

КОМП'ЮТЕРНІ ЗАСОБИ, МЕРЕЖІ ТА СИСТЕМИ

*A.M. Gupal, O.I. Ostapko,
T.J. Grachova, O.S. Vorobyov*

APPLYING OF COMPUTER APPLICATIONS IN TOOTH DISEASE DATA PROCESSING AND MODELING.

*This work describes applying of
modern computer systems in medical
researches of data processing and
modeling.*

*Рассмотрены результаты при-
менения современных компьютер-
ных средств для обработки ре-
зультатов медицинских исследо-
ваний и статистического моде-
лирования.*

*Розглянуто результати засто-
сування сучасних комп'ютерних
засобів для обробки результатів
медичних досліджень та ста-
стистичного моделювання.*

© А.М. Гупал, О.І. Остапко,
Т.Я. Грачова, О.С. Воробйов,
2009

УДК 681..3

А.М. ГУПАЛ, О.І. ОСТАПКО,
Т.Я. ГРАЧОВА, О.С. ВОРОБЙОВ

КОМП'ЮТЕРНІ ЗАСОБИ В МОДЕЛЮВАННІ ПРОЦЕСІВ СТОМАТОЛОГІЧНОЇ ЗАХВОРЮВАНОСТІ

Вступ. Проаналізувавши вислів історика ма-
тематичної науки К.А. Рибнікова „Матема-
тические модели – объект и орудие труда
математика», можна з впевненістю констату-
вати, що нинішній стан і можливості комп'ю-
терних засобів – техніки і програмного за-
безпечення дозволили цьому об'єкту стати
практичним джерелом реалізації прикладних
аспектів математики. Особливо це характер-
но для медицини. Результати обстежень у по-
єднанні зі статистичними інформаційними
потокami представляють собою підґрунтя
узагальнень і розрахунків з метою вдоскона-
лення роботи як самої медичної галузі, так
водночас, виступають і передумовою появи
нових підходів в області використання ін-
формаційних технологій.

У роботах [1–5] висвітлена проблема впли-
ву дефіциту, надлишку або дисбалансу мік-
роелементів, присутніх у людському органі-
змі, що може призводити до виникнення па-
тологічних станів, а в подальшому, і до пере-
ростання їх в конкретні захворювання. Но-
вим важливим елементом досліджень для
виявлення взаємозв'язків між станом стома-
тологічної захворюваності і хімічним скла-
дом елементів зубної емалі, наприклад, може
стати саме використання елементів матема-
тичного моделювання з використанням ком-
п'ютерних систем на стадії формалізації да-
ного процесу та при аналізі і прогнозуванні
ситуації.

Мета досліджень – визначення можливос-

ті використання комп'ютерних засобів для виявлення наявних взаємозв'язків і взаємодії між ступінню стоматологічної захворюваності і показниками хімічного складу, що опосередковано характеризують екологічну ситуацію у певному регіоні та по країні в цілому.

Предмет досліджень – вивчення та аналіз використання блоку функцій кореляційно-регресивного аналізу присутнього в стандартизованих програмних системах типу SPSS 14 та STATISTICA 6 при моделюванні процесів стоматологічної захворюваності.

Результати досліджень. Кореляційно-регресивний аналіз є одним із основних підходів у виявленні взаємозв'язків між показниками та визначенні тісноти цих зв'язків. Він представляє комплекс методів, за допомогою яких можна визначити вид рівняння для опису взаємозалежностей, провести розрахунки параметрів отриманих рівнянь, а також оцінити адекватність та істинність отриманих з рівнянь математичних моделей. У випадку досліджень по дитячій стоматологічній захворюваності виникає ситуація багатофакторної залежності.

Етап збору й обробки статистичної інформації, який був самим трудомістким, включав у себе проведення прямих медичних замірів, забір матеріалів, клініко-лабораторні аналізи, а також обчислення деяких показників і фіксування результатів у вигляді інформаційних потоків, що передбачали подальшу обробку їх за допомогою універсальних інтегрованих систем на персональному комп'ютері.

У дослідженнях використовувались найважливіші показники, характерні для визначення рівня стоматологічного здоров'я, а також показники хімічного складу емалі зубів та слини в ротовій порожнині дітей різних вікових категорій.

Екологічна група показників – це показники, що відображають стан атмосферного повітря, поверхневих вод та ґрунтів у розрізі їх хімічного складу.

Територіально-адміністративний розріз усіх показників – 4 регіони з різним економіко-природничим та екологічним станом від сільськогосподарського до виробничо-конгломеративного, від еколого-благополучного до майже катастрофічного. Це показники по місту Києву, Хмельницькій та Херсонській областях, а також по місту Маріуполь.

