

### ПОДВОДЯ ИТОГИ: COVID-19 В РОССИИ, США И ВЕЛИКОБРИТАНИИ

Юрий Благовещенский

Основным материалом для исследования динамики распространения COVID-19 являются суточные данные о числе инфицированных  $n(t)$  в той или иной стране. Естественно, что это далеко не полная информация об эпидемии: практически во всех центрах сбора статистических данных имеются сведения о числе умерших и о числе выздоровевших, а в отдельных источниках есть данные о тестировании населения<sup>1</sup>. Конечно, для полноты картины хотелось бы знать и о состоянии медицины в каждой из стран, и многое-многое другое, но это уже вне круга проблем, которые решает статистический анализ данных.

Нашим основным источником данных по России, США и Великобритании (GB) был сайт<sup>2</sup> со следующими логотипом и электронным адресом:

 **КОРОНАВИРУС** и <https://coronavirus-monitor.info/>.

В серии наших исследований с условным названием «прогнозы» (см. таблицу 1) мы изучали лишь динамику числа инфицированных, используя наряду с  $n(t)$  логарифмы общего числа людей  $N(t)$ , зарегистрированных на момент  $t$  как зараженных этим вирусом, и разные вычисляемые характеристики динамических рядов  $n(t)$  и  $N(t)$ .

Таблица 1. Список вывешенных на сайте фонда ИНДЕМ материалов с прогнозами о распространении COVID-19 с середины апреля по 1 июня включительно.

№ п/п	Дата	Гиперссылки доступа к соответствующим материалам
1	16.04.20	<a href="https://indem.ru/coronavirus/Prognoz-USA.pdf">https://indem.ru/coronavirus/Prognoz-USA.pdf</a>
2	17.04.20	<a href="https://indem.ru/coronavirus/Prognoz-dla-RF-using-GB.pdf">https://indem.ru/coronavirus/Prognoz-dla-RF-using-GB.pdf</a>
3	20.04.20	<a href="https://indem.ru/coronavirus/issled-pandem-a-19-1.pdf">https://indem.ru/coronavirus/issled-pandem-a-19-1.pdf</a>
4	21.04.20	<a href="https://indem.ru/coronavirus/issled-pandem-4.pdf">https://indem.ru/coronavirus/issled-pandem-4.pdf</a>
5	21.04.20	<a href="https://indem.ru/coronavirus/comment-prognoz-RF.pdf">https://indem.ru/coronavirus/comment-prognoz-RF.pdf</a>
6	01.05.20	<a href="https://indem.ru/coronavirus/issled-pandem-1may.pdf">https://indem.ru/coronavirus/issled-pandem-1may.pdf</a>
7	05.05.20	<a href="https://indem.ru/coronavirus/issled-pandem-6.pdf">https://indem.ru/coronavirus/issled-pandem-6.pdf</a>
8	15.05.20	<a href="https://indem.ru/coronavirus/chto-stalo-prognoz-ross1.pdf">https://indem.ru/coronavirus/chto-stalo-prognoz-ross1.pdf</a>
9	18.05.20	<a href="https://indem.ru/coronavirus/moskva-kak-mnogo.pdf">https://indem.ru/coronavirus/moskva-kak-mnogo.pdf</a>
10	01.06.20	<a href="https://indem.ru/coronavirus/Rossia-o-chem-govorit-statistika.pdf">https://indem.ru/coronavirus/Rossia-o-chem-govorit-statistika.pdf</a>

Этот материал – заключительное исследование представленной в таблице 1 серии. В сущности, на сегодня динамика эпидемии COVID-19 в России, США и Великобритании стала мало сопоставимой: в США явно протекает вторая волна заражений, Россия застряла на уровнях в 6-7тысяч новых инфицированных в сутки, а Великобритания добралась до фоновых значений в окрестности полутысячи и находится в неустойчивом равновесии.

