

## МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ЭПИДЕМИЙ: ОПЫТ COVID-19

Асташонок Д. А., Рагозик Е. Е., студенты

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Лобанок Л. В. – старший преподаватель

**Аннотация.** Работа посвящена анализу и сравнению математических моделей распространения инфекционных заболеваний на примере COVID-19. Рассматриваются модели SI, SIR и SEIR. Особое внимание уделяется роли инкубационного периода в прогнозировании эпидемического пика. Делается вывод, что для эффективного подавления пандемии решающее значение имеет своевременная корректировка стратегий на основе динамических моделей и данных геолокации.

**Ключевые слова.** Математическое моделирование эпидемий, компартментные модели, модель SI, модель SIR, модель SEIR, базовое репродуктивное число, инкубационный период, COVID-19, эпидемический пик, дифференциальные уравнения, пандемия, прогнозирование распространения инфекции, коллективный иммунитет, геолокационные данные.

В декабре 2019 года в китайском городе Ухань был зафиксирован первый случай заболевания, вызванного неизвестным до тех пор вирусом SARS-CoV-2. Менее чем за три месяца вирус пересёк границы более сотни государств. 11 марта 2020 года Всемирная организация здравоохранения объявила пандемию COVID-19 - событие, которое в XXI веке не имело аналогов ни по масштабу, ни по скорости распространения. К середине 2023 года число подтверждённых случаев заражения превысило 770 миллионов человек, а число смертей - 7 миллионов. Человечество оказалось перед острой необходимостью понять: как развивается эпидемия, когда наступит пик, что произойдёт при введении карантина - и можно ли вообще это предсказать? Ответ на эти вопросы лежал не только в медицине и вирусологии, но и в математике. Именно математическое моделирование оказалось тем инструментом, который позволил правительствам и учёным хотя бы приблизительно видеть будущее эпидемии

В математике существуют специальные компартментные модели - системы дифференциальных уравнений, разбивающие популяцию на группы (компарменты) в зависимости от эпидемиологического статуса человека. Эти модели позволяют решать важные задачи: просчитать возможное количество заболевших на единицу времени, определить «критический порог» - то число контактов и заражений, при котором эпидемия переходит в неуправляемую фазу, а также оценить эффект противоэпидемических мер ещё до их введения.

Центральным понятием всех эпидемиологических моделей является базовое репродуктивное число  $R_0$  - среднее число людей, которых один заражённый инфицирует в полностью восприимчивой популяции (среди лиц, не имеющих антител). Если  $R_0 > 1$ , эпидемия распространяется; если  $R_0 < 1$ , она угасает.

Для COVID-19 в начале пандемии  $R_0$  оценивался в диапазоне от 2,4 до 6,5 в зависимости от варианта вируса и региона. Математическая необходимость снижения репродуктивного числа до значений меньше единицы выступала основой для вводимых ограничений. Именно этот научный аргумент давал принимаемым мерам научное обоснование и оправдывал их в глазах общества.

В рамках работы рассмотрим модели: SI, SIR, SEIR. Для каждой модели проводим анализ устойчивости положений равновесия, вычислим базовое репродуктивное число  $R_0$  (порог эпидемической безопасности), результаты сопоставим с реальными данными пандемии COVID-19. Простейшая компартментная модель - SI (модель без выздоровления) - делит популяцию на восприимчивых (S, Susceptible) и инфицированных (I, Infected). Она была формализована ещё в начале XX века и описывает заболевания, при которых иммунитет не формируется (характерно для ряда хронических заболеваний). Система уравнений имеет вид:

$$\frac{dS}{dt} = \frac{-\beta \cdot S \cdot I}{N},$$

$$\frac{dI}{dt} = \frac{\beta \cdot S \cdot I}{N},$$

где  $\beta$  — скорость передачи инфекции,  $N = S + I$  — общая численность популяции. Правая часть уравнения для I описывает скорость «перетока» людей из восприимчивых в инфицированных. Решение этой системы — классическая сигмоидная кривая: сначала инфекция распространяется медленно, затем — лавинообразно, и в итоге весь S переходит в I. Для COVID-19 модель SI непригодна, поскольку не учитывает ни выздоровление, ни инкубационный период (принципиально важные явления для SARS-CoV-2). Представление модели SI приведено на рисунке 1.

