

УДК 614.2 + 303.4

## **ЭКОЛОГИЧЕСКИЕ (КОРРЕЛЯЦИОННЫЕ) ИССЛЕДОВАНИЯ В ЗДРАВООХРАНЕНИИ**

**А. М. Гржибовский<sup>1-3</sup>, С. В. Иванов<sup>4</sup>, М. А. Горбатова<sup>2</sup>**

<sup>1</sup> Национальный Институт Общественного Здравоохранения, г. Осло, Норвегия;

<sup>2</sup> Северный Государственный Медицинский Университет, г. Архангельск, Россия;

<sup>3</sup> Международный Казахско-Турецкий Университет им. Х.А. Ясави, г. Туркестан, Казахстан;

<sup>4</sup> Северо-Западный Государственный Медицинский Университет им. И.И. Мечникова, г. Санкт-Петербург, Россия.

В настоящей работе представлены основные принципы планирования, проведения и статистической обработки данных экологических (корреляционных) исследований в здравоохранении. Приведены практические примеры расчета размера выборки и оценки связи между факторами риска с помощью онлайн-калькуляторов и пакета для обработки эпидемиологических данных Epi Info. Авторы представляют практические примеры использования экологических (корреляционных) исследований в литературе, а также достоинства и недостатки данного типа исследований. Настоящая статья призвана дать общие сведения об экологических (корреляционных) исследованиях, и не заменяет прочтения специализированной литературы по клинической эпидемиологии.

**Ключевые слова:** принципы планирования исследований, экологические (корреляционные) исследования, статистическая обработка, эпидемиология, дизайн исследования.

## **ECOLOGICAL (CORRELATION) STUDIES IN HEALTH SCIENCES**

**A. M. Grjibovski<sup>1-3</sup>, S. V. Ivanov<sup>4</sup>, M. A. Gorbatova<sup>2</sup>**

<sup>1</sup> Norwegian Institute of Public Health, Oslo, Norway;

<sup>2</sup> Northern State Medical University, Arkhangelsk, Russia;

<sup>3</sup> International Kazakh - Turkish University named after H.A. Yassawe, Turkestan, Kazakhstan;

<sup>4</sup> North-Western State Medical University named after I.I. Mechnikov, St. Petersburg, Russia.

In this paper we present basic principles of planning and carrying out ecological (correlation) studies in health sciences as well as analyzing the data collected in these studies. We present practical examples of sample size and effect measures calculation for ecological studies using online calculators. Examples of studies using ecological (correlation) design from international peer-reviewed literature are also presented. This article provides general introduction to ecological studies and does not substitute specialized literature in clinical epidemiology.

**Keywords:** principles of research planning, ecological (correlation) studies, statistical analysis, epidemiology, study design.

## **ДЕНСАУЛЫҚ САҚТАУДАҒЫ КӨЛДЕНЕҢ ЭКОЛОГИЯЛЫҚ (КОРРЕЛЯЦИЯЛЫҚ) ЗЕРТТЕУЛЕР**

**А. М. Гржибовский<sup>1-3</sup>, С. В. Иванов<sup>4</sup>, М. А. Горбатова<sup>2</sup>**

<sup>1</sup> Қоғамдық Денсаулық сақтау Ұлттық Институты, Осло қ., Норвегия;

<sup>2</sup> Солтүстік Мемлекеттік Медициналық Университеті, Архангельск қ., Ресей;

<sup>3</sup> Х.А. Ясави ат. Халықаралық Қазақ – Түрік Университеті, Туркестан, Қазақстан;

<sup>4</sup> И. И. Мечников атынд. Солтүстік – Батыс мемлекеттік медициналық университеті, Санкт-Петербург қ., Ресей.

Осы жұмыста денсаулық сақтаудағы көлденең экологиялық (корреляциялық) жоспарлау, өткізу және статистикалық өңдеудің негізгі принциптері көрсетілген. Ері Info эпидемиологиялық мәліметтерді өңдеу үшін онлайн-калькуляторлар мен пакет көмегімен қауіп факторлары арасындағы байланысты таңдау және бағалау мөлшері есебінің практикалық мысалдары келтірілген. Авторлар әдебиеттегі көлденең экологиялық (корреляциялық) зерттеулерді пайдаланудың практикалық мысалдарын, сол сияқты осы зерттеу түрлері қасиеттері мен жетіспеушіліктерін келтіреді. Осы мақала көлденең экологиялық (корреляциялық) зерттеулер туралы жалпы мәліметтер беруге бейімделген, және клиникалық эпидемиология бойынша мамандандырылған әдебиетті оқуды ауыстырмайды.

**Негізгі сөздер:** зерттеулерді жоспарлау принциптері, экологиялық (корреляциялық) зерттеулер, статистикалық өңдеу, эпидемиология, зерттеу дизайны.

#### **Библиографическая ссылка:**

Гржибовский А. М., Иванов С. В., Горбатова М. А. Экологические (корреляционные) исследования в здравоохранении // Наука и Здравоохранение. 2015. № 5. С. 5-18.

Grijbovski A. M., Ivanov S. V., Gorbatova M. A. Ecological (correlation) studies in health sciences. *Nauka i Zdravooxranenie* [Science & Healthcare]. 2015, 5, pp. 5-18.

Гржибовский А. М., Иванов С. В., Горбатова М. А. Денсаулық сақтаудағы көлденең экологиялық (корреляциялық) зерттеулер // Ғылым және Денсаулық сақтау. 2015. № 5. Б. 5-18.

Целью настоящей статьи является формирование у начинающего исследователя ключевых представлений об экологических (корреляционных) исследованиях, области их применения в медицине, методологии проведения и корректном анализе полученных результатов.

#### **Экологические (корреляционные) исследования в научной и врачебной практике.**

Экологические (корреляционные) исследования (ecological studies) относятся к категории наблюдательных исследований, то есть при проведении научных изысканий исследователь собирает данные путем простого наблюдения событий в их естественном течении, без активного вмешательства в происходящие события [11, 2, 9]. Экологические исследования также являются аналитическими, так как они направлены на выяснение причин, лежащих в основе изучаемых явлений [11, 3].

Следует отметить принципиальные особенности экологических исследований, которые отличают их от других видов эпидемиологических исследований [11, 30, 13, 18]:

1. Объектом наблюдения являются не отдельные лица, а группы лиц, объединенные по какому-либо ключевому признаку (например, проживающие на территории одного района, города, области или государства).

2. Отсутствует исходное разделение изучаемой популяции на основную и контрольную группы.