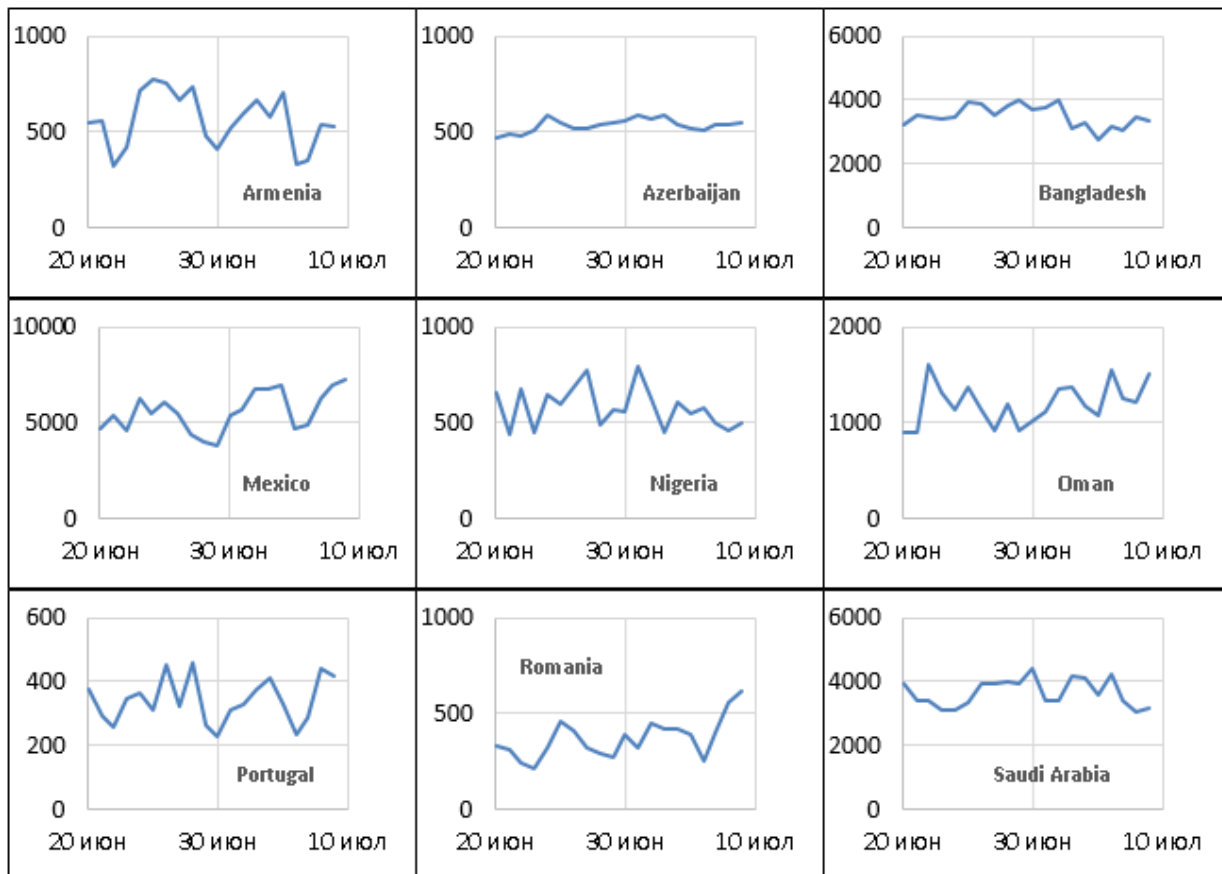
Мы полагаем, что фоновое состояние, когда постоянно возникают и затухают пусть небольшие, но опасные очаги распространения коронавируса SARS-CoV-2 среди населения, является тем наилучшим, которого могут достичь отдельные территории, пока не будет

<sup>1</sup> См., например, <https://ourworldindata.org/coronavirus-testing#our-checklist-for-covid-19-testing-data>.

<sup>2</sup> Приводимая на сайте статистика по России предоставляется Роспотребнадзором РФ. Статистика по странам мира формируется на основании данных от 18 официальных мировых статистических агентств. Обновление статистики происходит 1 раз в 30 минут. Два важных источника информации, портал открытых данных ЕС (<https://data.europa.eu/euodp/data/dataset/covid-19-coronavirus-data>) и данные, собираемые университетом Джонса Хопкинса (<https://coronavirus.jhu.edu/>), обсуждаются с точки зрения их достоинств и недостатков в <https://indem.ru/coronavirus/Univ-Hopkinsa-istochniki-dannyh.pdf>.

действенной против COVID-19 вакцины.

Исходя из этой гипотезы, мы исследовали динамику развития эпидемии от её истоков до 17 июля включительно с упором на последние полтора месяца, пытаюсь в разнообразных графиках разглядеть возможное будущее. Вообще говоря, флуктуация в некотором устойчивом интервале числа новых инфицированных за одни сутки является, на наш взгляд, очень распространенной ситуацией, о чем свидетельствует рисунок 1.



**Рис.1.** Данные суточного прироста числа инфицированных в 9-ти странах в период с 20 июня по 9 июля включительно (динамика).

