

РГП НА ПХВ «НАЦИОНАЛЬНЫЙ ЦЕНТР ОБЩЕСТВЕННОГО
ЗДРАВООХРАНЕНИЯ» МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ
РЕСПУБЛИКИ КАЗАХСТАН

**Ж.Т. Ерубает, Н.М. Муратова, Л.К. Касабекова, С.Б. Газезова,
Ф.А. Аблимитова, А.М. Балмаганбетов, Н. Кишор,
А.К. Сатылганова, Д.Ж.Мұқажан**

ПРОГНОЗИРОВАНИЕ ЗАБОЛЕВАЕМОСТИ КОРЬЮ

(Методические рекомендации)

Астана, 2023

УДК 616.915
ББК 55.14
П78

Рецензенты:

Текебаев К. О. – к.м.н., директор департамента профилактики инфекционных заболеваний НЦОЗ МЗ РК

Кишор Н. - PhD MS MPH, отдел по глобальной иммунизации, Центр США по контролю и профилактике заболеваний

Авторы:

Ерубает Ж.Т. – MPH, директор оперативного центра по чрезвычайным ситуациям в общественном здравоохранении НЦОЗ МЗ РК

Муратова Н.М. – главный эксперт управления контроля за вакциноуправляемыми инфекциями КСЭК МЗ РК

Касабекова Л.К. – менеджер здравоохранения отдела анализа и качества менеджмента ЦСЭЭ МЦ УДП РК

Газезова С.Б. – врач-эпидемиолог управления профилактики инфекционных и паразитарных заболеваний НПЦСЭЭиМ НЦОЗ МЗ РК

Аблимитова Ф.А. – врач-эпидемиолог управления профилактики инфекционных и паразитарных заболеваний НПЦСЭЭиМ НЦОЗ МЗ РК

Балмаганбетов А.М. – курсант двухгодичной программы прикладной эпидемиологии (FETP)

Кишор Н. – PhD MS MPH, отдел по глобальной иммунизации, Центр США по контролю и профилактике заболеваний

Сатылганова А.К. – главный специалист Оперативного центра по ЧС в области общественного здравоохранения НЦОЗ МЗ РК

Мұқажан Д.Ж. – главный специалист Оперативного центра по ЧС в области общественного здравоохранения НЦОЗ МЗ РК

Прогнозирование заболеваемости корью: Методические рекомендации / *Ерубает Ж.Т., Муратова Н.М. и др.* // Астана: Оперативный центр по чрезвычайным ситуациям в общественном здравоохранении, 2023. – 50 с.

ISBN 978-601-305-553-4

- Настоящие методические рекомендации описывают метод прогнозирования с целью оценки эпидемиологической ситуации и дальнейшего распространения заболеваемости корью в Республике Казахстан. Данный анализ использует EpiNow2, который основан на байесовском подходе латентных переменных с использованием языка вероятностного программирования Stan. Прогнозирование

осуществляется на 14 дней после последней отчетной даты, проводя экстраполяцию тренда и неопределенности из периода, непосредственно предшествующего последней дате, по которой доступны данные. Данный метод прогнозирования может быть использован Министерством здравоохранения для принятия необходимых ответных мер в случае роста количества случаев кори.

Утверждено и разрешено к изданию типографским способом РГП на ПХВ «Национальный научный центр развития здравоохранения имени Салидат Каирбековой» Министерства здравоохранения Республики Казахстан (протокол заседания Департамента развития науки и образования РГП на ПХВ ННЦРЗ им. Салидат Каирбековой) № 429 от «28» декабря 2023 года.

Содержание

Перечень терминов	5
Введение.....	6
Основная часть	8
Методология	8
Ограничения в методологии	10
Пример прогнозирования.....	11
Алгоритм прогнозирования заболеваемости корью	16
1. Установка программы R studio	17
2. Подготовка исходных данных для прогнозирования	17
3. Загрузка и инициирование пакетов	18
4.. Запуск прогнозирования	19
5. Создание отчета на основе результатов прогнозирования.....	24
Заключение	26
Список использованных источников	27
Приложение	28

Перечень терминов

1. Корь – вирусное заболевание, возбудителем которого является вирус из семейства парамиксовирусов. Вирус кори обычно передается через прямой контакт, а также по воздуху, инфицирует слизистую оболочку, а затем распространяется по организму.
2. Заболеваемость корью: Число новых случаев заболевания корью в определенном временном и пространственном интервале.
3. Прогнозирование: Процесс предсказания будущих событий или тенденций на основе анализа исторических данных и различных моделей.
4. Модель прогнозирования : Математическая модель, используемая для прогнозирования заболеваемости корью, которая учитывает различные факторы и параметры, влияющие на распространение болезни.
5. Факторы риска : Эти факторы могут включать наличие тест систем, количество подтвержденных случаев и другие переменные, которые могут повлиять на заболеваемость корью.
6. Сезонность: Регулярные колебания в уровне заболеваемости корью, связанные с временем года, климатическими факторами и другими факторами.
7. Моделирование распространения: Использование математических моделей для изучения и предсказания способов распространения заболевания корью в популяции.
8. Интервенции: Меры и стратегии, направленные на снижение заболеваемости корью, улучшение медицинской инфраструктуры и образования о заболевании.
9. R studio - это интегрированная среда разработки (IDE) для языка программирования R, который широко используется в статистическом анализе, научных исследованиях и прогнозировании. В контексте прогнозирования, R Studio представляет собой мощное программное обеспечение, которое предоставляет исследователям и аналитикам множество инструментов и библиотек для анализа данных, построения статистических моделей и создания прогнозов.
10. Отчетность и документирование: R Studio позволяет создавать отчеты и документацию к вашему анализу, что упрощает коммуникацию результатов и выводов вашей модели прогнозирования.

Введение

Корь – крайне заразное заболевание вирусного происхождения. Вирус из семейства парамиксовирусов является возбудителем этой опасной инфекции. Вирус кори распространяется при кашле и чихании, тесных личных контактах или непосредственном контакте с инфицированными выделениями из носоглотки. Во многих развивающихся странах Африки и Азии, показатели заболеваемости корью все еще остаются высоки. Более 95% летальных случаев от кори происходят в странах с низким доходом на душу населения и проблемами в инфраструктуре системы здравоохранения. Большинство смертельных случаев кори происходит из-за осложнений, связанных с этой болезнью. Чаще всего осложнения развиваются у детей в возрасте до пяти лет или у взрослых людей старше 30 лет. Самые серьезные осложнения включают слепоту, энцефалит (инфекцию, приводящую к отеку головного мозга), тяжелую диарею и связанную с ней дегидратацию, ушные инфекции и тяжелые инфекции дыхательных путей, такие как пневмония. Тяжелое течение кори более вероятно среди плохо питающихся детей младшего возраста, особенно тех, кто испытывает недостаток витамина А, или чья иммунная система ослаблена ВИЧ/СПИДом или другими болезнями. Регулярная противокоревая вакцинация детей в сочетании с кампаниями массовой иммунизации в странах с высокими показателями заболеваемости и смертности являются основными стратегиями общественного здравоохранения, направленными на уменьшение глобальной смертности от кори. Смертоносные вспышки кори могут также причинить ущерб инфраструктуре и службам здравоохранения страны, поэтому крайне важно отслеживать эпидемиологическую ситуацию и принимать своевременные ответные меры по борьбе с корью [1].

С начала 2023 года, по состоянию на 24 августа текущего года, в Республике Казахстан зарегистрировано 3624 подтвержденных случаев кори. 83% из них – дети. В связи с высоким количеством случаев кори в стране появляется острая необходимость в постоянном мониторинге эпидемиологической ситуации, с применением надежных статистических инструментов. С этой целью была разработана математическая модель для прогнозирования заболеваемости корью. Модель использует набор статистических инструментов EpiNow2, который оценивает изменяющееся во времени репродуктивное число, скорость роста и время удвоения. Прогнозирование осуществляется как на национальном, так и на региональном уровнях.

Прогнозирование заболеваемости корью является важным инструментом для оценки риска и разработки стратегий по предотвращению и борьбе с распространением этой инфекции. С учетом глобальных усилий по улучшению вакцинации и медицинской инфраструктуры, а также в свете внезапных эпидемических вспышек, прогнозирование заболеваемости корью становится актуальной задачей для органов здравоохранения и общества в

целом.