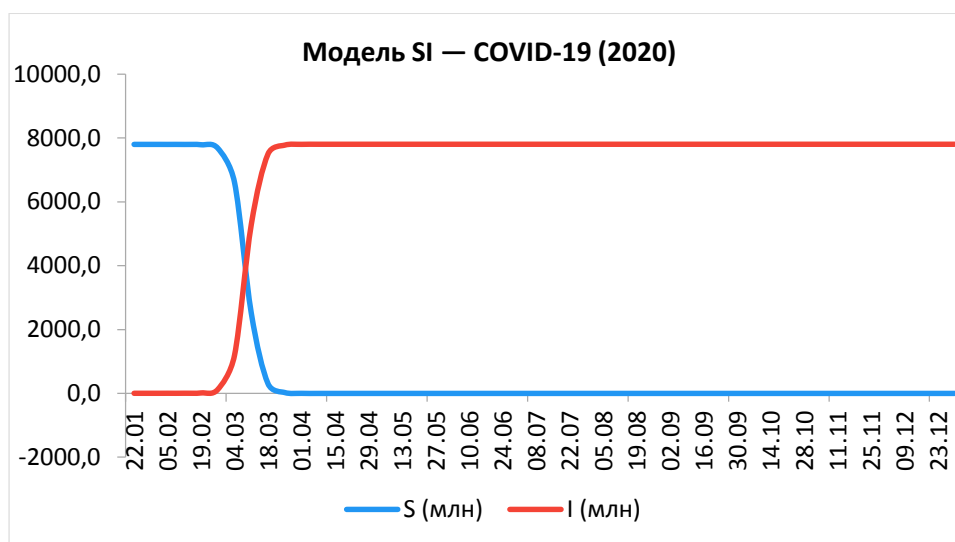


Рисунок 1 — Динамика модели SI (COVID-19, 2020 г.)

На графике видна классическая S-образная (сигмоидная) кривая: к концу 2020 г. практически всё население теоретически перешло бы в категорию инфицированных. Модель SI не учитывает выздоровление, поэтому она завышает масштаб эпидемии и применима лишь для демонстрации базового механизма передачи инфекции. Параметр  $\beta = 0,35$  соответствует оценкам начального темпа распространения COVID-19.

В 1927 году шотландские учёные Уильям Кермак и Андерсон МакКендрик предложили *модель SIR*, ставшую основой для всей математической эпидемиологии. Она добавляет третий компармент - R (Recovered/Removed) - выздоровевших, получивших иммунитет. Система дифференциальных уравнений:

$$\frac{dS}{dt} = \frac{-\beta \cdot S \cdot I}{N},$$

$$\frac{dI}{dt} = \frac{\beta \cdot S \cdot I}{N - \gamma \cdot I},$$

$$\frac{dR}{dt} = \gamma \cdot I,$$

где  $\gamma$  — скорость выздоровления (обратная к среднему времени болезни). Из этой системы выводится ключевой параметр — базовое репродуктивное число:

$$R_0 = \frac{\beta}{\gamma},$$

При  $R_0 > 1$  эпидемия развивается экспоненциально (один больной заражает в среднем больше одного человека), пока в популяции достаточно восприимчивых к вирусу людей. По мере распространения инфекции доля восприимчивых неизбежно сокращается, так как переболевшие переходят в категорию иммунных (выздоровевших, R).

В данной модели используется понятие эпидемического пика (критическая точка в динамике распространения инфекции, когда количество одновременно болеющих людей достигает своего максимума). Эпидемический пик наступает тогда, когда эффективное репродуктивное число  $Re = 1$ . Пик происходит из-за сокращения количества восприимчивых к вирусу людей (так как переболевшие переходят в категорию иммунных (выздоровевших, R)). После прохождения этой точки каждый носитель успевает передать вирус в среднем менее, чем одному человеку, и волна затухает [1].