Вся інформаційна база, заснована на вищезгаданих статистичних показниках, разом з висновками спеціалістів являла собою вхідні дані для аналізу з метою виявлення можливостей використання цієї бази для формалізації і представлення у вигляді, який дозволяв подальшу обробку даних за допомогою комп'ютерних технологій.

Вхідні інформаційні потоки зводились до стандартних форм через електронні таблиці EXCEL. Це дало можливість у подальшому використовувати їх для розрахунків по виявленню прямих і обернених кореляційних зв'язків та визначенню коефіцієнтів кореляції, а також побудови рівнянь регресії відразу в двох інтегрованих системах – SPSS та STATISTICA без допоміжних переробок.

Наведемо приклад вхідного інформаційного потоку (табл. 1):

Значення показників: КПВ – кількість уражених карієсом зубів, СРІ – кількість уражених секстантів пародонта, ЗК – кількість секстантів із зубним каменем, РМА – індекс тяжкості гінгівіту, СРЗ – показник рівня стоматологічного здоров'я – це статистично-звітні дані, присутні в медичних закладах.

ТАБЛИЦЯ 1. Показники стоматологічної захворюваності

№ п/п	ППП	КПВ	СРІ	ЗК	РМА	СРЗ
1	Дацюк Богдан	7	5	1	25	20
2	Слізаров Олексій	8	6	4	50	60
3	Кабанова Марія	3	6	4	46,4	70
4	Алексєєв Олексій	0	6	1	29,8	70
5	Жигуліна Христина	6	6	2	35,7	60
6	Капітонова Оксана	7	6	0	21,4	20
7	Коломієць Денис	3	5	3	49,4	70
8	Гавриленко Яна	2	6	1	28,6	70
9	Шумихін Сергій	2	5	1	33,4	70
10	Шубнікова Анастасія	0	6	3	40,5	70
11	Конищев Олександр	2	4	0	28,6	70
12	Чуб Олена	8	6	1	39,3	20

Група показників, представлених в табл. 2, має лабораторно-дослідницький характер. Вони виражають вміст наступних хімічних елементів: Са – кальцію, Р – фосфору, К – калію, Сl – хлору, Na – натрію, Pb – свинцю, Cd – кадмію, Cu – міді, Zn – цинку, Mn – магнію, Fe – заліза, Cr – хрому, Ni – нікелю, Со – кобальту, F – фтору, Sr – стронцію в емалі та ротовій порожнині.

ТАБЛИЦЯ 2. Хімічний склад емалі зубів

№ п/п	ППП	Са	Р	К	Сl	Na	Pb	Cd	Cu
1	Кабанова Марія	28,366	17,906	0,347	1,334	0,28	0,039	0,009	0,024
2	Алексєєв Олексій	27,952	17,843	0,085	0,95	0,271	0,034	0,01	0,034
3	Жигуліна Христина	25,115	14,772	0,215	1,206	0,295	0,031	0,008	0,03
4	Капітонова Оксана	25,521	16,271	0,18	1,112	0,368	0,026	0,005	0,025
5	Коломієць Денис	25,699	16,201	0,27	1,403	0,526	0,023	0,004	0,026
6	Гавриленко Яна	24,331	15,875	0,26	1,307	0,315	0,033	0,006	0,027
7	Шумихін Сергій	26,781	17,993	0,198	1,215	0,421	0,027	0,007	0,029
8	Шубнікова Анастасія	27,882	16,741	0,232	1,421	0,268	0,038	0,004	0,031
9	Конищев Олександр	28,524	16,338	0,273	1,115	0,438	0,041	0,005	0,023
10	Чуб Олена	25,432	17,232	0,224	1,312	0,331	0,024	0,008	0,025

Наведені в табл. 1, 2 фрагменти даних належать до однієї територіальної одиниці – міста Києва. Такі ж дані наявні і ще по трьох регіонах – Хмельницькій та Херсонській областях і місту Маріуполь.

На другому етапі моделювання після вивчення й аналізу був сформований вхідний інформаційний потік, елементи якого, наведені в табл. 3.