В данном исследовании представляется модель прогнозирования заболеваемости корью, основанная на анализе исторических данных, клинических показателей и факторах риска, влияющих на распространение болезни. Наше исследование направлено на создание инструмента, который поможет лучше понимать и прогнозировать динамику заболеваемости корью, а также способствовать разработке эффективных мер по предотвращению распространения этой опасной инфекции.

Прогнозирование включает в себя ряд элементов. Один из них — информация об объекте прогнозирования, раскрывающая его поведение в прошлом и настоящем, а также закономерности этого поведения.

Основной целью данной разработки является оценка эпидемиологической ситуации и прогнозирование дальнейшего распространения кори. Математический подход заключается в использовании имеющихся данных о некоторых характеристиках прогнозируемого объекта, их обработке математическими методами, получении зависимости, связывающей указанные характеристики со временем, и вычислении с помощью найденной зависимости характеристик объекта в заданный момент времени.

Этот подход предполагает применение моделирования. Прогнозирование в большинстве случаев является основой предупреждения над распространением заболеваемостью корью.

При возникновении вспышки заболеваемости прогнозируется ход развития обстановки, эффективность тех или иных намеченных мер по ликвидации роста заболеваемости, требуемый состав сил и средств. Наиболее важным из всех этих прогнозов является прогноз вероятности возникновения вспышки. Его результаты могут быть наиболее эффективно использованы для предотвращения распространения заболеваемости для заблаговременного снижения возможных потерь и ущерба, обеспечения готовности к ним, определения оптимальных превентивных мер.

Основная часть

Методология

Данный анализ использует набор статистических инструментов EpiNow2. Этот пакет оценивает изменяющееся во времени число воспроизводства, скорость роста и время удвоения с использованием ряда инструментов с открытым исходным кодом и текущих лучших практик. Он оценивает изменяющееся во времени число репродукций случаев по дате заражения (используя подход, аналогичный тому, который реализован в EpiEstim). Затем предполагаемые инфекции сопоставляются с наблюдаемыми данными (например, случаи по дате сообщения) с помощью ряда неопределенных распределений задержки и модели отчетности, которая может включать еженедельную периодичность. Неопределенность распространяется из всех входных данных в окончательные оценки параметров, помогая смягчить ложные выводы. Изменяющиеся во времени оценки воспроизводства и неопределенное время генерации также дают изменяющиеся во времени оценки темпов роста [2].

EpiNow2 использует байесовский подход латентных переменных с использованием языка вероятностного программирования Stan. Анализ начинается с оценки начального числа инфекций как свободного параметра с априорным показателем, основанным на начальном числе наблюдаемых случаев.

Для анализа используется R studio. R — это популярный пакет статистического анализа и машинного обучения, который обеспечивает управление данными и включает в себя тесты, модели, анализ и графику, а также позволяет управлять данными [3]. R Studio играет важную роль в области прогнозирования, предоставляя среду для разработки и реализации моделей, а также для анализа данных, что помогает аналитикам и исследователям принимать информированные решения на основе данных.

Для моделирования необходима информация о распределении времени генерации кори, структурированную как гамма-распределенная переменная; инкубационный период, оцененный как логнормально распределенный со средним значением 10-14 дней; распределение задержки регистрации, которое является логнормально распределенным и варьируется в зависимости от региона, в частности, отличается для города Алматы.

Для каждого последующего временного шага предыдущие расчетные инфекции (I_{t-1}) суммируются, взвешиваются неопределенной массовой функцией вероятности времени генерации ($\omega\omega$) и объединяются с оценкой R_t , чтобы получить заболеваемость в момент времени t ($I_t I_t$).

Для репродуктивного числа (R_0) используется логнормальное предшествующее значение со средним значением 1 и стандартным отклонением 1.

Затем строится кривая траекторий инфекции исходя из среднего количества зарегистрированных случаев (D_t) путем свертки по неопределенному инкубационному периоду и распределению задержки регистрации (свернутому в ϵ).

Затем предполагается, что наблюдаемые зарегистрированные случаи (C_t) генерируются отрицательной биномиальной моделью наблюдения с избыточной дисперсией ϕ (используя экспоненциальное априорное значение со средним значением 1) и средним D_t , умноженным на эффект дня недели с независимым параметром для каждого дня недели ($\omega_{t \bmod 7}$).

Временная вариативность контролируется с помощью приближенного гауссовского процесса с квадратичным экспоненциальным ядром (GP GP).

Кривая для каждого временного ряда строится независимо с помощью Марковской цепи Монте-Карло (MCMC). Используется как минимум 4 цепи с разогревом по 250 повторений и 2000 образцов после разогрева.

Модель прогнозируется на 14 дней после последней отчетной даты, проводя экстраполяцию тренда и неопределенности из периода, непосредственно предшествующего последней дате, по которой доступны данные.

Оценки отображаются и рассчитываются из апостериорного распределения ключевых оценок, представляющих интерес.

Эта система математически описывается следующим образом:

$$\begin{aligned}
 R_t &\sim R_{t-1} * GP \\
 I_t &= R_t \sum_T \omega_T * I_{t-T} \\
 D_t &= \sum_T \epsilon_t * I_{t-T} \\
 C_t &\sim NB(D_t * \omega_{t \bmod 7}, \phi)
 \end{aligned}$$

Рисунок 1. Математическое описание модели

Ограничения в методологии

Данный подход также имеет ряд ограничений. Во-первых, модель требует, чтобы доля зарегистрированных инфекций была постоянной в течение рассматриваемых 12 недель. Другими словами, требуется постоянство в фокусе метода эпиднадзора, уровне усилий, затраченных на тестирование, и определении случаев. Тем не менее, часто бывает так, что уровень недоучета в стране меняется в течение вспышки. Однако следует отметить, что любые изменения в процедурах тестирования при эпиднадзоре лишь временно искажают оценки, если после их изменения они снова начинают оставаться неизменными. Как долго смещение будет сохраняться в оценках репродуктивного числа будет зависеть от распределений серийного времени генерации и задержки, а также от масштаба длины гауссовского процесса, используемого в процессе оценки репродуктивного числа. Влияние погрешностей тестирования и других ошибок в отчетности варьируется между показателями распространения (случаи с положительным результатом тестирования, госпитализация, смерть с положительным результатом тестирования). По этой причине оценки основаны на зарегистрированных случаях смерти, и предоставляется инструментарий, позволяющий получить оценки для альтернативных наборов данных. Теоретически, оценки из разных источников должны быть сопоставимы при использовании данного подхода, однако если они фактически представляют разные субпопуляции, то между ними могут быть различия, которые потенциально могут быть полезны для интерпретации.

Кроме того, модель ограничена тем, насколько репрезентативно используемое нами распределение задержки от заражения до регистрации уведомления для данного района. Поскольку данные для оценки этого ограничены, было оценено бутстрапное глобальное распределение задержки, используя объединенные данные по каждой стране. В частности, задержка от появления инфекции до регистрации может особенно повлиять на увеличение числа случаев по дате возникновения заболевания, которое учитывает случаи, которые возникли, но еще не были зарегистрированы. Если истинная задержка от начала заболевания до уведомления в данной стране короче, чем наша глобальная задержка, то число случаев начала заболевания завышается, и наоборот, если истинная задержка больше, чем используемое нами распределение. Кроме того, известно, что оценки распределения задержки поступления информации в начале эпидемии необъективны и могут меняться со временем. Однако использование метода бутстрапной подвыборки смягчает эти проблемы, позволяя рассмотреть несколько распределений задержки, основанных на наблюдаемых данных, за счет увеличения неопределенности в оценках.

Данная модель также ограничена доступными данными. Например, общедоступные построчные списки содержат мало данных о статусе завоза

случаев. Это означает, что подсчеты случаев могут быть смещены в сторону увеличения за счет отнесения завозных случаев к местной передаче. Такое смещение особенно проблематично, когда число случаев невелико. К сожалению, в отсутствие данных этот вопрос можно изучить только с помощью анализа сценариев.

Примечательно, что прогноз является достоверным только при текущей вспышке кори, то есть для надежного прогноза необходимо большое число случаев.

