших (выписанных домой) из числа госпитализированных в связи с Covid-19
  - F. Число лиц, помещенных в отделения реанимации и интенсивной терапии в связи с Covid-19 инфекцией
  - G. Число лиц, находящихся в отделениях реанимации и интенсивной терапии в связи с Covid-19 инфекцией
2. Обобщённые данные Министерства здравоохранения (Santé publique France)
  - H. Общее число заболевших с начала эпидемии с подтвержденным диагнозом ( $\approx \Sigma A + L$  ?)
  - I. Общее число умерших ( $= \Sigma C + M$ )
  - J. Обобщённые данные госпитальной статистики (B,C,D,E,F)
  - K. Число случаев Covid-19 инфекции в специализированных медицинских учреждениях (EMS/ENPAD)
  - L. Число подтверждённых случаев Covid-19 инфекции в специализированных медицинских учреждениях (EMS/ENPAD): включено в пункт H
  - M. Число умерших по причине Covid-19 инфекции в специализированных медицинских учреждениях (EMS/ENPAD): включено в пункт I

# Качество данных

## Госпитальная отчетность (по территориям)



## Обобщенные данные на национальном уровне



### Качество данных неудовлетворительное:

- «Необъяснимая» цикличность
- Слишком сильные хаотические колебания

Big Data ? НЕТУ.....



Длительность болезни (госпитализации) ? НЕТУ.....

# Попытка прямого расчета

$$m_j = \frac{D_j}{P_j^H}$$

$j$  – день с начала эпидемии

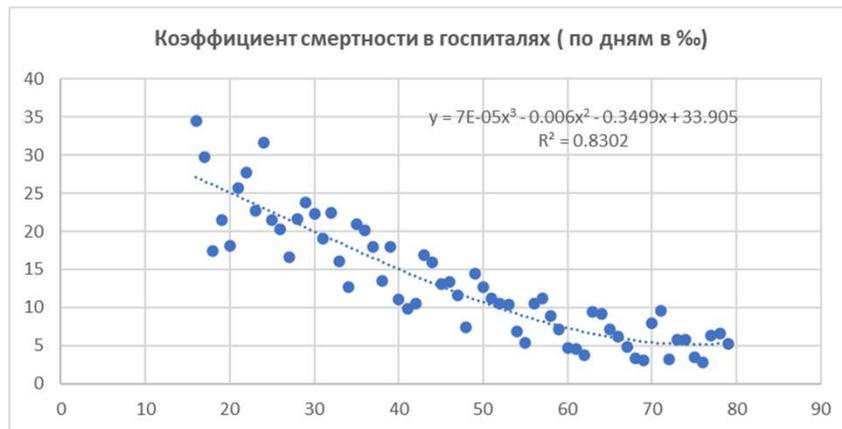
$D_j^H$  – число умерших в стационаре в день  $j$

$P_j^H$  – число, находящихся в госпиталях

$D_j$  – число умерших в день  $j$

$I_j$  – число новых случаев

$$\widehat{R}_j = \frac{\sum_j D_j}{\sum_j I_j}$$



Плохой результат!

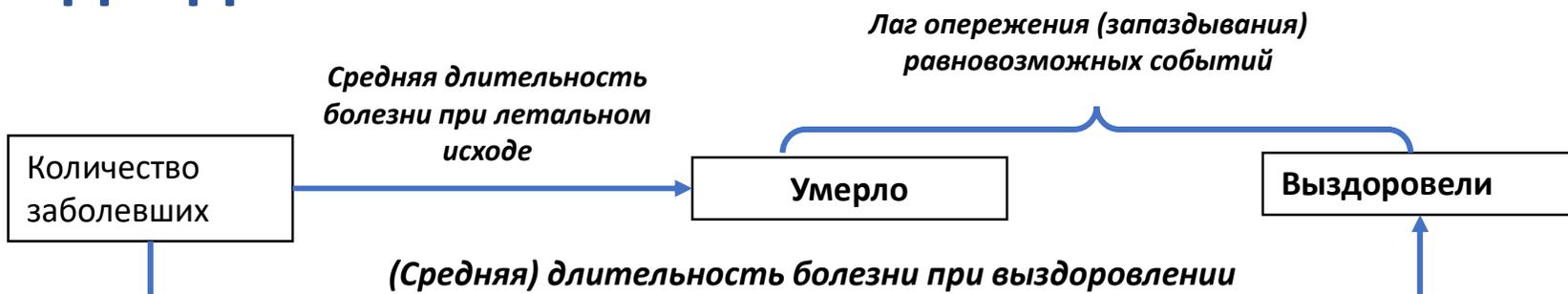
- Слишком выражена тенденция
- Тем не менее, средняя  $\sim 1,3\%$  при длительности болезни 14 дней дает величину риска, близкую к 16% (отношение общего умерших в госпиталях к числу госпитализированных)



Плохой результат! Но выглядит красиво...