Конечно, подобная ситуация наблюдается далеко не во всех странах, например, в этот же период возрастающую динамику суточного числа инфицированных имеют Алжир, Болгария, Индия, Израиль, Южная Африка. Естественно, что одновременно имеются страны, в которых число инфицированных в это же время снижается, так что динамические ряды разных государств представляют самое широкое многообразие траекторий.

Из всего этого напрашивается, по крайней мере, один вывод: пока нет никаких моделей, которые могли бы хоть как-то спрогнозировать подобные «хвосты» динамических рядов из предыдущих наблюдений. Так что выявление разных особенностей в динамике развития COVID-19, условий, при которых они появляются, и характеристик, по которым эти особенности можно обнаружить, остается основной задачей анализа COVID-данных<sup>3</sup>.

Первый прогноз для России мы делали, сравнивая её динамику с динамикой Великобритании на той же стадии развития эпидемии (см. материал №2 в таблице 1). Было отмечено их поразительное сходство, а именно, данные по России за 10.03-16.04 (38 наблюдений) и по GB за 29.02-06.04 (тоже 38 наблюдений) обладают высочайшей корреляцией:

<sup>3</sup> См. <https://indem.ru/coronavirus/Stat-problems-of-COVID-19.pdf>: материалы доклада, сделанного 27 мая 2020 года на семинаре «Прикладная статистика и моделирование реальных процессов» в ЦЭМИ.

и для самих значений  $N(t)$ , и для их логарифмов коэффициент  $r > 0,995$ .

Оказывается, что сходство динамик происходит и на падающих траекториях суточного числа инфицированных (рисунок 2).

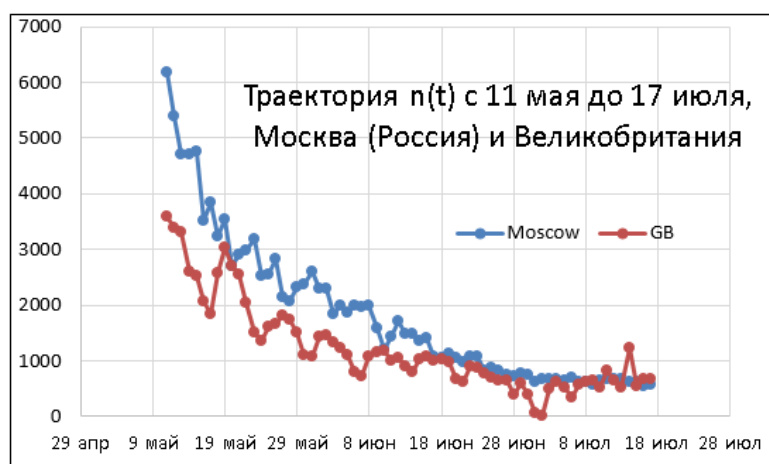


Рис.2. Кривые суточного числа инфицированных в Москве и GB с 11 мая по 17 июля.

Внешняя похожесть траекторий на рисунке 1 бросается в глаза, но и статистический показатель – коэффициент корреляции – подтверждает это:  $r > 0,91$ . Причем, 30-дневные «хвосты» этих кривых были бы почти совпадающими, если бы не «провал» 2 июля и неожиданный всплеск 14 июля в 1240 инфицированных, вдвое превысивший средний уровень.

Остановимся на «провале» 2 июля чуть подробнее. На сайте <https://covid.observer/>, представляющем данные из Центра системных наук и инженерии ун-та Дж. Хопкинса, мы находим в разделе «New Confirmed Cases», что число достоверно подтвержденных случаев 2 июля равно –29726 (минус 29 тысяч 726 инфицированных). Другими словами, в Великобритании из накопленного числа инфицированных за все время эпидемии 2 июля изъяли около 30 тысяч, причем, не удалось найти никаких сведений о причинах, по которым это было сделано. В то же самое время на исходном сайте <https://coronavirus.jhu.edu/map.html> ун-та Хопкинса таких данных нет, правда, снижение числа инфицированных почти до нуля 1-2 июля имеется. В итоге мы решили использовать данные ун-та Хопкинса.

В исследовании <https://indem.ru/coronavirus/exp-faza-COVID-19-bun.pdf>, посвященном экспоненциальной фазе пандемии в разных странах, Великобритания и Москва в качестве самостоятельного объекта были включены в число подробно анализируемых стран. Упор был на изучении экспоненциальной фазы эпидемии, которая в Англии длилась более месяца, начавшись 2-3 марта, а в Москве – около месяца (с 26 марта по 23 апреля).