Пример прогнозирования

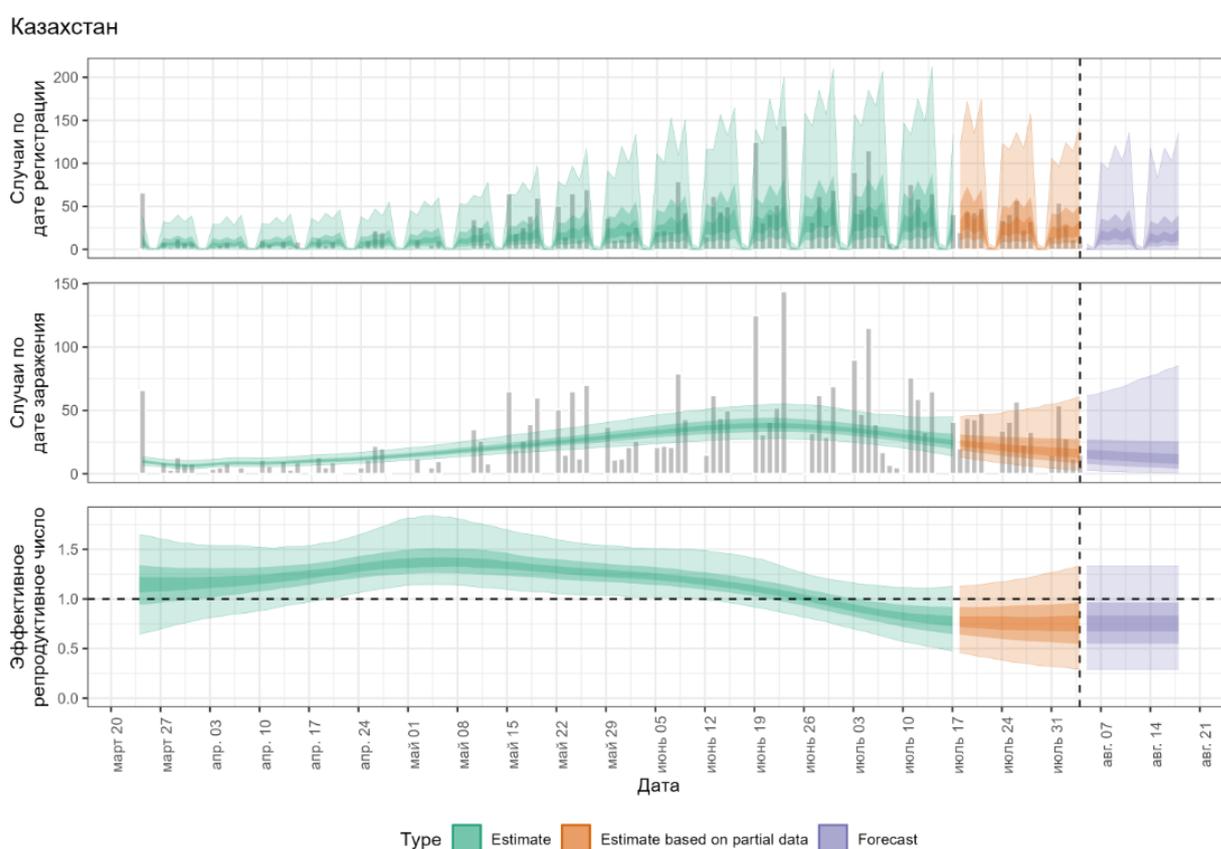


Рисунок 2. Прогноз по заболеваемости корью от 4.08.2023 на национальном уровне (Республика Казахстан)

Мера	Оценка
Новые подтвержденные случаи по дате заражения	15 (2 – 60)
Ожидаемое изменение в ежедневном количестве случаев	Вероятно уменьшение
Эффективное репродуктивное число	0.75 (0.29 – 1.3)
Темпы роста	-0.024 (-0.1 – 0.025)
Время увеличения/снижения в два раза (дни)	-29 (28 – -6.8)
Количество инфицированных на конец периода прогнозирования	11 (1 – 85)
Количество новых зарегистрированных случаев на конец периода прогнозирования	15 (0 – 136)
Кумулятивное количество инфицированных за период прогнозирования	181 (20 – 1022)
Кумулятивное количество зарегистрированных случаев за период прогнозирования	230 (65 – 813)

Таблица 1. Оценочные данные по заболеваемости корью от 4.08.2023 на национальном уровне (Республика Казахстан)

Алматинский регион

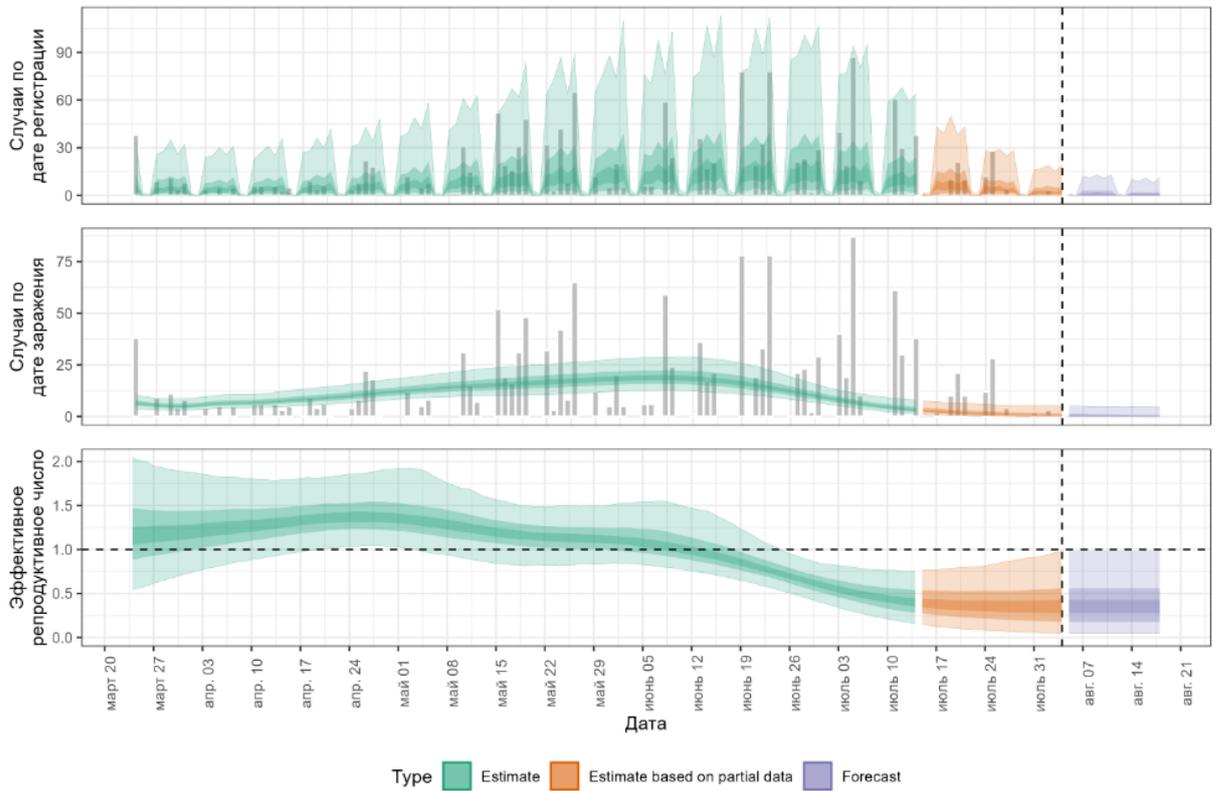


Рисунок 3. Прогноз по заболеваемости корью от 4.08.2023 на региональном уровне (Алматинский регион)

Мера	Оценка
Новые подтвержденные случаи по дате заражения	0 (0 – 5)
Ожидаемое изменение в ежедневном количестве случаев	Уменьшение
Эффективное репродуктивное число	0.35 (0.046 – 0.98)
Темпы роста	-0.088 (-0.24 – -0.002)
Время увеличения/снижения в два раза (дни)	-7.9 (-350 – -2.9)
Количество инфицированных на конец периода прогнозирования	0 (0 – 5)
Количество новых зарегистрированных случаев на конец периода прогнозирования	0 (0 – 11)
Кумулятивное количество инфицированных за период прогнозирования	3 (0 – 66)
Кумулятивное количество зарегистрированных случаев за период прогнозирования	14 (1 – 84)