ТАБЛИЦЯ 3. Вхідні дані для побудови регресивного рівняння

O	Ca	P	K	Cl	Na	Pb	Cd	Cu	Zn	Mn	Fe	Cr	Ni	Co	F	Sr	КПВ
51,387	28,366	17,906	0,347	1,334	0,28	0,039	0,009	0,024	0,036	0,012	0,028	0,01	0,015	0,007	0,171	0,029	4
52,594	27,952	17,843	0,085	0,95	0,271	0,034	0,01	0,034	0,051	0,012	0,034	0,004	0,013	0,013	0,06	0,04	5
58,201	25,115	14,772	0,215	1,206	0,295	0,031	0,008	0,03	0,029	0,01	0,025	0,003	0,011	0,008	0,005	0,036	4
56,379	25,521	16,271	0,18	1,112	0,368	0,026	0,005	0,025	0,027	0,009	0,024	0,006	0,009	0,007	0,033	0,031	5
55,745	25,699	16,201	0,27	1,403	0,526	0,023	0,004	0,026	0,021	0,003	0,021	0,004	0,006	0,005	0,016	0,027	5
53,831	24,331	15,875	0,26	1,307	0,315	0,033	0,006	0,027	0,033	0,008	0,023	0,005	0,008	0,006	0,012	0,029	6
54,721	26,781	17,993	0,198	1,215	0,421	0,027	0,007	0,029	0,038	0,007	0,025	0,007	0,012	0,008	0,05	0,028	4
53,283	27,882	16,741	0,232	1,421	0,268	0,038	0,004	0,031	0,024	0,011	0,028	0,003	0,01	0,007	0,007	0,039	7
57,323	28,524	16,338	0,273	1,115	0,438	0,041	0,005	0,023	0,027	0,009	0,029	0,003	0,007	0,005	0,005	0,041	5
56,638	25,432	17,232	0,224	1,312	0,331	0,024	0,008	0,025	0,022	0,004	0,031	0,004	0,005	0,009	0,009	0,033	7

Це приклад вхідного інформаційного потоку, що характеризує склад емалі зубів. Подібні дані були присутні і в описі інформаційного масиву по хімічному складу слини в ротовій порожнині дітей.

Використання вхідної інформації для розрахунків регресивних залежностей за допомогою інтегрованої системи SPSS дозволили отримати відразу значення коефіцієнтів кореляції та рівнянь регресивної залежності стоматологічної захворюваності дітей від факторів хімічного складу емалі і слини, за допомогою яких у подальшому можна через підстановку в рівняння опосередкованих показників екологічного стану прогнозувати рівень стоматологічної захворюваності.

Розрахунки проводились з використанням блоку <АНАЛІЗ> за допомогою методу <Backward stepwise> – покрокове виключення змінних.

Розглянемо рівняння регресії, що виражає залежність рівня стоматологічної захворюваності від хімічного складу емалі:

$$\text{КПВ} = 24,228 - 0,404 \cdot \text{O} + 186,851 \cdot \text{Fe} - 226,796 \cdot \text{Co} - 14,655 \cdot \text{F},$$

де O – кисень; Fe – залізо; Co – кобальт; F – фтор.

Модель, виражену цим рівнянням, назвемо моделлю по емалі, тому що дані, на яких вона побудована, це об'єднання інформації по всіх географічних регіонах з урахуванням статистичних і лабораторних показників по емалі зубів.

Аналіз адекватності моделі по емалі. Залежність між відгуком – КПВ і предикаторами O, Fe, Co і F досить значна. Коефіцієнт детермінації R^2 на рівні 70 % ($R^2 = 0.697$).

Значення показників дисперсійного аналізу представлені в блоці ANOVA (*Analysis of Variance*). Наявний ріст значень коефіцієнта Фішера на всіх кроках

розрахунків доводить вибір кращої моделі. Величини стандартизованих регресивних коефіцієнтів *Beta* дозволяють порівнювати вклад кожного предикатора у прогнозовану величину відгуку (КПВ) і ранжовані таким чином: по кисню – 0,629, по залізу – +0,433, по кобальту – – 0,340, по фосфору – – 0,243, причому, від’ємне значення коефіцієнта свідчить про те, що при його збільшенні значення відгуку падають, додатне – зростають.

Власні коефіцієнти кореляції (*Partial Cor*) показують міру впливу одного предикатора на відгук за умови, що інші предикатори залишаються на тому ж рівні. Саме ці показники виражають обґрунтованість виключення того чи іншого показника з моделі. Так, за 13 кроків при роботі методу <Backward stepwise> поступово були виключені змінні Cd, Ca, Ni.

Показник статистики Durbin-Watson, рівний 2,022, характеризує присутність серійної залежності між залишками і свідчить про наявність залежності всередині вхідних даних, за якими розраховувалась модель, що може впливати на адекватність моделі процесу, який вивчався, тобто визначенню впливу певних факторів на рівень захворюваності.

Розглянемо рівняння регресії, отримане з використанням клінічних даних по хімічному складу слини.

$$CPI = 37,8 + 1008,1 \cdot Pb + 12,2 \cdot Cu + 65,7 \cdot Mn - 128,7 \cdot Fe - 10,2 \cdot Ni,$$

де CPI – кількість уражених секстантів пародонта; Pb – свинець; Cu – мідь, Mn – магній; Fe – залізо; Ni – нікель.