Для COVID-19 с  $R_0 \approx 3$  это означает, что ~67% популяции должны обладать иммунитетом, чтобы остановить распространение. В начале пандемии классическая модель SIR стала фундаментом для прогнозов, так как помогла понять, когда наступит пик и справятся ли больницы с нагрузкой. Однако SIR обладает существенным ограничением: она предполагает, что человек становится заразным мгновенно после контакта с вирусом. Представление модели SIR приведено на рисунке 2 [2].

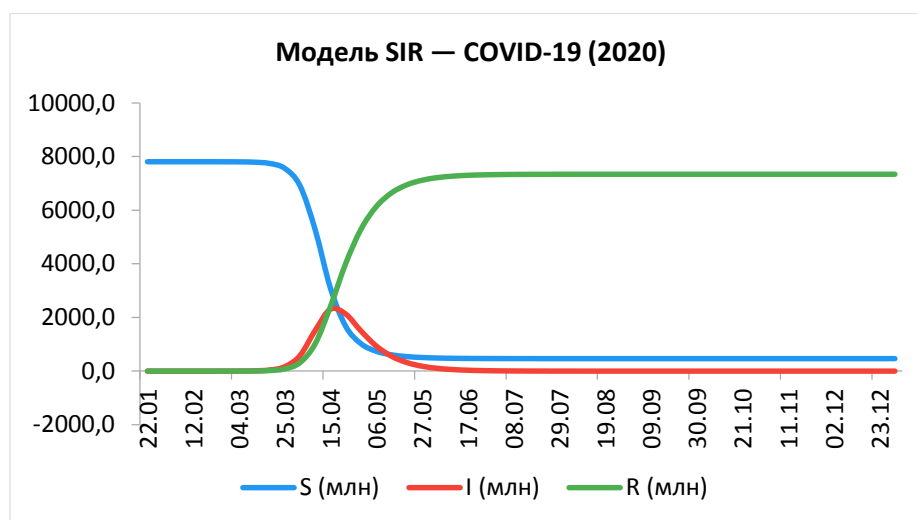


Рисунок 2 — Динамика модели SIR (COVID-19, 2020 г.)

График модели SIR демонстрирует эпидемический пик (кривая I) в весенние месяцы 2020 г., после чего число активных случаев начинает снижаться за счёт перехода переболевших в группу R. При  $R_0 = \beta/\gamma = 3,0$  для остановки эпидемии необходим иммунитет примерно у 67 % населения.

Для более точного моделирования COVID-19 ученые перешли к модели SEIR, которая предназначена для вирусов с выраженным инкубационным периодом, где добавлен четвёртый компартмент E (Exposed) - люди, уже заражённые, но ещё не заразные. Именно это объясняет, почему так трудно контролировать COVID-19: человек уже несёт вирус, но ещё не знает об этом. Пока официальная статистика показывала лишь первых больных, в фазе «E» уже находились тысячи скрытых носителей, которые должны были стать заразными через несколько дней. Модель SEIR наглядно объяснила властям, почему нельзя медлить с карантином: если ждать явного всплеска болезни, будет уже поздно, так как невидимая волна зараженных уже запущена. Модель SEIR описывается системой:

$$\frac{dS}{dt} = \frac{-\beta \cdot S \cdot I}{N},$$

$$\frac{dE}{dt} = \frac{\beta \cdot S \cdot I}{N - \sigma \cdot E},$$

$$\frac{dI}{dt} = \sigma \cdot E - \gamma \cdot I,$$

$$\frac{dR}{dt} = \gamma \cdot I,$$

где  $\sigma$  — скорость выхода из инкубационного периода ( $1/\sigma$  — средняя длина инкубационного периода). Базовое репродуктивное число в модели SEIR принимает вид:

$$R_0 = \frac{\beta \cdot \sigma}{[(\sigma + \mu)(\gamma + \mu)]},$$

где  $\mu$  — естественная смертность. Метод матрицы следующего поколения (Next Generation Matrix) позволяет формально вычислить  $R_0$  как спектральный радиус матрицы  $FV^{-1}$ , где F — матрица новых заражений, V — матрица переходов. Для базовой SEIR без смертности:

$$R_0 = \rho(FV^{-1}) = \frac{\beta \sigma}{[(\sigma)(\gamma)] = \beta/\gamma},$$