3. Источником данных для исследования часто является официальная медико-статистическая информация о первичной и общей заболеваемости, социально-демографических характеристиках популяции, распространенности среди населения факторов риска и т.д.

4. Статистическая обработка полученных данных проводится с помощью корреляционного или регрессионного анализа.

5. В ходе исследования проводится оценка воздействия экологических, социальных, демографических и прочих факторов на здоровье популяции, но не на здоровье отдельных индивидуумов.

Экологические исследования в настоящее время намного менее популярны, чем, например, в 1950-60-е годы, когда они послужили основой для многочисленных научных гипотез, которые в дальнейшем проверялись в ходе последующих исследований.

Например, экологические исследования сыграли существенную роль в изучении факторов риска развития онкологических заболеваний. Международные экологические исследования диетических предпочтений населения разных стран позволили обнаружить, что ряд продуктов животного

происхождения, а также сахар, подсластители и некоторые жиры могут выступать факторами риска развития онкопатологии, в то время как продукты растительного происхождения могут иметь онкопротективное значение [16].

Экологическое исследование, проведенное в Японии, позволило сделать вывод о том, что рост заболеваемости болезнью Альцгеймера в Японии в период с 1985 до 2007 года, возможно, связан с переходом населения Японии от традиционных японских диетических предпочтений к западно-европейской диете [17].

В научной практике экологические исследования служат источником ценной информации для формирования исследовательских гипотез о возможном влиянии различных факторов на формирование заболеваемости и смертности групп населения.

Например, в результате экологического исследования, проведенного в 18-ти развитых странах и посвященного изучению смертности от ишемической болезни сердца (ИБС), была обнаружена обратная связь между смертностью от ИБС и потреблением вина среди населения различных стран. Была выдвинута гипотеза о профилактической роли вина в отношении смерти от ИБС, которая в дальнейшем была изучена уже на индивидуальном уровне: было выяснено, что употребление вина повышает концентрацию в крови липопротеидов высокой плотности, которые оказывают кардиопротективное действие [29, 11].

**Методология проведения экологических (корреляционных) исследований.**

Схема экологического исследования представлена на рисунке 1.



**Рис. 1. Схема экологического исследования.**

Массив данных для экологического исследования, как правило, извлекается из официальных статистических сведений, характеризующих изучаемую популяцию: из регистров заболеваемости, социально-демографических показателей населения, действующих на данную популяцию природных и антропогенных факторов риска и прочее.

В основе статистической обработки данных экологических исследований лежит корреляционный анализ.

Термин «корреляция» был впервые введен Ж. Кювье в 1806 году, математическое обоснование метода было предложено О. Браве, а в 1886 г. Ф. Гальтон применил его для биомедицинских исследований [10]. Корреляционный анализ используется для определения силы и направления связи между переменными.

Существуют несколько коэффициентов корреляции, которые имеют разные способы расчета [5, 4, 1, 7, 14, 8], однако наиболее популярными являются следующие:

1. Коэффициент корреляции Пирсона ( $r$ ) – параметрический.

2. Коэффициент корреляции Спирмена ( $\rho$ ) – непараметрический.

Коэффициент корреляции Пирсона наиболее широко используется в экологических исследованиях, поэтому необходимо указать условия его применения [5, 6]:

1. Обе переменные являются количественными и непрерывными.

2. Как минимум один из изучаемых признаков, а лучше оба, имеют нормальное (Гауссово) распределение.

3. Зависимость между переменными носит линейный характер.

4. Вариабельность одной переменной не зависит от значения другой переменной, то есть разброс значений одной из переменных будет примерно одинаковым для всех значений другой переменной (условие гомоскедастичности).

5. Независимость участников исследования друг от друга.

6. Парность наблюдений (значение показателя здоровья и значения показателя, характеризующего действующий фактор, изучаются у одних и тех же участников исследования).

7. Объем выборки не менее 25 наблюдений [15].

8. Репрезентативность выборки.

Если условия применения коэффициента корреляции Пирсона не выполняются, следует использовать непараметрические коэффициенты корреляции, например, коэффициент корреляции Спирмена, расчет которого основан на использовании не исходных значений признаков, а их рангов [5, 1, 8]. Если в данной ситуации применять коэффициент корреляции Пирсона, полученные результаты будут искаженными, а выводы – сомнительными. Следует заметить, что если условия применения коэффициента корреляции Пирсона выполняются, то можно использовать и непараметрические коэффициенты, но мы рекомендуем в этом случае использовать именно коэффициент корреляции Пирсона, так как он обладает большей статистической мощностью по сравнению с непараметрическими коэффициентами корреляции [5].

Недостатком коэффициента корреляции Пирсона является его высокая чувствительность к наличию «выскакивающих» величин (выбросов). Даже небольшое количество выбросов может значительно исказить значение коэффициента и буквально «уничтожить» зависимость между величинами, поэтому всегда следует анализировать выбросы и выяснять, не является ли какой-либо выброс следствием ошибки ввода данных.

Значение любого коэффициента корреляции интерпретируется согласно статистическим пособиям [5, 10, 8] следующим образом:

–  $r = 1$  – наличие полной (функциональной) связи между признаками.

–  $r \geq 0,7$  – наличие сильной связи между признаками.

–  $0,3 < r < 0,7$  – наличие связи между признаками средней силы.

–  $r \leq 0,3$  – наличие слабой связи между признаками.

–  $r = 0$  – отсутствие связи между признаками.

При обратной связи, соответственно, коэффициент корреляции будет иметь отрицательное значение, но интерпретация его значения остается той же. Таким образом, спектр возможных значений коэффициента корреляции находится в пределах от -1 до 1.

Следует учесть, что данная классификация носит во многом условный характер, поэтому для понимания степени «тесноты» связи между признаками используется коэффициент детерминации, который рассчитывается как коэффициент корреляции в квадрате ( $r^2$ ). Коэффициент детерминации показывает, какую долю вариабельности одного из изучаемых признаков способно объяснить влияние другого признака. Например, даже при значении коэффициента корреляции Пирсона  $r=0,8$ , одна переменная может отвечать только за 64% вариабельности другой переменной.

Для исследователя важно учитывать, что в результате корреляционного анализа невозможно установить причинно-следственные связи, поэтому неправомерно делать выводы о влиянии одного явления на другое на основании одного лишь корреляционного анализа данных. Исследователю не следует путать

корреляционную и причинно-следственную связь. Юмористическим примером такого заблуждения служит интерпретация сильной положительной корреляционной связи между количеством гнезд аистов и количеством новорожденных в Копенгагене в ранние послевоенные годы, так как данный результат не может служить доказательством того, что детей приносят аисты [12].