Таблица 2. Оценочные данные по заболеваемости корью от 4.08.2023 на региональном уровне (Алматинский регион)

Регионы	Прогнозируемое количество подтвержденных случаев кори за период прогнозирования (на 14.07.2023)	Фактическое количество подтвержденных случаев кори по состоянию на 14.07.2023
Акмолинская	4	3
Актюбинская	78	0
Алматинская	24	83
Атырауская	29	34
ВКО	Данные не вошли в анализ из-за низкого количества случаев	3
Жамбылская	217	123
ЗКО	5	3
Карагандинская	5	9
Костанайская	40	8
Кызылординская	8	14
Мангистауская	3	14
Павлодарская	1	5
СКО	0	2
Туркестанская	2	0
г. Алматы	180	201
г. Астана	36	33
г.Шымкент	41	0

Абайская	0	5
Жетысуская	22	12
Улытауская	1	1
Республика Казахстан	636	553

Таблица 3. Прогнозные и фактические значения количества подтвержденных случаев кори по состоянию на 14.07.2023

Для отображения достоверности математического модели, было проведено сравнение прогнозируемых и фактических показателей заболеваемости корью с 30 июня по 14 июля 2023 года. Сравнение проведено по регионам и по всей Республике Казахстан. (Таблица 3).

Алгоритм прогнозирования заболеваемости корью:

1. Установка программы R studio
2. Подготовка исходных данных для прогнозирования
3. Загрузка и инициирование пакетов
4. Запуск прогнозирования
5. Создание отчета на основе результатов прогнозирования

Важное техническое примечание: поскольку прогнозирование занимает продолжительную часть времени (несколько часов), при запуске прогнозирования, компьютер на котором проводится анализ, должен быть постоянно подключен к питанию. Также, убедитесь, что на вашем компьютере отключен “спящий режим”, во избежание остановки прогнозирования. Кроме того, необходимо учесть, что скорость обработки данных зависит от технических характеристик компьютера, на котором выполняется анализ.

1. Установка программы R studio

Загрузить и установить R на компьютер можно с сайта <http://www.r-project.org/>

2. Подготовка исходных данных для прогнозирования

Информация о случаях кори (по состоянию на 12.07.2023 года)

Регiónы	случаев с подозрением на корь в нарастании	случаев с подозрением на корь за 11.07	диагнозы сняты	диагноз подтвержден в карантинах	диагноз подтвержден за 11.07	местный	защитной	возраст					прививочный статус					в работе		
								до года	от 1 до 14 лет	от 15 до 18 лет	19 лет и старше	привито		непривито			неизвестный статус			
												однократно	двукратно	отказ от вакцинации	без отвода	иные причины (уточнить)			недостаток возраста	
1 Акмоловская	34		13	20		19	1 (Египет)	5	13		2		3	1	7	3		5	1	1
2 Актобская	68		8	29		24	5(Египет-1, Турция-1, Узбекистан-1)	4	23	1	1	2	1	17	4		4	1	31	
3 Алматыская	846	28	58	347	3	347		56	227	16	48	33	29	105	65	2	56	57	441	
4 Атырауская	224	4	53	147		141	6 (Турция-3, Таиланд-2, Сауд. Аравия-1)	24	101	2	20	11	3	71	16		24	22	24	
5 ВКО	6	1	1	4	1	3	Турция-1		3		1			1	1	1		1	1	
6 Жамбылская	362	13	52	297	13	297		50	180	11	56	30	14	78	54		50	71	13	
7 ЗКО	37		16	17		15	2 (Ф. Узбекистан)	2	7		8	3	3	1	5		2	3	4	
8 Карагандинская	28	1	5	19		11	8 (Турция-3, Узбекистан-2, Египет-1, РФ-1, Шри-Ланка-1)	2	17				1	6	1		2	9	4	

Рисунок 4. Пример Microsoft Excel с данными по заболеваемости корью

Для корректной работы кода, данные по заболеваемости корью должны быть строго отсортированы и упорядочены в формате, показанном на рисунках 5, 6 и 7:

Имя папки/файла	Дата изменения	Тип	Размер
data	12.07.2023 16:46	Папка с файлами	
inst	12.07.2023 16:46	Папка с файлами	
int	12.07.2023 20:31	Папка с файлами	
output	12.07.2023 16:46	Папка с файлами	
R	13.07.2023 16:10	Папка с файлами	
.gitattributes	26.06.2023 21:51	Файл "GITATTRIBUTES"	1 КБ
.gitignore	26.06.2023 21:51	Файл "GITIGNORE"	1 КБ
.here	26.06.2023 21:51	Файл "HERE"	0 КБ
data	05.07.2023 14:33	Сжатая ZIP-папка	535 КБ
LICENSE	26.06.2023 21:51	Файл	2 КБ
README.md	26.06.2023 21:51	Файл "MD"	1 КБ

Рисунок 5. Формат хранения данных

Имя	Дата изменения	Тип	Разм
april	12.07.2023 16:46	Папка с файлами	
july	12.07.2023 16:46	Папка с файлами	
june	12.07.2023 16:46	Папка с файлами	
march	12.07.2023 16:46	Папка с файлами	
may	12.07.2023 16:46	Папка с файлами	

Рисунок 6. Формат хранения данных (по месяцам)

имя	дата изменения	тип	размер
03.07.2023	05.07.2023 14:30	Лист Microsoft Excel ...	38 КБ
04.07.2023	05.07.2023 14:30	Лист Microsoft Excel ...	38 КБ
05.07.2023	05.07.2023 14:30	Лист Microsoft Excel ...	38 КБ
06.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	39 КБ
07.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	39 КБ
08.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	39 КБ
09.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	39 КБ
10.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	39 КБ
11.07.2023	12.07.2023 16:30	Лист Microsoft Excel ...	38 КБ
12.07.2023	05.09.2023 12:11	Лист Microsoft Excel ...	39 КБ

Рисунок 7. Формат хранения данных (по дням)

3. Определение показателей и факторов для прогнозирования

Необходимо подготовить Excel файл для его дальнейшего использования в программе R studio. Пример Microsoft Excel файла показан на Рисунке 8.

region	super_region	rep_avg	rep_plus_minus	mean_sd	sd_sd	max	pop
Абайская	Восток	3	2	0.01227867	0.009708269	7	611888
Акмолинская	Столичный регион	3	2	0.01227867	0.009708269	7	785708
Актюбинская	Запад	3	2	0.01227867	0.009708269	7	916750
Алматинская	Алматинский регион	5	2.5	0.01227867	0.009708269	10	1478496
Атырауская	Запад	3	2	0.01227867	0.009708269	7	681241
ВКО	Восток	3	2	0.01227867	0.009708269	7	732966
Жамбылская		3	2	0.01227867	0.009708269	7	1209665
Жетысуская		3	2	0.01227867	0.009708269	7	698757
ЗКО	Запад	3	2	0.01227867	0.009708269	7	683327

Рисунок 8. Шаблон Microsoft Excel файла для загрузки в программу R studio

В столбце *region* необходимо вводить название соответствующей области.

В столбце *super_region* необходимо ввести название объединенных регионов, согласно географическому расположению. Распределение регионов может быть выполнено в произвольной форме.

Пример распределения регионов:

Объединенный регион	Входящие в него области/города
Восток	Абайская, Восточно-Казахстанская
Столичный	Акмолинская область и г. Астана
Запад	Актюбинская, Атырауская, Мангистауская, Западно-Казахстанская
Алматинский	Алматинская область и г. Алматы
Центр	Карагандинская, Улытауская
Юг	Туркестанская область и г. Шымкент

Таблица 4. Распределение областей по регионам

Столбец *rep_avg* содержит информацию о среднем времени между регистрацией случая и появлением симптомов.

Столбец *rep_plus_minus* показывает отклонение от среднего времени между регистрацией случая и появлением симптомов (*rep_avg*).

Столбцы *mean_sd* и *sd_sd* отражают параметры стандартного отклонения, взятые из литературных источников.

Столбец *max* показывает максимальное возможное среднее время между регистрацией случая и появлением симптомов (*rep_avg*).

Столбец *pop* содержит информацию о количестве населения в соответствующем регионе/области.

4. Загрузка и инициирование пакетов

Для загрузки и инициирования пакетов необходимо ввести код, представленный в Приложение 1 и нажать кнопку Run.