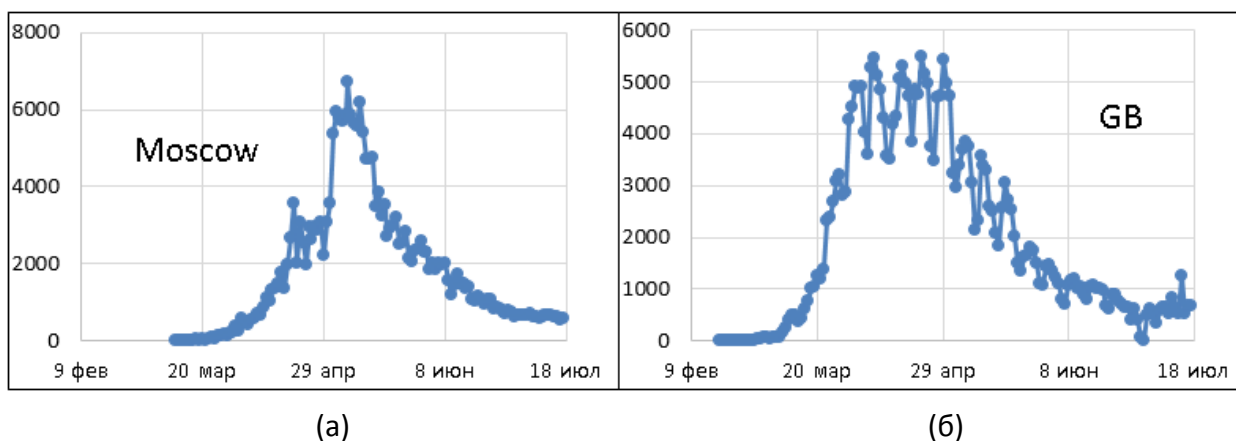
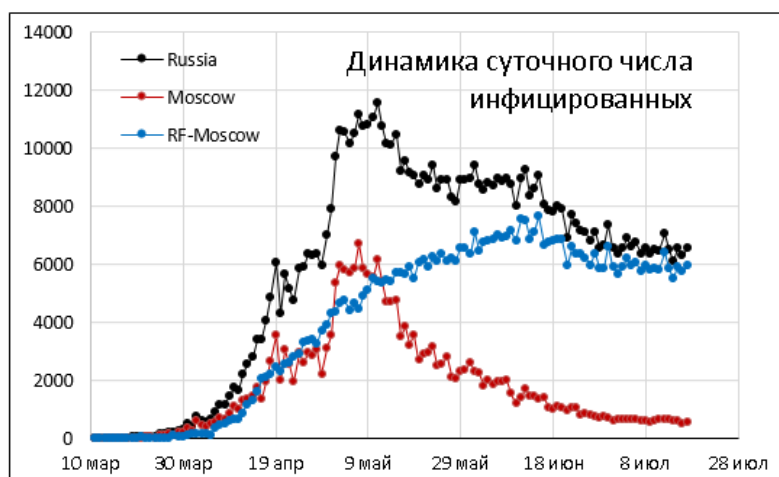


Рис.3. Динамика суточного числа инфицированных: (а) Москва; (б) Великобритания.

Диаграммы (а) и (б) на рисунке 3 различаются в главном всего двумя-тремя особенностями.

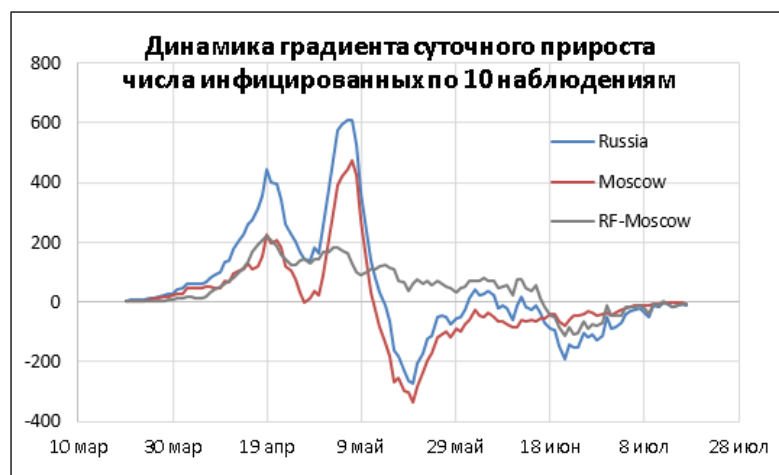
Это, во-первых, период флуктуаций в Москве с 21-30 апреля в интервале 2-3 тысяч инфицированных, а во-вторых, то, что флуктуации в окрестности максимума в GB длятся почти весь апрель, а в Москве всего лишь 10 дней (3-12 мая). Наконец, можно отметить, что размер колебаний в хвостовой части (после 11 мая) существенно больше в Великобритании. Но эти различия почти наверняка вызваны не сутью эпидемии или страны, а действиями органов власти, что следует из указанного выше исследования.