Другим важным аспектом интерпретации результатов корреляционного анализа служит то, что выявленные связи между двумя признаками могут быть обусловлены влиянием некоего третьего фактора. Например, было выяснено, что вероятность рождения у женщины ребенка с синдромом Дауна тесно коррелирует с количеством родов у матери до момента настоящей беременности. Но это вовсе не означает, что количество предшествующих родов влияет на риск рождения больного ребенка. Очевидно, что в данном случае имеется третий фактор, который связан с обеими переменными – возраст женщины. Поэтому между вероятностью рождения ребенка с синдромом Дауна имеется корреляционная, но ни в коей мере не причинно-следственная связь [5].

В более сложных случаях, когда одновременно изучается воздействие многих факторов, применяются различные варианты регрессионного анализа. Например, казахстанскими исследователями недавно в международной научной литературе было опубликовано сравнительно большое количество экологических исследований с использованием многомерных методов статистического анализа данных по теме воздействия климатических факторов на здоровье населения [19-28].

Возвращаясь к рассмотрению методологии экологических (корреляционных) исследований, необходимо выделить их основные преимущества:

1. Простота выполнения и отсутствие существенных финансовых затрат (при условии доступа к необходимой информации).
2. Возможность изучения воздействия большого спектра факторов на здоровье больших популяций.
3. Оптимальность использования с целью изучения гомогенных воздействий в случае, когда эти воздействия проще

определить и измерить на популяционном, а не на индивидуальном уровне.

4. Возможность применения для изучения редких заболеваний.

5. Возможность предварительной проверки исследовательских гипотез.

Наряду с преимуществами, экологические исследования имеют и недостатки:

1. Отсутствие возможности проследить связи между воздействием и исходом, то есть невозможность прямого применения результатов исследования на индивидуальном уровне (так называемая «экологическая ошибочность»).

2. Невозможность определения причинно-следственных связей.

3. Проблема доступа к необходимым данным в некоторых странах и регионах.

4. Результаты экологических исследований могут иметь ограниченную валидность (достоверность) из-за особенностей сбора данных на популяционном уровне.

#### **Алгоритм проведения и анализа результатов экологического (корреляционного) исследования.**

Для создания у начинающего исследователя представлений о непосредственном проведении данного вида исследований, в качестве рабочего примера проведем исследование с использованием реальных эпидемиологических данных, источником которых является электронный ресурс [www.garpminder.org](http://www.garpminder.org).

Цель данного исследования – оценить связь между государственными расходами на здравоохранение и средней продолжительностью жизни населения (по состоянию на 2010 г.).

Факторным признаком в данном случае выступает уровень государственных расходов на здравоохранение, а изучаемым результативным признаком – средняя продолжительность ожидаемой жизни. В исследование включены страны, имеющие относительно сходный уровень достатка населения и расположенные на карте мира относительно недалеко друг от друга.

Массив имеющихся данных представлен в таблице 1.

Таблица 1.

**Государственное финансирование здравоохранения и средняя продолжительность жизни в различных странах (по состоянию на 2010 г.).**

| № п/п | Страна     | Государственное финансирование здравоохранения (USD на 1 человека) | Средняя продолжительность жизни населения (лет) |
|-------|------------|--|---|
| 1.    | Албания    | 577  | 77  |
| 2.    | Беларусь   | 786  | 70  |
| 3.    | Болгария   | 947  | 73  |
| 4.    | Венгрия    | 1469   | 74  |
| 5.    | Греция     | 2853   | 80  |
| 6.    | Грузия     | 522  | 74  |
| 7.    | Испания    | 3027   | 82  |
| 8.    | Кипр       | 1842   | 79  |
| 9.    | Латвия     | 1093   | 72  |
| 10.   | Литва      | 1299   | 72  |
| 11.   | Македония  | 791  | 75  |
| 12.   | Мальта     | 2261   | 79  |
| 13.   | Молдова    | 360  | 69  |
| 14.   | Польша     | 1476   | 76  |
| 15.   | Португалия | 2818   | 79  |
| 16.   | Румыния    | 811  | 73  |
| 17.   | Сербия     | 1169   | 74  |
| 18.   | Словакия   | 2818   | 79  |
| 19.   | Словения   | 2060   | 75  |
| 20.   | Узбекистан | 184  | 68  |
| 21.   | Украина    | 519  | 68  |
| 22.   | Хорватия   | 1514   | 77  |
| 23.   | Черногория | 1155   | 74  |
| 24.   | Чехия      | 2051   | 77  |
| 25.   | Эстония    | 1226   | 74  |

На основании данной таблицы с 25-ю наблюдениями ( $n=25$ ) проведем корреляционный анализ имеющихся данных.