4.. Запуск прогнозирования

Код, описанный в этом шаге алгоритма, включает в себя прогнозирование заболеваемости для всей Республики Казахстан, а также отдельно для регионов.

В командной строке ввести следующий код и нажать кнопку Run:

```
library(here)

source(here("R/src.R"))

#load measles data
measles_reg <- load_measles_data()

#create super region data
measles_superreg <- read_csv(here("inst/ref.csv")) |>
  select(region, super_region) |>
  right_join(measles_reg, by = "region") |>
  filter(!is.na(super_region)) |>
  group_by(date, super_region) |>
  summarise(confirm = sum(confirm)) |>
  rename(region = super_region)

#create national data
measles_nat <- measles_reg |>
  group_by(date) |>
  summarise(confirm = sum(confirm))

epi_params <- get_epi_params()
```

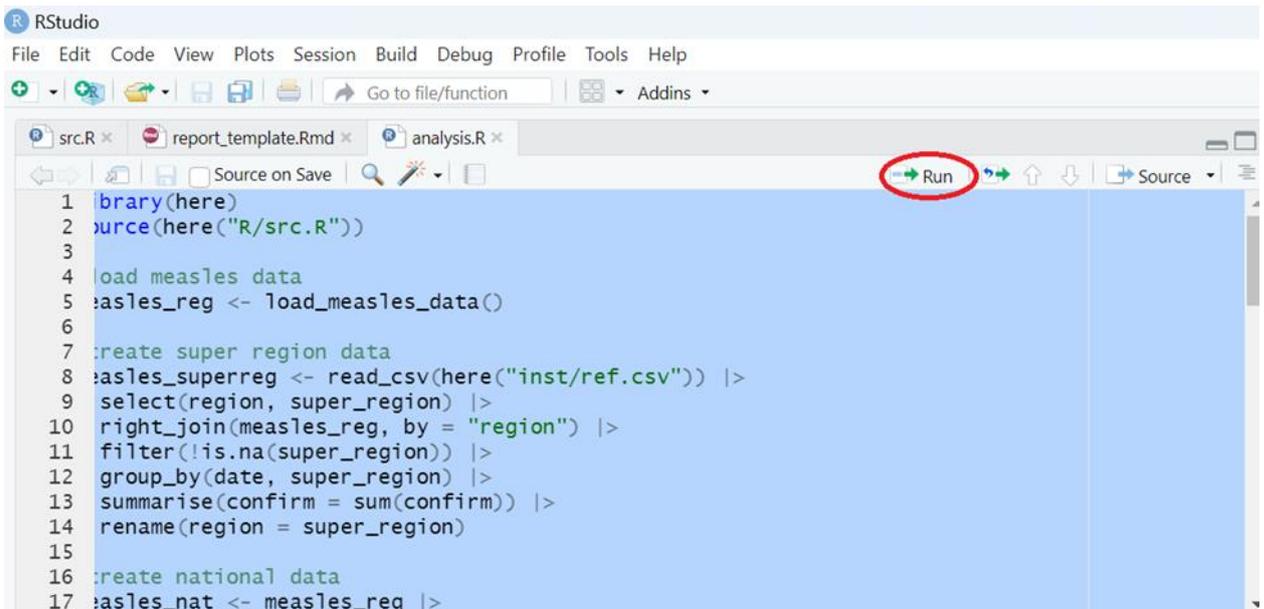


Рисунок 9. Запуск кода для прогнозирования

После нажатия кнопки Run, необходимо дождаться, когда перестанет высвечиваться символ Stop:

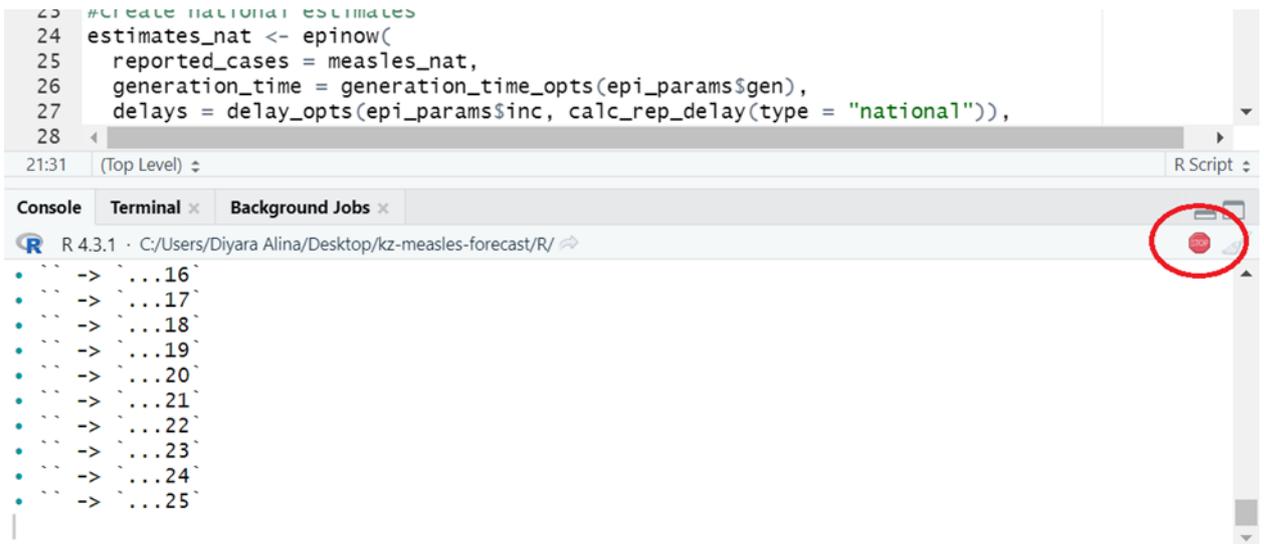


Рисунок 10. Процесс обработки кода

Необходимо дождаться, когда символ Stop исчезнет (Рисунок 11).

```

14 rename(region = super_region)
15
16 #create national data
17 measles_nat <- measles_reg |>
18

```

21:31 (Top Level) ↕ R Script ↕

Console Terminal x Background Jobs x

R 4.3.1 · C:/Users/Diyara Alina/Desktop/kz-measles-forecast/R/ ↗

```

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
`summarise()` has grouped output by 'date'. You can override using the `.groups` argument.
>
> #create national data
> measles_nat <- measles_reg |>
+   group_by(date) |>
+   summarise(confirm = sum(confirm))
>
> epi_params <- get_epi_params()
> |

```

Рисунок 11. Завершение обработки кода

Далее необходимо ввести следующий код и нажать кнопку Run:

```

#create national estimates
estimates_nat <- epinow(
  reported_cases = measles_nat,
  generation_time = generation_time_opts(epi_params$gen),
  delays = delay_opts(epi_params$inc, calc_rep_delay(type = "national")),
  rt = rt_opts(),
  stan = stan_opts(cores = future::availableCores()-1, control = list(adapt_delta =
0.99)),
  verbose = interactive(),
  horizon = 14
)

write_rds(estimates_nat, here("int/estimates_nat.rds"))

rm(estimates_nat)

gc()

estimates_superreg <- regional_epinow(

```

```

reported_cases = measles_superreg,
generation_time = generation_time_opts(epi_params$gen),
delays = create_reg_delay(inc = epi_params$inc, rep =
calc_rep_delay("super_region")),
rt = rt_opts(),
stan = stan_opts(cores = future::availableCores()-1, control = list(adapt_delta =
0.99)),
verbose = interactive(),
horizon = 14
)

```

```

write_rds(estimates_superreg, here("int/estimates_superreg.rds"))

```

```

rm(estimates_superreg)

```

```

gc()

```

```

estimates_reg <- regional_epinow(
reported_cases = measles_reg,
generation_time = generation_time_opts(epi_params$gen),
delays = create_reg_delay(inc = epi_params$inc, rep = calc_rep_delay("region")),
rt = rt_opts(),
stan = stan_opts(cores = future::availableCores()-1, control = list(adapt_delta =
0.99)),
verbose = interactive(),
horizon = 14
)

```

```

write_rds(estimates_reg, here("int/estimates_reg.rds"))

```

```
rm(estimates_reg)
```

```
gc()
```

```
#create and translate key figures
```

```
extract_translate_figs(file = here("int/estimates_nat.rds"), type = "national")
```

```
extract_translate_figs(file = here("int/estimates_superreg.rds"), type =  
"super_region")
```

```
extract_translate_figs(file = here("int/estimates_reg.rds"), type = "region")
```

На этом этапе так же необходимо дождаться завершения обработки кода.