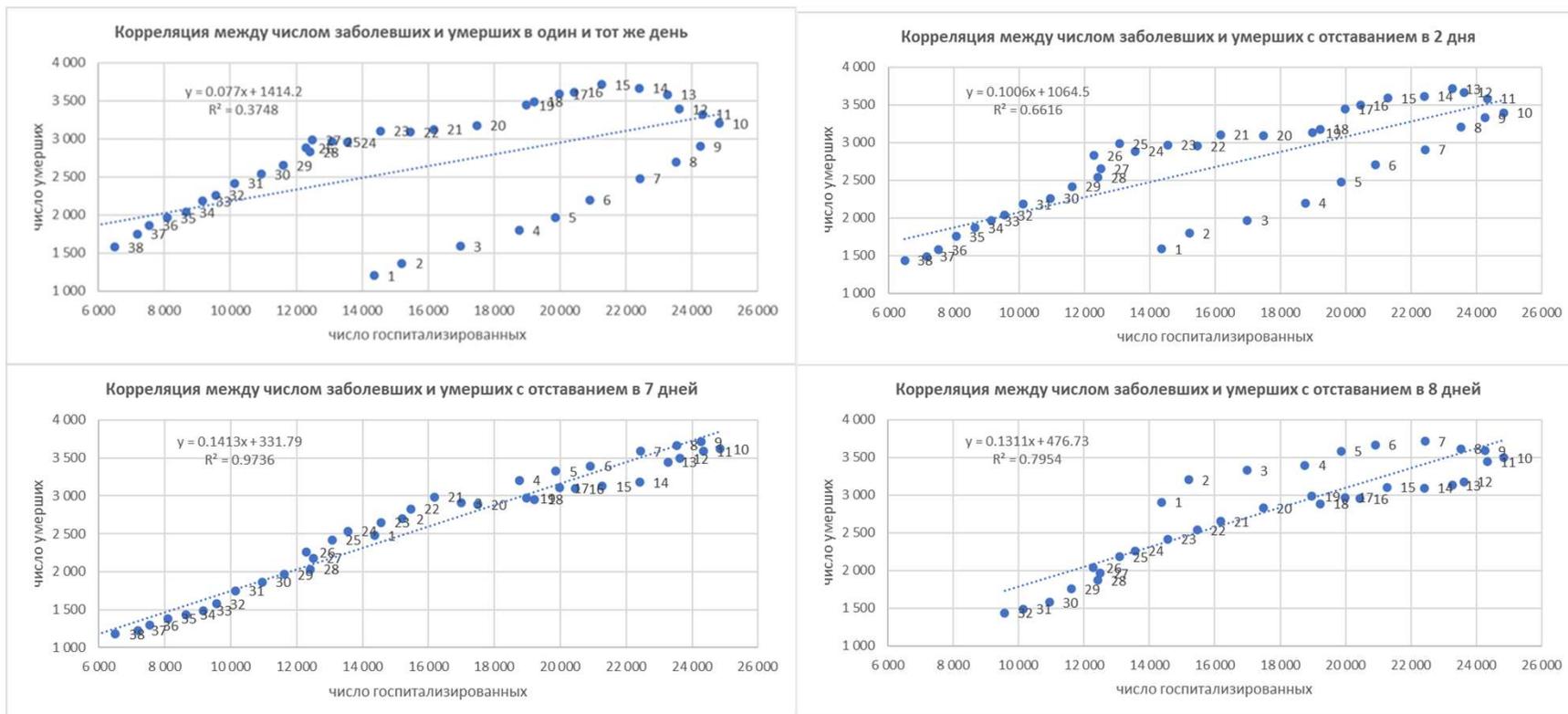
- Было бы правдой, если все заболели в начале эпидемии и болели бы в течение 2-х месяцев
- Вероятно, здесь **отражен эффект запаздывания** (рост числа заболевших быстрее роста числа умерших)
- По достижении пика **показатель стабилизируется**

# Подходы и гипотезы:



## Гипотеза для оценки лага ( $L_n$ ):

число умерших должно коррелировать с числом заболевших (госпитализированных)  $n$  дней назад независимо от этапа эпидемии (подъем, пик или спад).