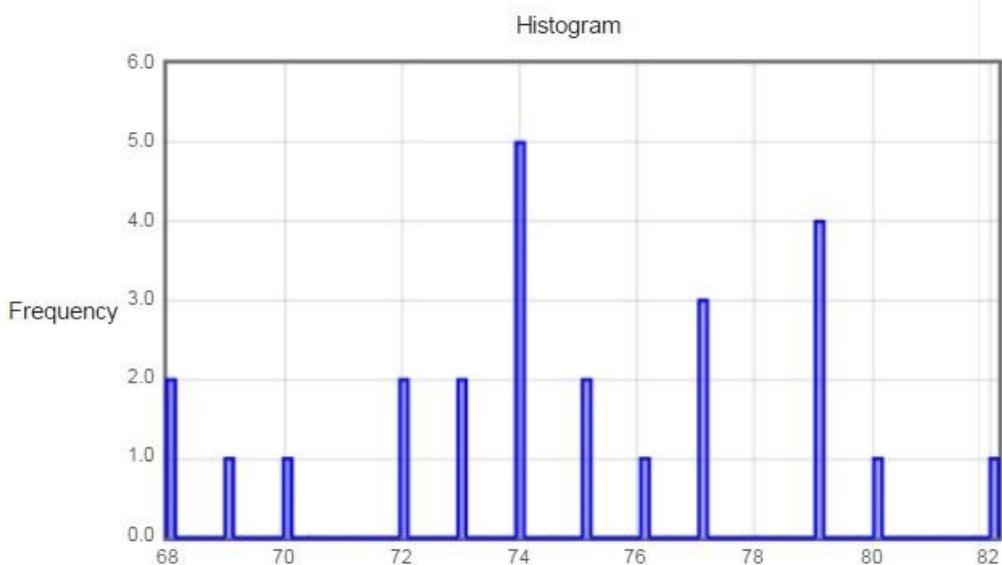
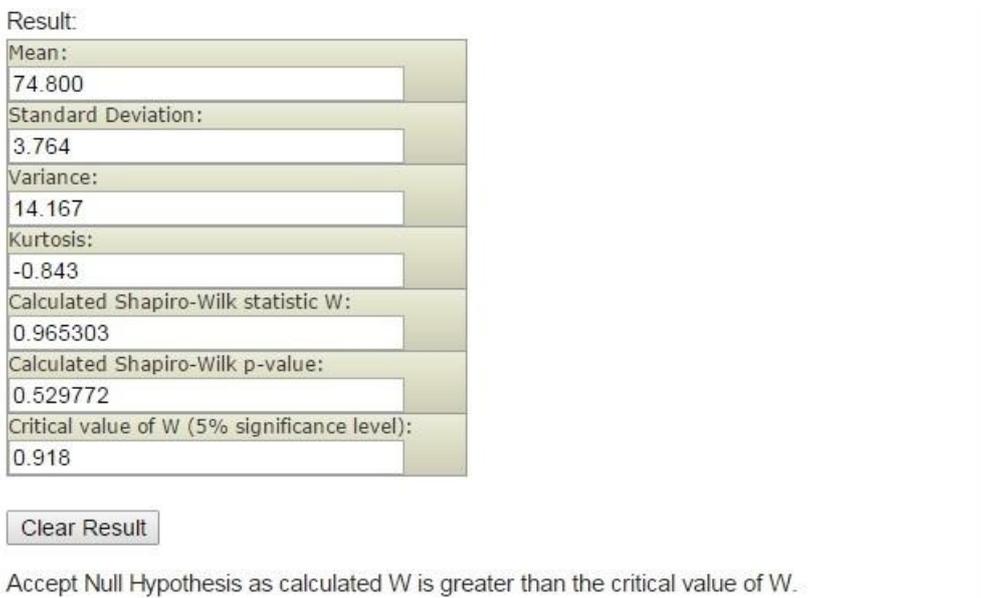
В процессе статистической обработки данных в статье авторами сознательно будут представляться исключительно бесплатные программные продукты – онлайн-калькуляторы. Авторы считают, что для начинающего исследователя важно минимизировать расходы на проведение научной работы, в том числе на закупку статистических программ, таких как «SPSS», «Statistica» и других, которые, разумеется, представляют более удобный функционал для работы с данными.

Перед тем, как приступить к корреляционному анализу, необходимо проверить, удовлетворяют ли подлежащие обработке данные (количественные и непрерывные) условиям применения коэффициента корреляции Пирсона.

В первую очередь следует проверить тип распределения данных. Проверку распределения можно провести с использованием онлайн-калькулятора для расчета критерия Шапиро-Уилка [5] <http://scistatcalc.blogspot.ru/2013/10/shapiro-wilk-test-calculator.html>. Данный критерий используется для проверки распределения небольших выборок (в отличие от критерия Колмогорова-Смирнова) и интерпретируется следующим образом: если полученный в результате расчета критерия уровень статистической значимости меньше 0,05, то распределение считается отличным от нормального (гауссовского), и в этом случае требуется решать вопрос об использовании непараметрических критериев. Вторым, более простым способом проверки распределения является построение и рассмотрение гистограмм или квантильных диаграмм.

Представленный онлайн-калькулятор требует ввода данных по каждому из вариационных рядов, либо загрузки данных в формате MS Excel, что менее трудоемко.

Результаты расчета критерия Шапиро-Уилка для вариационного ряда показателей средней продолжительности жизни представлены на рисунке 2.



**Рис. 2. Проверка типа распределения с помощью онлайн-калькулятора <http://scistatcalc.blogspot.ru/2013/10/shapiro-wilk-test-calculator.html>**

Рассчитанная статистическая значимость критерия («Calculated Shapiro-Wilk p-value») равна 0,530 (значение критерия  $W=0,965$ ), то есть на формальных основаниях мы можем принять нулевую гипотезу об отсутствии различий между нашим распределением и нормальным распределением, поскольку полученная статистическая значимость больше критического значения  $p=0,05$ .

Следует заметить, что представленная гистограмма не сильно похожа на нормальное распределение. Данный факт связан с тем, что при небольшом числе наблюдений маловероятно получить классическую гистограмму нормального распределения в форме колокола. Поэтому, в данном случае полагаем, что нет оснований считать данное распределение отличным от нормального.

Оценка типа распределения вариационного ряда финансирования здравоохранения показывает значение критерия  $W=0,931$  ( $p=0,093$ ). В данном случае распределение формально также не отличается от нормального, хотя тот факт, что значение  $p$  близко к критическому, должен привлечь внимание исследователя. Таким образом, основное условие расчета коэффициента корреляции Пирсона выполняется.

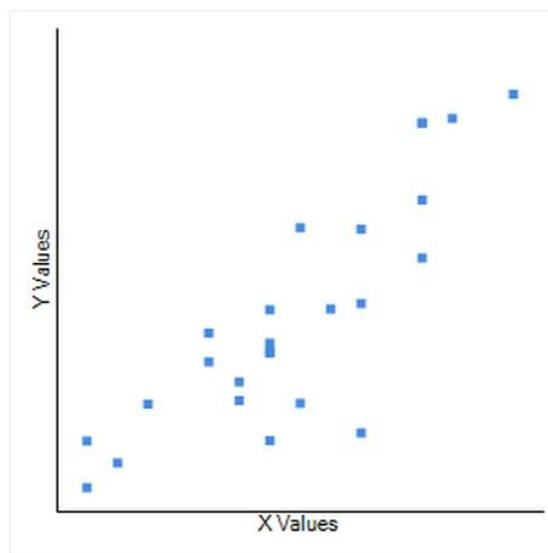
Линейность зависимости между переменными и наличие гомоскедастичности можно проверить путем построения диаграммы рассеяния (скаттерограммы), в том числе с помощью MS Excel или других программ (скаттерограмма будет представлена ниже). Условия независимости

участников исследования (государств), парности наблюдений и прочие условия также выполняются, поэтому в данном случае правомерно проводить корреляционный анализ с расчетом коэффициента корреляции Пирсона.

Для решения этой задачи воспользуемся следующим онлайн-калькулятором, размещенным на сайте: <http://www.socscistatistics.com/tests/pearson/Default2.aspx>.