5. Создание отчета на основе результатов прогнозирования

Правой кнопкой мыши нажать на файл в формате RMD и выбрать опцию “Открыть с помощью R studio”. Таким образом, открывается шаблон для создания отчетной формы по прогнозированию.

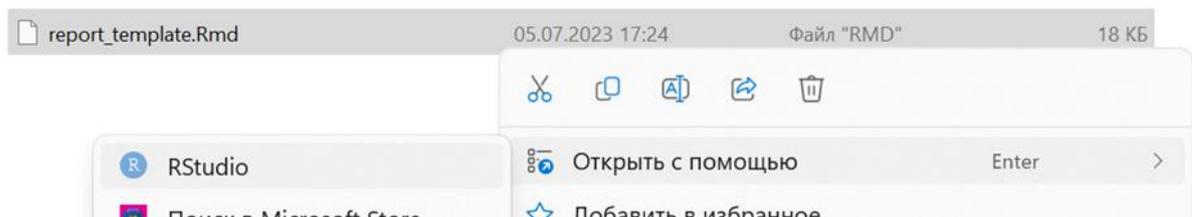


Рисунок 12. Запуск шаблона отчета

После открытия данного файла необходимо ввести актуальные данные согласно шаблону. Далее сохранить отчет в формате HTML.

Цель

Результаты

Примечания

Благодарности

Контактная информация

Цель

Результаты

Примечания

Благодарности

Контактная информация

Отчет по прогнозированию заболеваемости корью в Республике Казахстан

2023-08-18

Цель

- Данный отчет создан с целью оценки эпидемиологической ситуации и дальнейшего распространения заболеваемости корью в Республике Казахстан.
- На **18.08.2023** в Республике Казахстан зарегистрировано **3285** случаев кори.
- Данный отчет может быть использован Министерством здравоохранения для принятия необходимых ответных мер.

Примечания

- Дата проведения анализа: 2023-08-18
- Источник данных: Отчетная информация о случаях кори на **18.08.2023** г.

Результаты

1. Национальный уровень

Processing math: 100%

Прогноз Сводная таблица

Рисунок 13. Пример отчета по прогнозированию

Заключение

В настоящих методических рекомендациях описывается модель прогнозирования заболеваемости корью, основанную на анализе подтвержденных случаев данных и рядах факторов, влияющих на распространение этого заболевания. Полученные результаты позволяют сделать следующие выводы:

На основе данной модели, мы можем предсказать тенденции в распространении кори в будущем с высокой точностью, что может быть полезно для планирования и мониторинга программ по профилактике и лечению этого заболевания.

В целом, данная модель представляет собой значимый инструмент для борьбы с корью и может способствовать улучшению здоровья населения. Дальнейшие исследования в области здравоохранения могут помочь внедрить разработанные методы в практику и добиться снижения заболеваемости корью в будущем.

Данная методика прогнозирования заболеваемости корью может использоваться для эффективного планирования профилактических и контрольных мероприятий по борьбе с этой опасной инфекцией. Внедрение методов прогнозирования с использованием эффективных статистических инструментов позволит определять и планировать направления и объем ответных мер.

Список использованных источников

1. <https://www.who.int/ru/news-room/fact-sheets/detail/measles>
2. <https://epiforecasts.io/EpiNow2/index.html>
3. <https://www.ibm.com/docs/en/watson-studio-local/1.2.3?topic=data-rstudio>
4. <https://www.who.int/teams/immunization-vaccines-and-biologicals/immunization-analysis-and-insights/surveillance/surveillance-for-vpds/laboratory-networks/measles-and-rubella-laboratory-network/manual-for-the-laboratory-based-surveillance-of-measles-rubella-and-congenital-rubella-syndrome> - World Health Organization (WHO) Measles Surveillance Manual: Руководство ВОЗ по сбору и анализу данных о кори предоставляет основные методы сбора и анализа данных о заболеваемости корью
5. <https://pubmed.ncbi.nlm.nih.gov/22534001/> -Simons E, Ferrari M, Fricks J, et al. Assessment of the 2010 global measles mortality reduction goal: results from a model of surveillance data. Lancet. 2012;379(9832):2173-2178.: Эта статья предоставляет информацию о глобальных усилиях по снижению смертности от кори.
6. Metcalf CJE, Børnstad ON, Grenfell BT, et al. Interrupting measles transmission in West Africa with oral polio vaccine. Science. 2016;354(6316):630-632.: Эта статья оценивает методы прерывания передачи кори .
7. <https://jamanetwork.com/journals/jama/fullarticle/2764824>
Jewell NP, Lewnard JA, Jewell BL. Predictive Mathematical Models of the COVID-19 Pandemic: Underlying Principles and Value of Projections. JAMA. 2020;323(19):1893-1894.: Эта статья о моделях COVID-19 может предоставить полезные принципы для разработки модели заболеваемости корью.
8. https://lit-review.ru/guides/epid_dict.pdf
: Эта книга предоставляет обширные методологические сведения о статистических методах в эпидемиологии, что может быть полезно при анализе данных.

Приложение

Код для шага Загрузка и инициирование пакетов

```
#key libraries

library(readxl)

#library(tidycl)

library(tidyverse)

library(lubridate)

library(EpiNow2)

library(patchwork)

#' #' Function to clean erroneous dates
#' #' @description reads in dates as a characters and performs cleaning steps
#' #' specifically for data provided by KZ MoH
#' #' @param date_to_check str: a date formatted as a string
#' #' @returns date formatted value
#' error_check_date <- function(date_to_check){
#'
#'   if(grepl("r.p", date_to_check)){
#'     if(nchar(date_to_check) > 8){
#'       a <- gsub("[^0-9.]","", date_to_check) |>
#'         str_split("\\.")
#'       as_date(paste0(a[[1]][3], "-", a[[1]][2], "-", a[[1]][1]))
#'     }else{
#'       as_date(paste0(gsub("[^0-9]","", date_to_check), "-01-01"))
#'     }
#'   }
```

```

#' }else{
#'   if(grepl("r", date_to_check)){
#'     if(unlist(gregexpr('r', date_to_check))<6){
#'       a <- gsub("[^0-9.]", "", date_to_check) |>
#'         str_split("\\.")
#'
#'       as_date(paste0(a[[1]][1], "-", a[[1]][2], "-", a[[1]][3]))
#'     }else{
#'       a <- gsub("[^0-9.]", "", date_to_check) |>
#'         str_split("\\.")
#'
#'       as_date(paste0(a[[1]][3], "-", a[[1]][2], "-", a[[1]][1]))
#'     }
#'   }else{
#'     if(grepl("ж", date_to_check)){
#'
#'       if(unlist(gregexpr('ж', date_to_check))<6){
#'         a <- gsub("[^0-9.]", "", date_to_check) |>
#'           str_split("\\.")
#'
#'         as_date(paste0(a[[1]][1], "-", a[[1]][2], "-", a[[1]][3]))
#'       }else{
#'         a <- gsub("[^0-9.]", "", date_to_check) |>
#'           str_split("\\.")
#'
#'         as_date(paste0(a[[1]][3], "-", a[[1]][2], "-", a[[1]][1]))

```



```

#' z <- xlsx_formats(file_location)
#'
#' #identify cells that are colored "red"
#' red <- y |>
#'   filter(local_format_id %in%
#'     which(z$local$fill$patternFill$fgColor$rgb %in% c("FFFF0000",
#' "FF5B9BD5", "FFED7D31"))) |>
#'   select(sheet, address) |>
#'   mutate(status = "confirmed")
#'
#' #row numbers for where it is red
#' guide_confirmed <- left_join(y, red, by = c("sheet", "address")) |>
#'   select(sheet, address, row, col, status) |>
#'   filter(col == color_col_num &
#'     sheet == soi &
#'     row != 1 &
#'     status == "confirmed")
#'
#' #all rows that are "red"
#' x_confirmed <- x[guide_confirmed$row-2,] |>
#'   select(province, dob = DOB, exposure = "Date of exposure",
#'     report = "Date of report", hosp = "Date of hospitalization")
#'
#'
#'
#' x_confirmed <- x_confirmed |>
#'   rowwise() |>