Если  $R_0 < 1$ , вирус не может поддерживать себя в популяции. Больной передает инфекцию меньше, чем одному человеку, цепочки быстро обрываются, и система стремится к бессимптомному равновесию — периоду, когда болезнь исчезает сама собой. При  $R_0 > 1$  ситуация кардинально меняется: эпидемия не просто вспыхивает, а переходит в эндемическое равновесие (состояние, при котором вирус не исчезает, но и не растет бесконечно). Устанавливается баланс населения: количество новых заражений совпадает с количеством выздоравливающих или умирающих (в медицине их называют «выбывшими из цепочки»). В итоге в популяции всегда остается стабильный процент инфицированных.

Принципиальное преимущество SEIR перед SIR состоит в том, что она учитывает задержку между моментом заражения и началом заразности. Именно этот период инкубации был ключевой

особенностью COVID-19, делавшей его столь опасным: человек распространял вирус ещё до появления симптомов.

Модель SEIR стала основой для расширенных версий, применявшихся при COVID-19: SEIRD (с учётом смертности), SEIHR (с учётом госпитализации), SEIQR (с карантинными мерами) и стохастических вариантов, учитывающих случайную природу передачи инфекции на малых популяциях. Схематическое представление модели SIER приведено на рисунке 2.

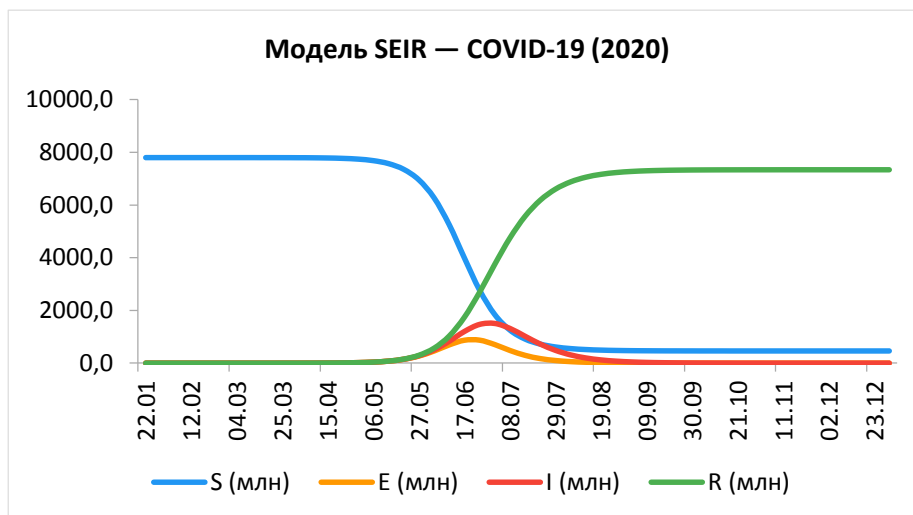


Рисунок 3 — Динамика модели SEIR (COVID-19, 2020 г.)

Модель SEIR добавляет кривую E (инкубационный период,  $\sigma^{-1} \approx 5,1$  дня), которая опережает рост числа явных больных (I). Это наглядно объясняет скрытое распространение COVID-19: в момент регистрации первых тысяч случаев десятки миллионов человек уже находились в бессимптомной фазе. Пик кривой I ниже, чем в SIR, так как инкубационный период замедляет нарастание заразности. SEIR лучше всего из рассмотренных моделей описывает реальную динамику первой волны пандемии. Далее в таблице 1 представлено сравнение основных показателей всех изученных выше моделей.

Таблица 1 — Сравнение компартментных моделей

Критерий	SI	SIR	SEIR
Компартменты	S (восприимчивые) I (инфицированные)	S (восприимчивые) I (инфицированные) R (выздоровевшие)	S (восприимчивые) E (уже заражённые, но не заразные) I (инфицированные) R (выздоровевшие)
Уравнений	2	3	4
Учёт иммунитета	Нет	Да	Да
Инкубационный период	Нет	Нет	Да
$R_0$ (COVID-19)	Не применима	$\approx 2,4-3,0$ ~67% популяции должны обладать иммунитетом, чтобы остановить распространение	$\approx 2,4-6,5$ в фазе «E» уже находятся тысячи скрытых носителей, которые должны стать заразными через несколько дней
Возможность расширения	Ограничена	Умеренная	Высокая
Учёт смертности	Нет	Нет	Да (в расширении SEIRD)
Применимость к COVID-19 по волнам	Нет	Только 1-я волна	1-я и частично 2-я волны