Способ ввода данных и схематичный график корреляционной связи (скаттерограмма) между расходами на здравоохранение («X Values») и средней продолжительностью жизни («Y Values») представлены на рисунке 3.

| X Values | Y Values |
|----------|----------|
| 80       | 2853     |
| 79       | 2818     |
| 79       | 2818     |
| 79       | 2261     |
| 77       | 2051     |
| 75       | 2060     |
| 77       | 1514     |
| 76       | 1476     |
| 74       | 1469     |
| 74       | 1226     |
| 68       | 184      |
| 82       | 3027     |
| 72       | 1093     |
| 73       | 947      |
| 70       | 786      |
| 68       | 519      |
| 69       | 360      |
| 74       | 1169     |
| 75       | 791      |
| 74       | 522      |
| 72       | 1299     |



The value of R is 0.8446. This is a strong positive correlation, which means that high X variable scores go with high Y variable scores (and vice versa).

The value of  $R^2$ , the coefficient of determination, is 0.7133.

**Рис. 3. Ввод данных, скаттерограмма и результаты расчета коэффициента корреляции Пирсона с помощью онлайн-калькулятора <http://www.socscistatistics.com/tests/pearson/Default2.aspx>**

На представленной скаттерограмме видно, что имеется тенденция увеличения продолжительности жизни с увеличением расходов на здравоохранение. Онлайн-калькулятор также представляет результаты расчета коэффициента корреляции Пирсона «The value of R is 0.8446». Таким образом, значение коэффициента корреляции Пирсона составляет  $r = 0,84$ , что говорит о сильной положительной корреляционной связи между

продолжительностью жизни и финансированием здравоохранения.

Программа также рассчитывает значение коэффициента детерминации  $r^2 = 0,84^2 = 0,71$ . Данное значение говорит о том, что 71% вариабельности продолжительности жизни может быть связан с государственным финансированием здравоохранения.

Помимо значения коэффициента корреляции можно рассчитать статистическую

значимость данного значения с помощью другого онлайн-калькулятора, расположенного на том же электронном ресурсе: <http://www.socscistatistics.com/pvalues/pearsondistribution.aspx>.

Ввод данных и результаты расчета представлены на рисунке 4 (следует учесть, что в данных калькуляторах десятичные разряды отделяются от целых чисел не запятыми, а точками).

R Score:

N:

Significance Level:

- 0.01
- 0.05
- 0.10

The P-Value is < 0.00001. The result is significant at  $p < 0.05$ .

**Рис. 4. Расчет уровня статистической значимости для коэффициента корреляции Пирсона с помощью онлайн-калькулятора <http://www.socscistatistics.com/pvalues/pearsondistribution.aspx>**

Полученное значение статистической значимости ( $p < 0,00001$ ), меньшее, чем 0,05, говорит о том, что полученное значение коэффициента корреляции Пирсона статистически значимо (то есть статистически значимо отличается от нуля). Однако не стоит в публикациях обозначать более трех знаков после запятой, поэтому запись  $p < 0,001$  для данного примера будет наиболее корректной.

Поскольку целью исследования является определение параметра, характеризующего генеральную совокупность, следует рассчитать также и 95% доверительный интервал (confidence interval – CI) для значения коэффициента корреляции Пирсона. Для этого воспользуемся очередным онлайн-калькулятором, расположенным на сайте <http://faculty.vassar.edu/lowry/rho.html>. Способ ввода данных и результаты расчета представлены на рисунке 5.

|            |                                   |  |
|------------|-----------------------------------|--|
| <b>r =</b> | <input type="text" value="0.84"/> | <input type="button" value="Reset"/><br><input type="button" value="Calculate"/> |
| <b>n =</b> | <input type="text" value="25"/>   |  |

*0.95 and 0.99 Confidence Intervals of rho*

|      | Lower Limit                        | Upper Limit                        |
|------|------------------------------------|------------------------------------|
| 0.95 | <input type="text" value="0.666"/> | <input type="text" value="0.927"/> |
| 0.99 | <input type="text" value="0.587"/> | <input type="text" value="0.943"/> |

**Рис. 5. Расчет доверительного интервала для коэффициента корреляции Пирсона с помощью онлайн-калькулятора <http://faculty.vassar.edu/lowry/rho.html>**

Таким образом, рассчитанный 95% доверительный интервал для коэффициента корреляции Пирсона составляет от 0,67 до 0,93.

Итак, по результатам данного экологического исследования установлено, что между средней продолжительностью жизни населения и государственным финансированием здравоохранения существует тесная корреляционная связь:  $r = 0,84$  (95% CI 0,67-0,93),  $p = 0,012$ ,  $n = 25$ .

#### Примеры экологических (корреляционных) исследований.

Примером варианта экологического исследования может служить работа,

опубликованная в International Journal of Cancer в 1997 году [31]. В основу исследования были заложены статистические сведения органов здравоохранения Японии о потреблении населением основных нутриентов и продуктов, и данные о смертности населения от онкологических заболеваний за длительный период времени (1955-1993 гг.). В ходе данного исследования проводился анализ связи между состоянием питания населения и смертностью от разных видов рака, причем учитывался и отсроченный (через 10 лет) эффект имеющихся особенностей питания населения. Фрагмент результатов данного исследования представлен в таблице 2.

Таблица 2.

**Значения коэффициентов корреляции между потреблением некоторых нутриентов и продуктов питания и смертностью от рака желудка в Японии (за период 1955-1993 гг.).**

| Нутриенты / продукты питания | Мужчины             |                        | Женщины             |                        |
|------------------------------|---------------------|------------------------|---------------------|------------------------|
|                              | Исходные показатели | Отсроченные показатели | Исходные показатели | Отсроченные показатели |
| Углеводы                     | 0,960               | 0,991                  | 0,968               | 0,991                  |
| Белки                        | -0,751              | -0,908                 | -0,757              | -0,908                 |
| Жиры                         | -0,825              | -0,975                 | -0,839              | -0,974                 |
| Соль                         | 0,647               | 0,951                  | 0,638               | 0,944                  |
| Витамин С                    | -0,746              | -0,911                 | -0,752              | -0,910                 |
| Рис                          | 0,958               | 0,959                  | 0,959               | 0,963                  |
| Молоко                       | -0,896              | -0,993                 | -0,908              | -0,991                 |