```

```

#' mutate(dob = error_check_date(dob),
#'         exposure = error_check_date(exposure),
#'         report = error_check_date(report),
#'         hosp = error_check_date(hosp)) |>
#' mutate(status = "confirmed",
#'         sheet = soi)
#'
#'
#'
#' if(include_primary){
#'   white <- y |>
#'     filter(local_format_id %in%
#'            which(is.na(z$local$fill$patternFill$fgColor$rgb) |
#'                  z$local$fill$patternFill$fgColor$rgb == "FFFFFFFF")) |>
#'     select(sheet, address) |>
#'     mutate(status = "primary_case")
#'
#'   guide_primary <- left_join(y, white, by = c("sheet", "address")) |>
#'     select(sheet, address, row, col, status) |>
#'     filter(col == color_col_num &
#'            sheet == soi &
#'            row != 1 &
#'            status == "primary_case")
#'
#'   x_primary <- x[guide_primary$row-2,] |>
#'     select(province, dob = DOB, exposure = "Date of exposure",
#'            report = "Date of report", hosp = "Date of hospitalization")

```

```

#'
#'
#' x_primary <- x_primary |>
#'   rowwise() |>
#'   mutate(dob = error_check_date(dob),
#'          exposure = error_check_date(exposure),
#'          report = error_check_date(report),
#'          hosp = error_check_date(hosp)) |>
#'   mutate(status = "primary_case",
#'          sheet = soi)
#'   return(bind_rows(x_primary, x_confirmed) |>
#'          filter(!is.na(province)))
#'
#' }else{
#'   return(x_confirmed)
#' }
#'
#' }
#'
#' #' Function to iterate through all sheets and load in data
#' #' @description Load in all data from KZ MoH measles line list
#' #' @param file_location str: location of KZ MoH measles line list
#' #' @returns tibble: machine readable line list
#' load_measles_data <- function(file_location){
#'
#'   sheets <- excel_sheets(file_location)

```

```

#'
#' measles <- lapply(sheets, function(x){
#'   extract_measles_data(file_location = file_location,
#'                         soi = x,
#'                         include_primary = T)
#' }) |>
#'   bind_rows()
#'
#'
#' }

#' Create key epidemiological parameters for estimation
#'
#' @description Create key epidemiological parameters for estimation
#' @returns list: generation time, incubation period and reporting delay
distributions
get_epi_params <- function(){

  #pulled from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6198657/
  generation_time <- list(
    "mean" = 11.7,
    "mean_sd" = 0,
    "sd" = 3,
    "sd_sd" = 0,
    "max" = 30,
    "dist" = "gamma"
  )
}

```

```

)

#using estimation of measles incubation period of 10-14 days
incubation_period <- list(
  "mean" = log(12),
  "mean_sd" = 0,
  "sd" = log(2),
  "sd_sd" = 0,
  "max" = 21,
  "dist" = "lognormal"
)

return(list(
  inc = incubation_period,
  gen = generation_time
))

}

#' Create reporting delay
#' @description Function to create reporting delay priors
#' @param rep_delay_file str: location of file
#' @param type str: national, super_region, region
#' @return named list as needed by EpiNow2
calc_rep_delay <- function(type, rep_delay_file = here("inst/ref.csv")){

```

```

#national

if(type == "national"){
  y <- read_csv(rep_delay_file) |>
  mutate(tot = sum(pop),
         weight = pop / tot,
         rep_avg = log(rep_avg),
         rep_sd = (log(rep_plus_minus)/1.96)^2) |>
  summarize(
    mean = weighted.mean(rep_avg, weight),
    mean_sd = max(mean_sd),
    sd = weighted.mean(rep_sd, weight),
    sd_sd = max(sd_sd),
    max = max(max),
    dist = "lognormal"
  ) |>
  as.list()

return(y)
}

if(type == "super_region"){

  x <- read_csv(rep_delay_file) |>
  group_by(super_region) |>
  mutate(tot = sum(pop),
         weight = pop / tot,

```

```

    rep_avg = log(rep_avg),
    rep_sd = (log(rep_plus_minus)/1.96)^2) |>
summarize(
  mean = weighted.mean(rep_avg, weight),
  mean_sd = max(mean_sd),
  sd = weighted.mean(rep_sd, weight),
  sd_sd = max(sd_sd),
  max = max(max),
  dist = "lognormal"
) |>
drop_na()

y <- x |>
group_by(super_region) |>
group_split(.keep = F) |>
lapply(as.list)

names(y) <- x$super_region

return(y)

}

if(type == "region"){

  x <- read_csv(rep_delay_file) |>

```

```

group_by(region) |>
mutate(tot = sum(pop),
       weight = pop / tot,
       rep_avg = log(rep_avg),
       rep_sd = (log(rep_plus_minus)/1.96)^2) |>
summarize(
  mean = weighted.mean(rep_avg, weight),
  mean_sd = max(mean_sd),
  sd = weighted.mean(rep_sd, weight),
  sd_sd = max(sd_sd),
  max = max(max),
  dist = "lognormal"
) |>
drop_na()

y <- x |>
group_by(region) |>
group_split(.keep = F) |>
lapply(as.list)

names(y) <- x$region

return(y)

}

```

```
}
```

```
#' Read in measles data in format provided by KZ MoH
```

```
#' @description Read in measles data in format provided by KZ MoH
```

```
#' @param data_folder str: location of data, defaults to here("data")
```

```
#' @return tibble of region specific measles data
```

```
load_measles_data <- function(data_folder = here("data")){
```

```
  files_full <- list.files(data_folder, full.names = T, recursive = T)
```

```
  dates <- list.files(data_folder, recursive = T) |>
```

```
    sapply(function(x){
```

```
      str_split(x, "/")[[1]][2] |>
```

```
        str_split("\\.") %>%
```

```
        {paste0(.[[1]][3], "-", .[[1]][2], "-", .[[1]][1])}
```

```
    })
```

```
  x <- lapply(1:length(files_full), function(n){
```

```
    read_excel(files_full[n], skip = 4) %>%
```

```
      {
```

```
        .[,c(2,6)]
```

```
      } |>
```

```
      set_names(c("region", "confirm")) |>
```

```
      filter(!is.na(region) & region != "PK") |>
```

```

    replace_na(list("confirm" = 0)) |>
    mutate(date = as_date(dates[n]))

  }) |>
  bind_rows()

  tibble(
    "region" = unique(x$region),
    "confirm" = 0,
    "date" = as_date("2023-03-23")
  ) |>
  bind_rows(x) |>
  arrange(date) |>
  group_by(region) |>
  mutate(confirm = confirm - lag(confirm, default = confirm[1])) |>
  mutate(confirm = ifelse(confirm < 0, 0, confirm))

}

#' Creates reporting delay by region
#' @description helper function to create region specific variable reporting delay
#' @param inc list: incubation period - does not vary by region
#' @param rep list: reporting delay as provided in inst/ref.csv
#' @return nested list with total delay (incubation and reporting) per region
create_reg_delay <- function(inc, rep){

```

```
x <- names(rep)
```

```
y <- list()
```

```
for(i in 1:length(x)){
```

```
  y[[i]] <- delay_opts(epi_params$inc, rep[[x[i]])
```

```
}
```

```
names(y) <- x
```

```
return(y)
```

```
}
```

```
#' Extract output of EpiNow and write out key functions and tables
```

```
#' @description helper function to extract data from EpiNow2
```

```
#' @param file str: location of EpiNow2 output
```

```
#' @param type str: national, region, or super_region
```

```
#' @returns Writes out key figures and tables in format to be read by RMD
```

```
extract_translate_figs <- function(file, type){
```

```
  data <- read_rds(file)
```

```
  if(type == "national"){
```

```
    dir.create(here("output/national"), recursive = T, showWarnings = F)
```

```

sum_cases <- data$estimates$samples |>
  as_tibble() |>
  filter(variable %in% c("infections", "reported_cases") &
    type == "forecast") |>
  group_by(variable, sample) |>
  summarise(value = sum(value))|>
  group_by(variable) |>
  summarise(
    median = round(median(value),0),
    ll = round(quantile(value, 0.05),0),
    ul = round(quantile(value, 0.95),0)) |>
  mutate(
    measure = case_when(
      variable == "infections" ~ "Кумулятивное количество инфицированных
за период прогнозирования",
      variable == "reported_cases" ~ "Кумулятивное количество
зарегистрированных случаев за период прогнозирования"
    ),
    estimate = paste0(median, " (", ll, " -- ", ul, ")")
  ) |>
  select(measure, estimate)

incident <- data$estimates$samples |>
  as_tibble() |>
  filter(variable %in% c("infections", "reported_cases") &
    type == "forecast") |>