Важно отметить, что Москва на начальном этапе распространения COVID-19 представляла как бы всю Россию (до 6 апреля почти 80% новых инфицированных в России давала Москва), тогда как с 11 мая её доля стала резко падать и после 16 июня стала составлять меньше 15% (на самом «хвосте» меньше 10%, см. рисунок 4).



**Рис.4.** Динамика суточного числа инфицированных: Россия, Москва и Россия без Москвы.

Из приведенной диаграммы трудно что-либо спрогнозировать даже на ближайшую пару недель, определить, как поведет себя кривая числа инфицированных SARS-CoV-2 в России, будь это Россия в целом или Россия без Москвы, или же сама Москва. Особенно хорошо это видно на следующей диаграмме (рисунок 5).



**Рис.5.** По оси ординат откладываются коэффициенты углов наклона прямой, аппроксимирующей градиент траектории суточного числа инфицированных по 10 дням, предшествующим расчетной дате (включая её саму).

Если  $t$  – расчетная дата, то по точкам  $Q(s; t) = (t - s + 1, n(t - s + 1))$ ,  $1 \leq s \leq 10$ , проводится методом наименьших квадратов прямая  $y(s; t) = a(t) \cdot s + b(t)$ , и её коэффициент  $a(t)$  принимается за градиент траектории суточного числа инфицированных в точке  $t$ . Эти градиенты и представлены в кривых на диаграмме рисунка 5. Они с очевидностью выходят

на нуль для всех трех объектов, будучи до этого в отрицательной зоне. А это означает, что падение числа инфицированных прекращается и выходит на некоторое плато.

На этой диаграмме, кстати, отчетливо видно, что градиенты Москвы и России в целом ведут себя чрезвычайно похожим образом практически до 7 июня, после чего градиент России как бы подражает динамике градиентов России регионов (без Москвы), что мы и отмечали, оправдывая проводимое сравнение GB с Москвой, а не с Россией.

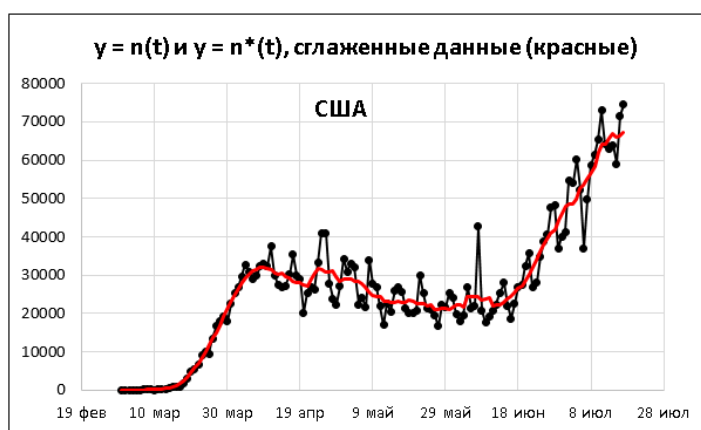
Подобные диаграммы с динамикой градиентов для США и Великобритании тоже весьма показательны (рисунок 6), хотя и не сулят пока ничего хорошего.



**Рис.6.** Градиенты динамических рядов, подобные представленным на кривых рисунка 5, для (а) Великобритании и (б) США. Стрелки объясняются в тексте.

Первое, что бросается в глаза, это – очень большие колебания у кривой градиентов в США. Но это обусловлено ещё и тем, расчетные значения  $a(t)$  не нормированы и пропорциональны (в среднем) масштабу числа инфицированных за одни сутки в расчетный период. А вот тенденции в динамике градиентов в Великобритании (Рис.6(а), май-июль) и в США (Рис.6(б), июнь-июль), помеченные красными стрелками, связаны с политикой ослабления принятых ранее мер по борьбе с COVID-19.