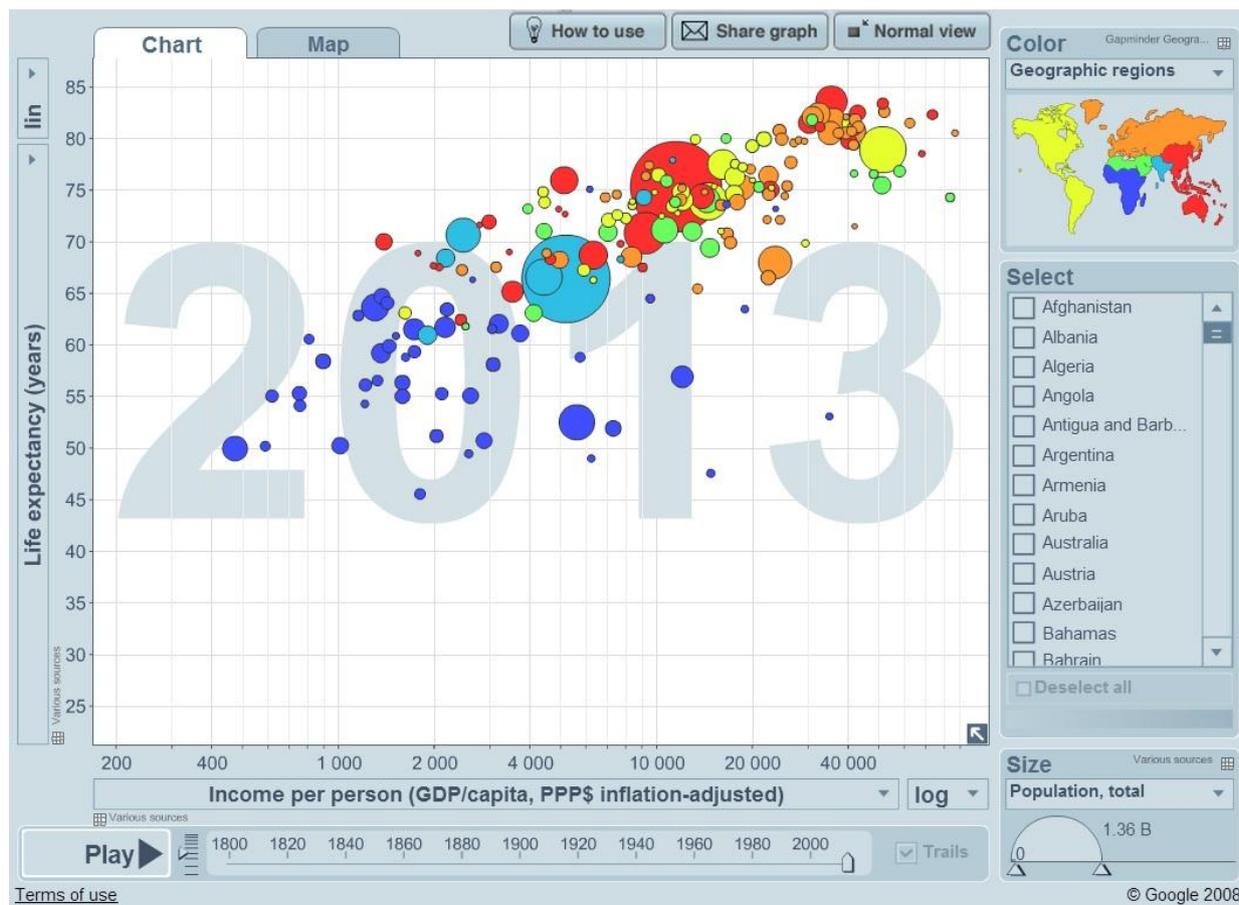
Следует обратить внимание на то, что некоторые коэффициенты корреляции имеют отрицательный знак, что говорит об обратной связи между изучаемыми явлениями. Например, между употреблением в пищу углеводов и смертностью от рака желудка имеется сильная положительная корреляционная связь (чем больше углеводов в питании населения, тем больше смертность от рака желудка), в то время как для жиров и белков эта связь отрицательная (чем больше жиров или белков в питании, тем меньше смертность от рака желудка). Данное исследование продемонстрировало связь между особенностями пищевого статуса населения и смертностью от онкологических заболеваний. Исследователи отметили, что данное исследование является важным для постановки гипотезы о связи диеты и развития рака и установили необходимость дальнейших исследований для подтверждения этиологической роли

различных видов диет и отдельных нутриентов в развитии онкопатологии. Однако, следует помнить, что даже такие высокие коэффициенты корреляции не дают права делать выводы о причинно-следственных связях между фактором риска и исходом.

Для начинающего исследователя, интересующегося вопросами проведения экологических исследований, несомненный интерес представляет упоминавшийся ранее интернет-ресурс [www.garminder.org](http://www.garminder.org).

Данный ресурс в наглядной интерактивной форме представляет полученные из официальных международных источников (Всемирная организация здравоохранения и др.) данные о продолжительности жизни, заболеваемости населения, некоторых факторах риска, социо-демографических показателях и проч.

Основной рабочий экран данного ресурса представлен на рисунке 6.



**Рис. 6. Основной экран интернет-ресурса [www.garminder.org](http://www.garminder.org). Представлена графическая зависимость средней продолжительности жизни населения (в годах) от логарифма среднего дохода населения (в USD на человека) в 2013 году.**

Исследователь может самостоятельно выбрать год наблюдения и популяционные показатели, которые будут располагаться по осям абсцисс и ординат, после чего с помощью сформированной системой скаттерограммы оценить связь между изучаемыми показателями. Каждой из точек скаттерограммы является государство, диаметр точки зависит от населения страны, а цвет – от региона мира, к которому данная страна принадлежит. При наведении курсора на определенную страну на осях абсцисс и ординат появляются соответствующие этой точке количественные данные, на основании которых можно самостоятельно сформировать таблицу и рассчитать коэффициент корреляции для полученного массива данных (разумеется, предварительно убедившись, что представленная зависимость носит линейный характер).

Таким образом, мы рассмотрели один из самых популярных типов исследования – экологические, которые могут быть хорошим отправным пунктом для начинающих исследователей. В заключительном выпуске журнала мы познакомим читателей с экспериментальными исследованиями в здравоохранении.

#### **Литература:**

1. Банержи А. Медицинская статистика понятным языком: вводный курс. М. : Практическая медицина, 2007. 287 с.
2. Власов В. В. Эпидемиология : учебное пособие для вузов. М. : ГЭОТАР-Медиа, 2004. 448 с.
3. Воробьев К. П. Формат современной журнальной публикации по результатам клинического исследования. Часть 3 // Укр. мед. часопис. 2008. №2. С. 150-160.