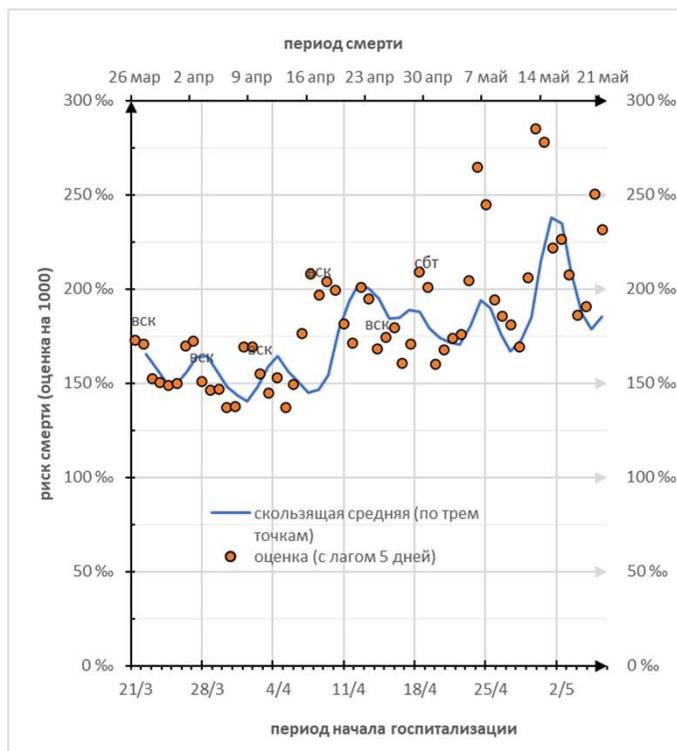
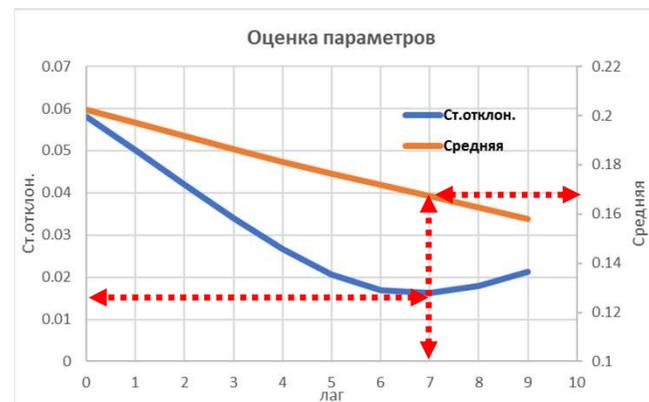


# Прямая оценка риска на основе гипотезы о длительности болезни

## Алгоритм:

- Сглаживание исходных данных по 5 или 7 точкам
- Оценка риска по сглаженным данным
- Сглаживание риска по 3 точкам
- Оценка
  - дисперсии (min 0.016) и
  - центральной тенденции (0.17) сглаженного риска

(метод максимального правдоподобия?)

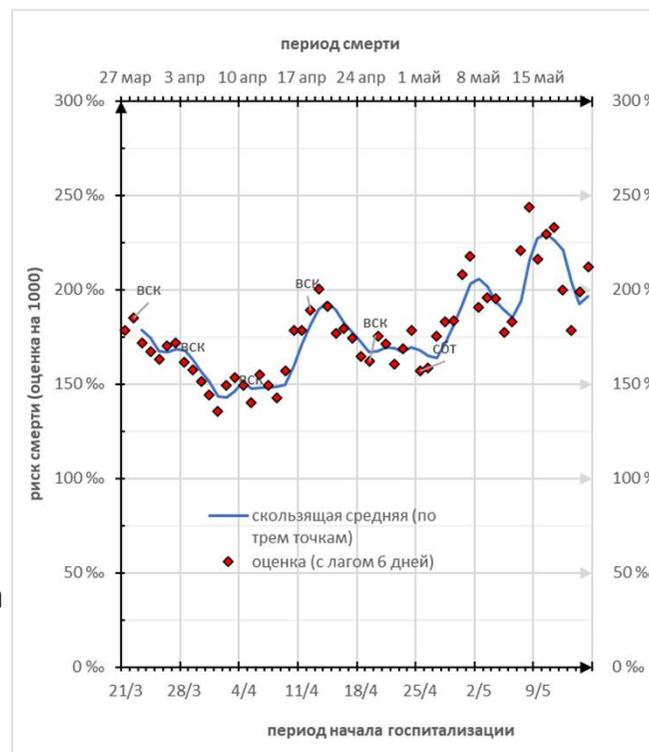


## Комментарии:

Не удалось погасить ни хаотичные, ни «сезонные» колебания

Проявляется тенденция к повышению на спаде эпидемии (изменение лагов?)