Назвемо модель, виражену цим рівнянням, моделлю по слині.

Розрахунки проводились також з використанням блоку <АНАЛІЗ> за допомогою методу <Backward stepwise>.

Аналіз адекватності моделі по слині. Залежність між відгуком – CPI і предикаторами на межі допустимого ($R^2 = 67\%$). Пояснюється це не досить значною кількістю даних, які було можливо використати в розрахунках. Певне значення тут має дисперсійний аналіз ANOVA. У цій частині представлено ріст значень коефіцієнта Фішера по всіх кроках розрахунків.

Власні коефіцієнти кореляції (*Partial Cor*) виразили обґрунтованість виключення того чи іншого показника з моделі. Так, при використанні методу <Backward stepwise> поступово були виключені змінні Cd, Zn, Cr.

Саме у випадку не зовсім повної адекватності найважливіше значення має показник статистики Durbin-Watson. У даному випадку ця характеристика рівна 1,22.

Дещо низькі значення коефіцієнта детермінації, отримані внаслідок моделювання, спонукали до спроби вдосконалення даного регресивного рівняння. Так, за допомогою інструментарію *Multiple Regression* пакету STATISTICA при використанні розширеної опції – *stepwise or ridge regression* – гребенева регресія отримано таке рівняння регресивної залежності:

$$КПВ = 12,007 - 0,174 \cdot O + 139,422 \cdot Fe - 190,115 \cdot Co - 6,818 \cdot F + 100,901 \cdot Sr - 85,179 \cdot Pb.$$

Результати розрахунків з використанням файлу даних, який вміщує вхідні дані по чотирьох регіонах, отриманий через установку динамічного обміну да-

ними між файлом формату "Лист Microsoft Excel" та таблицею STATISTICA формату "sta" наведені в табл. 4:

ТАБЛИЦЯ 4. Параметри рівняння гребеневої регресії

N=40	Beta	Std.Err. of Beta	B	Std.Err. of B	t(33)	p-level
Intercept	0	0	12,007	6,04172	1,98733	0,055237
Sr	0,688174	0,192016	100,901	28,15361	3,58393	0,001077
Co	-0,284879	0,093037	-190,115	62,08838	-3,06201	0,004352
Fe	0,323007	0,091729	139,422	39,59382	3,52131	0,001279
Pb	-0,350965	0,155285	-85,179	37,68735	-2,26014	0,030535
O	-0,271352	0,160806	-0,174	0,10334	-1,68745	0,100948
F	-0,113230	0,100340	-6,818	6,04154	-1,12847	0,267261

Вдалося довести коефіцієнт детермінації до рівня майже 80 % ($R^2 = 0.78224126$). Всі коефіцієнти рівняння (крім коефіцієнта при змінній F) статистично значимі – скрізь рівень значимості ($p - level$) нижче 0,1.

Практичною перевіркою якості такої моделі можуть служити результати, отримані при підстановці замість змінних їх фізичних значень. Так, при розрахунках прогнозу з використанням даних по вмісту хімічних елементів в емалі постійних зубів маємо повний збіг прогнозу по місту Києву, а саме – фактична кількість уражених карієсом зубів показник КПВ = 5.

Висновки. Внаслідок проведених досліджень вдалося побудувати математичні моделі залежності стоматологічної захворюваності дітей від показників хімічного складу за допомогою комп'ютерних систем аналізу й обробки даних. Доведена теоретична адекватність моделей, виражених рівняннями регресії. Використання комп'ютерних засобів при моделюванні процесів стоматологічної захворюваності дозволило систематизувати наявні статистичні дані та створити апарат аналізу і прогнозування ситуації.

1. Агулова Л.П., Опалинская А.М., Кирьянов В.С. Характерные особенности реакций объектов различной природы, чувствительных к изменению космофизических факторов и действию искусственных слабых электромагнитных полей // Биофизические и климатические аспекты гелиобиологии. – Л.: Наука, 1989. – С. 160 – 181.
2. Буштуева К.А., Случанко И.С. Методы и критерии оценки состояния здоровья населения в связи с загрязнением окружающей среды. – М.: Медицина, 1979. – 68 с.
3. Звиняцковский Л.И. Комплекс показателей здоровья населения как критерий качества окружающей среды // Научные основы гигиены окружающей среды. – Усть-Каменогорск, 1986. – С. 89 – 92.
4. Славин М.Б. Методы системного анализа в медицинских исследованиях. – М.: Медицина, 1989. – 304 с.
5. Біостатистика / В.Ф. Москаленко, О.П. Гульчій, М.В. Голубчиков та ін. – К.: Книга плюс, 2009. – 184 с.

Отримано 18.05.2009