4. Гланц С. Медико-биологическая статистика. М. : Практика, 1998. 459 с.
5. Гржибовский А. М. Корреляционный анализ // Экология человека. 2008. №9. С. 50-60.
6. Гржибовский А. М. Типы данных, проверка распределения и описательная статистика // Экология человека. 2008. №1. С. 52-58.
7. Доннели-мл. Р. А. Статистика. М. : Астрель : АСТ, 2007. 367 с.
8. Зайцев В.М., Лифляндский В.Г., Маринкин В.И. Прикладная медицинская статистика. СПб. : Фолиант, 2003. 428 с.
9. Зуева Л. П., Яфаев Р. Х. Эпидемиология : учебник. СПб : ООО «Издательство Фолиант», 2008. 752 с.
10. Лакин Г.Ф. Биометрия. М. : Высшая школа, 1990. 350 с.
11. Флетчер Р. Клиническая эпидемиология. Основы доказательной медицины: пер. с англ. / Р. Флетчер, С. Флетчер, Э. Вагнер. М. : Медиа Сфера, 1998. 352 с.
12. Anderson M., Whitcomb P. RSM simplified: optimizing processes using response surface methods for design of experiments. London : Taylor and Francis, 2005.
13. Beaglehole R., Bonita R. Basic epidemiology. World Health Organization, Geneva, 1993.
14. Cleopas T.J. et al. Statistics Applied to Clinical Trials. 4th ed. Springer, 2009.
15. David F. Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples. Cambridge : Cambridge University Press, 1938.
16. Grant W. B. A Multicountry Ecological Study of Cancer Incidence Rates in 2008 with Respect to Various Risk-Modifying Factors // Nutrients. 2014. Vol. 6. P 163-189.
17. Grant W. B. Trends in Diet and Alzheimer's Disease During the Nutrition Transition in Japan and Developing Countries // Journal of Alzheimer's Disease. 2014. Vol. 38. P. 611-620.
18. Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects // Int J Epidemiol. 2001. Vol. 30. P. 1343-1350.
19. Grjibovski A., Nurgaliyeva N., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Cold temperatures and mortality from diseases of circulatory system in Astana, Kazakhstan: a registry-based study // Eur. J. Epidemiol. 2013. Vol. 28. P. S169.
20. Grjibovski A. M., Adilbekova B., Aynabekova B., Kozhakhmetova G., Nurgaliyeva N., Kosbayeva A., Menne B. Cold temperatures and daily counts of ambulance calls for asthma exacerbations in Astana, Kazakhstan in 2006–2010: a registry-based study // Eur. J. Epidemiol. 2013. Vol. 28. P. S75.
21. Grjibovski A. M., Adilbekova B., Kozhakhmetova G., Nurgaliyeva N., Kosbayeva A., Aynabekova B., Menne B. Associations between Cold temperatures and ambulance calls in Astana, Kazakhstan // Eur. J. Public Health. 2013. Vol. 23 (Suppl.1). P. 258.
22. Grjibovski A. M., Kosbayeva A., Menne B. The effect of ambient air temperature and precipitation on monthly counts of salmonellosis in four regions of Kazakhstan, Central Asia, in 2000-2010 // Epidemiol. Infect. 2014. Vol. 142. P. 608-615.
23. Grjibovski A.M., Kosbayeva A. Climate variations and Salmonella infection in Astana, Kazakhstan: a time-series analysis // Eur. J. Public Health. 2012. Vol. 22 (Suppl. 2). P. 162.
24. Grjibovski A. M., Kozhakhmetova G., Nurgaliyeva N., Adilbekova B., Kosbayeva A., Menne B. Air temperature and suicides in Astana, Kazakhstan: a time-series analysis // Eur. J. Public Health. 2013. Vol. 23 (Suppl. 1). P. 258.
25. Grjibovski A. M., Nurgaliyeva N., Adilbekova B., Kozhakhmetova G., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Associations between high summer temperatures and cerebrovascular mortality in Astana, Kazakhstan: a time-series analysis // Eur. J. Public Health. 2013. Vol. 23 (Suppl. 1). P. 258-259.
26. Grjibovski A. M., Nurgaliyeva N., Kosbayeva A., Menne B. No association between temperature and deaths from cardiovascular and cerebrovascular diseases during the cold season in Astana, Kazakhstan – the second coldest capital in the world // Int. J. Circumpol. Health. 2012. Vol. 71. P. 19769.
27. Grjibovski A. M., Nurgaliyeva N., Kosbayeva A., Sharbakov A., Seysembekov T., Menne B. Effect of high temperatures on daily counts of mortality from diseases of circulatory system in Astana, Kazakhstan // Medicina (Kaunas). 2012. Vol. 48. P. 640-646.

28. Grijbovski A. M., Nurgaliyeva N., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Cold temperatures and mortality from diseases of circulatory system in Astana, Kazakhstan // *Eur. J. Epidemiol.* 2013. Vol. 28. P. 169.

29. Leger A. S., Cochrane A. L., Moore F. Factors associated with cardiac mortality in developed countries with particular reference to the consumption of wine // *Lancet.* 1979. Vol. 1. P. 107-1020.

30. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods // *Annual Review of Public Health.* 1995. Vol. 16. P. 61-81.

31. Tominaga S., Kuroishi T. An ecological study on diet/nutrition and cancer in Japan // *Int. J. Cancer.* 1997. Vol. 2-6 (Suppl. 10). P. 2-6.

#### References:

1. Banerzhi A. *Meditinskaya statistika ponyatnyim yazykom: vvodnyy kurs. [Introduction to medical statistics].* M. : *Prakticheskaya meditsina*, 2007. P. 287. [in Russian]

2. Vlasov V. V. *Epidemiologiya : uchebnoe posobie dlya vuzov. [Epidemiology: the textbook]* M. : *GEOTAR-Media*, 2004. P. 448. [in Russian]

3. Vorob'ev K. P. Format sovremennoy zhurnalnoy publikatsii po rezul'tatam klinicheskogo issledovaniya. Chast 3. [The modern format of journal publications on the results of clinical studies. Part 3] // *Ukr. med. Chasopis [Ukrainian medical journal].* 2008. No. 2. PP. 150-160. [in Russian]

4. Glants S. *Mediko-biologicheskaya statistika [The biomedical statistics].* M. : *Praktika*, 1998. PP. 459. [in Russian]

5. Grijbovski A. M. *Correlatsionnij analiz [Correlation analysis] // Ekologiya cheloveka [Human Ecology].* 2008. No. 9. PP. 50-60. [in Russian]