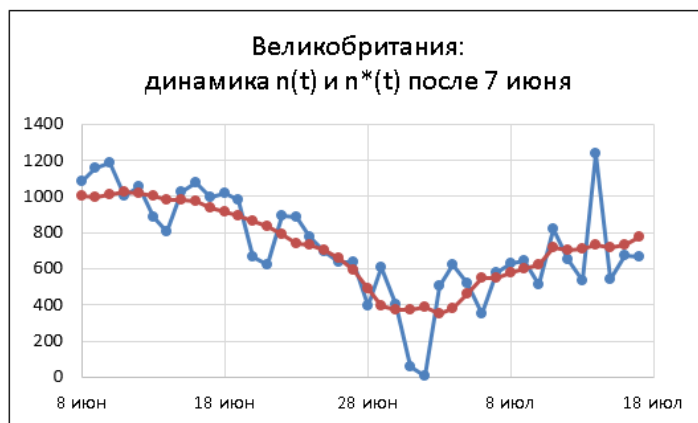
В США дополнительным катализатором роста числа инфицированных стали массовые выступления, начавшиеся в конце мая под лозунгами общественного движения Black Lives Matter и охватившие почти всю Америку. Это, на наш взгляд, стало главной причиной второй волны эпидемии в США, о чем ярко свидетельствует рисунок 7.



**Рис.7.** Динамика суточных данных  $n(t)$  о числе инфицированных в США. Красным цветом выделена траектория средних по 7 дням  $n^*(t) = \frac{1}{7} \sum_{i=-3}^{i=3} n(t+i)$ .

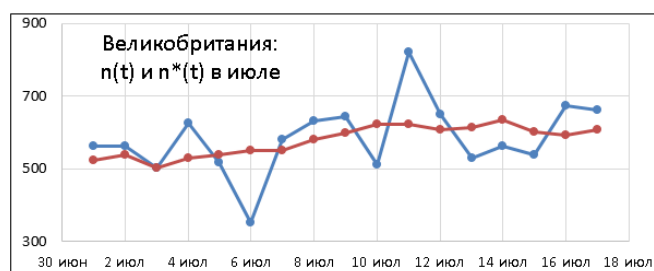
В Великобритании (в отличие от США) помеченная на Рис.6(а) красной стрелкой тенденция не столь очевидна и, возможно, не отражает реальное положение дел, являясь всего лишь следствием ошибок в статистических данных. Для оценки этой возможности посмотрим на

динамику суточного числа инфицированных в другом масштабе (рисунок 8).



**Рис.8.** Динамика суточных данных о числе инфицированных 08.06-17.07: синие точки представляют исходные данные  $n(t)$ , а красные – сглаженные  $n^*(t)$ , см. рисунок 7.

Мы видим, что траектория сглаженных данных о числе инфицированных в Великобритании растет в июле примерно с 400 до 800, причем, это происходит не только за счет провала вниз и скачка за 1200 зараженных. Например, заменив эти «экстремальные» значения на среднее 17 наблюдений в июле ( $\approx 562$ ), мы все равно получим рост, хотя и небольшой, в пределах от 500 до 600 в среднем (рисунок 9).



**Рис.9.** Скорректированные июльские данные суточного числа инфицированных в Великобритании: исходные и сглаженные.

Подводя итоги всему этому анализу, мы можем воспользоваться дополнительной информацией за 18-22 июля (время подготовки этого текста), содержащей пять новых наблюдений числа инфицированных в рассматриваемых странах. Во-первых, для Великобритании, к сожалению, отмеченная красной стрелкой тенденция на Рис.6(а) получает дополнительные аргументы «За» (в последние три дня число заражений за сутки превысило 800 человек). В США наблюдается «небольшое» снижение на 15-20 тысяч инфицированных в день от максимума в 78 тысяч 16-18 июля. Москва по-прежнему колеблется вокруг цифры в 600 инфицированных за день. У России в целом, как и у России без Москвы есть слабый намек на снижение, но всё равно планка в 6 тысяч для России остается актуальной.

Из проведенного анализа мы можем сделать пока один важный вывод: в силу разных естественных и социально-политических причин динамика суточного числа инфицированных SARS-CoV-2 во многих странах «застраивает» на некотором уровне, на преодоление которого необходимо либо достаточно длительное время, либо включения новых мер борьбы с COVID-19. И с этого уровня вполне может начинаться вторая волна эпидемии, так что плато в динамике числа инфицированных требует особого внимания и анализа.

[\*\*Переход к COVID-19: АНАЛИЗ И ПРОГНОЗ\*\*](#)