**Оценка риска составляет  $17 \pm 1,6\%$ , что близко к отношению умерших к заболевшим по данным на вторую декаду мая.**



## Оценка риска на основе отношения числа умерших к числу выздоровевших в когорте

### Исходные условия (предположения):

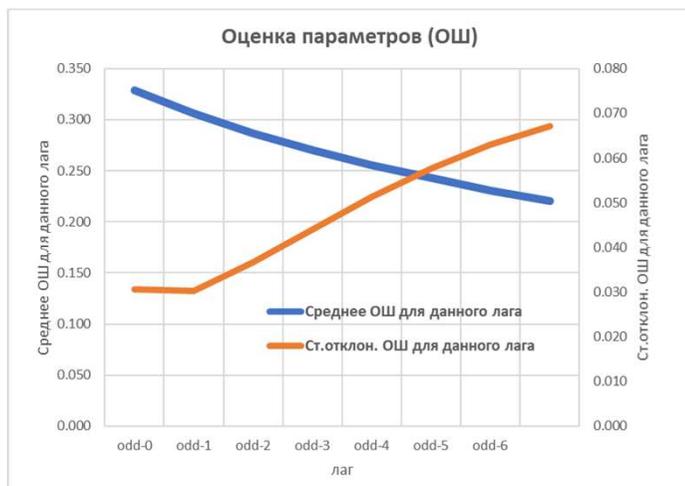
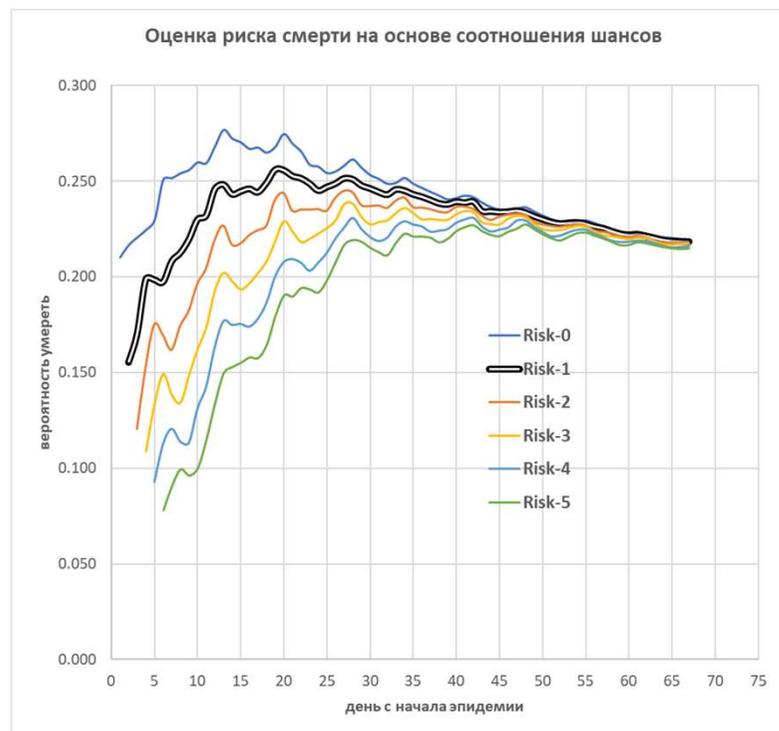
Выздоровевшие к дню  $j$  принадлежат к той же когорте заболевших, что и умершие в день  $(j - L)$ , или  $L$  дней назад.

Сглаживание обеспечивается за счет аккумуляции наблюдений по мере развития эпидемии

Отношение шансов, сложившееся ко дню  $j$ , или  $O_j$  рассчитывается по формуле:

$$O_j = \frac{\sum_{k=1}^{j-L} D_k}{\sum_{i=1}^j C_i}$$

$j$  – день с начала эпидемии  
 $D_k$  – число умерших в стационаре в день  $k$   
 $P_j^H$  – число находящихся в госпиталях  
 $C_i$  – число вернувшихся домой в день  $i$   
 $L$  – лаг или разница в длительности заболевания при альтернативных исходах



### Комментарии:

Не удалось получить стационарной динамики

Видна конверсия к уровню 21% к концу эпидемии и понижающая тенденция

Среднее значение риска где-то **22±2%**, что заметно выше прямой оценки риска

## Выводы

- Доступные национальные данные позволяют оценить общую картину динамики эпидемии коронавируса во Франции
- В то же время для детального статистического анализа эти данные использовать затруднительно, поскольку, по всей видимости, их динамика подвержена очень сильному влиянию административной практики лечебных учреждений
- Косвенные оценки показывают очень высокий (близкий к 20%) уровень летальности заболеваний, вызванных (обусловленных) коронавирусной инфекцией
- Не исключено, что столь высокий уровень летальности зависит от особенностей регистрации первичного диагноза и эффективности сбора информации вне госпитального звена национальной системы здравоохранения
- В любом случае, смертность от коронавируса во Франции кажется очень высокой и вряд ли можно предположить, что детализация и улучшение исходной информации приведут к существенной коррекции предварительных оценок.

**СПАСИБО  
за внимание!**