6. Grijbovski A. M. *Tipy dannyh, proverka raspredeleniya i opisatel'naya statistika [Types of data, distributions and descriptive statistics] // Ekologiya cheloveka [Human Ecology].* 2008. No. 1. PP. 52-58. [in Russian]

7. *Donneli-m. R. A. Statistika.* M. : *Astrel* : *AST*, 2007. P.367. [in Russian]

8. Lakin G. F. *Biometriya [Biometria].* M. : *Vyshshaya shkola*, 1990. 350 p. [in Russian]

9. Zueva L. P., Yafaev R. H. *Epidemiologiya : uchebnik [Epidemiology: the textbook].* SPb : *OOO «Izdatelstvo Foliant»*, 2008. P. 752 . [in Russian]

10. Zajcev V. M., Lifyandskij V. G., Marinkin V.I. *Prikladnaya medicinskaya statistika [Applied medical statistics].* SPb. : *Foliant*, 2003. 428 p. [in Russian]

11. Fletcher R. et al. *Klinicheskaya ehpidemiologiya. Osnovy dokazatel'noj mediciny [Clinical epidemiology. Basics of the evidence-based medicine] / R. Fletchtr, C. Fletcher, E. Vagner. M. : Media Sphere*, 1998. 352 p. [in Russian]

12. Anderson M., Whitcomb P. *RSM simplified: optimizing processes using response surface methods for design of experiments.* London : *Taylor and Francis*, 2005.

13. Beaglehole R., Bonita R. *Basic epidemiology.* World Health Organization, Geneva, 1993.

14. Cleopas T. J. et al. *Statistics Applied to Clinical Trials.* 4th ed. Springer, 2009.

15. David F. *Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples.* Cambridge : *Cambridge University Press*, 1938.

16. Grant W. B. *A Multi-country Ecological Study of Cancer Incidence Rates in 2008 with Respect to Various Risk-Modifying Factors // Nutrients.* 2014. Vol. 6. P 163-189.

17. Grant W. B. *Trends in Diet and Alzheimer's Disease During the Nutrition Transition in Japan and Developing Countries // Journal of Alzheimer's Disease.* 2014. Vol. 38. P. 611-620.

18. Greenland S. *Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects // International Journal of Epidemiology.* 2001. Vol. 30. P. 1343-1350.

19. Grijbovski A., Nurgaliyeva N., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Cold temperatures and mortality from diseases of circulatory system in Astana, Kazakhstan: a registry-based study // *Eur. J. Epidemiol.* 2013. Vol. 28. P. S169.

20. Grijbovski A. M., Adilbekova B., Aynabekova B., Kozhakhmetova G., Nurgaliyeva N., Kosbayeva A., Menne B. Cold temperatures and daily counts of ambulance calls for asthma exacerbations in Astana, Kazakhstan in 2006–

2010: a registry-based study // Eur. J. Epidemiol. 2013. Vol. 28. P. S75.

21. Grjibovski A. M., Adilbekova B., Kozhakhmetova G., Nurgaliyeva N., Kosbayeva A., Aynabekova B., Menne B. Associations between Cold temperatures and ambulance calls in Astana, Kazakhstan // Eur. J. Public Health. 2013. Vol. 23 (Suppl.1). P. 258.

22. Grjibovski A. M., Kosbayeva A., Menne B. The effect of ambient air temperature and precipitation on monthly counts of salmonellosis in four regions of Kazakhstan, Central Asia, in 2000-2010 // Epidemiol. Infect. 2014. Vol. 142. P. 608-615.

23. Grjibovski A. M., Kosbayeva A. Climate variations and Salmonella infection in Astana, Kazakhstan: a time-series analysis // Eur. J. Public Health. 2012. Vol. 22 (Suppl. 2). P. 162.

24. Grjibovski A. M., Kozhakhmetova G., Nurgaliyeva N., Adilbekova B., Kosbayeva A., Menne B. Air temperature and suicides in Astana, Kazakhstan: a time-series analysis // Eur. J. Public Health. 2013. Vol. 23 (Suppl. 1). P. 258.

25. Grjibovski A. M., Nurgaliyeva N., Adilbekova B., Kozhakhmetova G., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Associations between high summer temperatures and cerebrovascular mortality in Astana, Kazakhstan: a time-series analysis. // Eur. J. Public Health. 2013. Vol. 23 (Suppl. 1). P. 258-259.

26. Grjibovski A. M., Nurgaliyeva N., Kosbayeva A., Menne B. No association between temperature and deaths from cardiovascular and cerebrovascular diseases during the cold season in Astana, Kazakhstan – the second coldest capital in the world // Int. J. Circumpol. Health. 2012. Vol. 71. P. 19769.

27. Grjibovski A. M., Nurgaliyeva N., Kosbayeva A., Sharbakov A., Seysembekov T., Menne B. Effect of high temperatures on daily counts of mortality from diseases of circulatory system in Astana, Kazakhstan // Medicina (Kaunas). 2012. Vol. 48. P. 640-646.

28. Grjibovski A. M., Nurgaliyeva N., Sharbakov A., Seysembekov T., Kosbayeva A., Menne B. Cold temperatures and mortality from diseases of circulatory system in Astana, Kazakhstan // Eur. J. Epidemiol. 2013. Vol. 28. P. S169.

29. Leger A. S., Cochrane A. L., Moore F. Factors associated with cardiac mortality in developed countries with particular reference to the consumption of wine // Lancet. 1979. Vol. 1. P. 107-1020.

30. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods // Annual Review of Public Health. 1995. Vol. 16. P. 61-81.

31. Tominaga S., Kuroishi T. An ecological study on diet/nutrition and cancer in Japan // Int. J. Cancer. 1997. Vol. 2-6 (Suppl. 10). P. 2-6.

#### **Контактная информация:**

**Гржибовский Андрей Мечиславович** – доктор медицины, магистр международного общественного здравоохранения, Старший советник Национального Института Общественного Здравоохранения, г. Осло, Норвегия; Директор Архангельской международной школы общественного здоровья, г. Архангельск, Россия; Профессор Международного Казахско-Турецкого Университета им. Х.А. Ясыви, г. Туркестан, Казахстан.

**Почтовый адрес:** INFA, Nasjonalt folkehelseinstitutt, Postboks 4404 Nydalen, 0403 Oslo, Norway.

**E-mail:** Andrej.Grijbovski@gmail.com

**Телефон:** +4745268913 (Норвегия), +79214717053 (Россия), +77471262965 (Казахстан)