Пандемия COVID-19 стала беспрецедентным испытанием для математических моделей — причём не только в плане их точности, но и в плане управленческих последствий их применения. В январе-марте

2020 года группа Imperial College London во главе с Нилом Фергюсоном применила расширенную SEIR-модель с пространственной структурой и параметрами:  $R_0 \approx 2,4$ ; инкубационный период  $\sigma^{-1} \approx 5,1$  дня; период заразности  $\gamma^{-1} \approx 2,9$  дня. Прогноз без вмешательства: ~510 000 смертей в Великобритании и ~2,2 млн в США. Именно этот прогноз стал основанием для введения жёстких локдаунов [3].

Параллельно в России группа учёных из ИПМ РАН применяла модифицированную SEIR-модель для оценки динамики пандемии в Москве. Калиброванные параметры:  $R_0 \approx 3,0$  в начале пандемии; снижение до  $R_0 \approx 1,1-1,3$  после введения ограничений. Модель корректно предсказала первый пик заболеваемости в Москве (май 2020), однако ошиблась в прогнозе второй волны осенью того же года.

В период COVID-19 были совершенно достаточно грубые математические ошибки, в результате которых эпидемия распространилась слишком быстро. В первую очередь, заметим, что игнорировалась неоднородности популяции, так как классические SEIR-модели предполагают однородное смешивание: каждый заражённый с одинаковой вероятностью контактирует с любым восприимчивым. В реальности структура контактов сильно зависит от возраста, профессии и места жительства. COVID-19 показал: пожилые люди умирали в десятки раз чаще молодых, а модели этого не учитывали. В первые месяцы пандемии реальная доля инфицированных была неизвестна из-за недостаточного тестирования. Модели калибровались по смертям, но смертность сама зависела от перегруженности больниц. Все модели предполагали, что люди адекватно реагируют на риск и соблюдают ограничения. Реальное поведение людей оказалось значительно сложнее: часть населения игнорировала рекомендации, другая часть чрезмерно ограничивала контакты.

Пандемия COVID-19 трансформировала классические модели в цифровые инструменты. Современное математическое моделирование развивается по 3 основным направлениям: использование Big Data и нейросетей для мгновенного обновления параметров заразности, переход к агентным моделям для учета личных контактов и распространителей, а также интеграция данных геолокации для прогнозирования перемещения вируса между регионами. Это превращает теоретические формулы в прикладные системы, которые способны учитывать биологию вируса и поведение людей в едином комплексе.

COVID-19 доказал, что математические модели - это не абстракция, а критический инструмент для принятия важных мировых решений. Главный урок пандемии заключается в том, что лучшей математической модели не существует, так как точность прогноза зависит от качества данных и своевременности их применения. Опыт Южной Кореи, Германии, Тайвани (там модели применялись проактивно, что способствовало тому, что эпидемию удавалось контролировать значительно эффективнее) показал: математика не останавливает болезнь сама по себе, но она дает время и понимание, необходимые для спасения миллионов жизней. Будущее эпидемиологии стоит за гибридными системами, сочетающими строгость формул с гибкостью искусственного интеллекта.

**Список использованных источников:**

1. Kermack, W.O. *A contribution to the mathematical theory of epidemics* / W.O. Kermack, A.G. McKendrick // *Proceedings of the Royal Society A*. – 1927. – Vol. 115, no. 772. – P. 700–721.
2. He, S. *SEIR modeling of the COVID-19 and its dynamics* / S. He, Y. Peng, K. Sun // *Nonlinear Dynamics*. – 2020. – Vol. 101, no. 3. – P. 1667–1680.
3. Сонин, К.И. *Моделирование пандемии COVID-19: методы и результаты* / К.И. Сонин // *Российский журнал менеджмента*. – 2020. – Т. 18, № 4. – С. 551–558.