```

```

filter(date == max(date)) |>
group_by(variable) |>
summarise(
  median = round(median(value),0),
  ll = round(quantile(value, 0.05),0),
  ul = round(quantile(value, 0.95),0)) |>
  mutate(
    measure = case_when(
      variable == "infections" ~ "Количество инфицированных на конец
периода прогнозирования",
      variable == "reported_cases" ~ "Количество новых зарегистрированных
случаев на конец периода прогнозирования"
    ),
    estimate = paste0(median, " (", ll, " -- ", ul, ")")
  ) |>
select(measure, estimate)

```

```

data$summary |>
as_tibble() |>
mutate(
  estimate = case_when(
    estimate == "Likely increasing" ~ "Вероятно увеличение",
    estimate == "Increasing" ~ "Увеличение",
    estimate == "Stable" ~ "Стабильный",
    estimate == "Likely decreasing" ~ "Вероятно уменьшение",
    estimate == "Decreasing" ~ "Уменьшение",

```

```

  T ~ estimate
),
measure = case_when(
  measure == "New confirmed cases by infection date" ~ "Новые
подтвержденные случаи по дате заражения",
  measure == "Expected change in daily cases" ~ "Ожидаемое изменение в
ежедневном количестве случаев",
  measure == "Effective reproduction no." ~ "Эффективное репродуктивное
число",
  measure == "Rate of growth" ~ "Темпы роста",
  measure == "Doubling/halving time (days)" ~ "Время
увеличения/снижения в два раза (дни)"
)
)|>
select(measure, estimate) |>
bind_rows(incident, sum_cases) |>
select("Мера" = measure, "Оценка" = estimate) |>
write_rds(here("output/national/table.rds"))

b <- data$plots$infections +
labs(x = "Дата", y = "Случаи по дате заражения") +
  theme(
axis.text.x = element_blank(),
axis.title.x = element_blank(),
axis.ticks.x = element_blank()
)

```

```

a <- data$plots$reports +
labs(x = "Дата", y = "Случаи по\ндате регистрации") +
  theme(
axis.text.x = element_blank(),
  axis.title.x = element_blank(),
axis.ticks.x = element_blank()
)

c <- data$plots$R +
labs(x = "Дата", y = "Эффективное\препродуктивное число")

plot <- a/b/c

plot <- plot + plot_annotation(title = "Казахстан") + plot_layout(guides =
"collect") & theme(legend.position = 'bottom')

ggsave(here("output/national/summary.png"), plot,
  height = 7, width = 10, dpi = 300)

}else{

  head <- here(paste0("output/",type))

dir.create(head, recursive = T, showWarnings = F)

regions <- names(data$regional)

```

```

ggsave(paste0(head, "/summary_plot.png"),
       plot = data$summary$summary_plot,
       height = 8, width = 10, dpi = 300)

data$summary$summarised_results$table |>
  as_tibble() |>
  rename(estimate = `Expected change in daily cases`) |>
  mutate(
    estimate = case_when(
      estimate == "Likely increasing" ~ "Вероятно увеличение",
      estimate == "Increasing" ~ "Увеличение",
      estimate == "Stable" ~ "Стабильный",
      estimate == "Likely decreasing" ~ "Вероятно уменьшение",
      estimate == "Decreasing" ~ "Уменьшение",
      T ~ estimate
    ) |>
  select(
    "Регион" = Region,
    "Новые подтвержденные случаи по дате заражения" = `New confirmed
cases by infection date`,
    "Эффективное репродуктивное число" = `Effective reproduction no.`,
    "Темпы роста" = `Rate of growth`,
    "Время увеличения/снижения в два раза (дни)" = `Doubling/halving time
(days)`
  ) |>

```

```

write_rds(paste0(head,"/summary_table.rds"))

error <- ""

for(i in regions){

  if("error" %in% names(data$regional[[i]])){

    error <- paste0(error, " | ", i)

  }else{

    reg_loc <- paste0(head,"/",i)

    dir.create(reg_loc, recursive = T, showWarnings = F)

    sum_cases <- data$regional[[i]]$estimates$samples |>
      as_tibble() |>
      filter(variable %in% c("infections", "reported_cases") &
        type == "forecast") |>
      group_by(variable, sample) |>
      summarise(value = sum(value))|>
      group_by(variable) |>
      summarise(
        median = round(median(value),0),
        ll = round(quantile(value, 0.05),0),

```

```

  ul = round(quantile(value, 0.95),0) |>
mutate(
  measure = case_when(
    variable == "infections" ~ "Кумулятивное количество инфицированных
за период прогнозирования",
    variable == "reported_cases" ~ "Кумулятивное количество
зарегистрированных случаев за период прогнозирования"
  ),
  estimate = paste0(median, " (", ll, " -- ", ul, ")")
) |>
select(measure, estimate)

incident <- data$regional[[i]]$estimates$samples |>
as_tibble() |>
filter(variable %in% c("infections", "reported_cases") &
  type == "forecast") |>
filter(date == max(date)) |>
group_by(variable) |>
summarise(
  median = round(median(value),0),
  ll = round(quantile(value, 0.05),0),
  ul = round(quantile(value, 0.95),0) |>
mutate(
  measure = case_when(
    variable == "infections" ~ "Количество инфицированных на конец
периода прогнозирования",
    variable == "reported_cases" ~ "Количество новых зарегистрированных

```

случаев на конец периода прогнозирования"

```
),  
  estimate = paste0(median, " (", ll, " -- ", ul, ")")  
) |>  
select(measure, estimate)
```

```
data$regional[[i]]$summary |>
```

```
as_tibble() |>
```

```
mutate(  
  estimate = case_when(  
    estimate == "Likely increasing" ~ "Вероятно увеличение",  
    estimate == "Increasing" ~ "Увеличение",  
    estimate == "Stable" ~ "Стабильный",  
    estimate == "Likely decreasing" ~ "Вероятно уменьшение",  
    estimate == "Decreasing" ~ "Уменьшение",  
    T ~ estimate  
  ),  
  measure = case_when(  
    measure == "New confirmed cases by infection date" ~ "Новые  
подтвержденные случаи по дате заражения",  
    measure == "Expected change in daily cases" ~ "Ожидаемое изменение  
в ежедневном количестве случаев",  
    measure == "Effective reproduction no." ~ "Эффективное репродуктивное  
число",  
    measure == "Rate of growth" ~ "Темпы роста",  
    measure == "Doubling/halving time (days)" ~ "Время  
увеличения/снижения в два раза (дни)"
```

```

)
)|>
select(measure, estimate) |>
bind_rows(incident, sum_cases) |>
select("Мера" = measure, "Оценка" = estimate) |>
write_rds(paste0(reg_loc, "/table.rds"))

```

```

b <- data$regional[[i]]$plots$infections +
  labs(x = "Дата", y = "Случаи по\ндате заражения") +
  theme(
    axis.text.x = element_blank(),
    axis.title.x = element_blank(),
    axis.ticks.x = element_blank()
  )

```

```

a <- data$regional[[i]]$plots$reports +
  labs(x = "Дата", y = "Случаи по\ндате регистрации") +
  theme(
    axis.text.x = element_blank(),
    axis.title.x = element_blank(),
    axis.ticks.x = element_blank()
  )

```

```

c <- data$regional[[i]]$plots$R +
  labs(x = "Дата", y = "Эффективное\нрепродуктивное число")

```

```
plot <- a/b/c
```

```
plot <- plot + plot_annotation(title = i) + plot_layout(guides = "collect") &  
theme(legend.position = 'bottom')
```

```
ggsave(paste0(reg_loc, "/summary.png"), plot,  
        height = 7, width = 10, dpi = 300)
```

```
}
```

```
write_rds(error, paste0(head, "/errors.rds"))
```