СПОСІБ ПРОГНОЗУВАННЯ КІЛЬКОСТІ ЗАХВОРЮВАНЬ У ДІТЕЙ ЖИТОМИРЩИНИ

Лойко Є.Є., Ротштейн О.П., Катєльніков Д.І., Лойко Л.С.

ВДМУ, ВДТУ, ВДПУ, м.Вінниця

Прогнозування або передбачення кількості захворювань на рівні вибраної адміністративної одиниці (місто, район, регіон тощо) або спеціалізованого лікувального закладу є суттєвим моментом планування профілактично-лікувальних програм різної спрямованості та контролю за їх виконанням.

Відомі в науці методи вирішення задач прогнозування різних дисретних послідовностей (сукупність значень в фіксовані моменти), які виникають не тільки в медицині, але й фізиці, техніці, соціології тощо, потребують багаторічних досліджень та великої вибірки експериментальних даних [5]. Найбільш поширеними є методи, побудовані на базі ймовірнісно-статистичного апарату [2]. Але їх використання потребує значної кількості експериментальних і статистичних даних, які не завжди вдається зібрати в умовах подій, що відбулися недавно, наприклад, аварія на ЧАЕС, повінь в Закарпатті тощо.

Крім цього, нетривалість прогнозування дискретних послідовностей обумовлена необхідністю екстраполяції даних про минуле на майбутній час з урахуванням невідомого закону про явище, що лежить в основі процесу, який вивчається.

Останнім часом в задачах прогнозування широкої популярності набувають штучні нейронні мережі. Вони розглядаються, як схожі до людського мозку універсальні функціональні моделі, що навчаються розпізнавати невідомі закономірності [10]. Але, як і випадку з ймовірнісно-статистичними методами, навчання нейронної мережі потребує великої вибірки експериментальних даних [10].

Академік В. М. Глушков писав: Унікальною властивістю людмького зору являється здібність сприймати закномірнісні лінійні ділянки в експериментальних залежностях" [3].

$$\ldots x_1^{i-1}\}\ \{x_1^{i}\ x_2^{i}\ x_3\ x_4^{i}\}\ \{x_1^{i+1}\ \ldots$$

високосний рік

де i – номер чотирирічного періоду;

x^{i}_{1} – кількість захворювань за два роки до високосного (опреділяючий рік);

x^{i}_{2} – кількість захворювань за один рік до високосного (перехідний рік);

x^{i}_{3} – кількість захврювань у високосному році (рік прогресування тенденції);

x^{i}_{4} – кількість захворювань в наступному році після високосного (розрішаючий рік або рік хронізації).

Введені позначення будуть використані в подальшому при формуванні закономірностей, необхідних для прогнозування.

Більше того, людині властиво бачити і нелінійні тенденції, котрі простежуються в експериментальних залежностях. При цьому важливу роль відіграє можливість описання таких залежностей з допомогою природньо-мовних виразів, як наприклад, кількість захворювання "не змінюється", "дещо збільшується або зменшується", "сильно збільшується", "після підйому настає спад" тощо. Формальні виразки таких залежностей можливі з використанням засобів нечіткої логіки [4, 7].

На відміну від класичної двохканальної логіки, в якій елемент може належати (1) або не належати (0) деякій множині, нечітка логіка дозволяє оцінювати належність елемента проміжними (між 0 та 1) значеннями [4].

При цьому "нечітка" абсолютно не еквівалентно терміну "не точна", оскільки за допомогою настройки нечіткої моделі на експериментальні дані можна досягнути скільки завгодно високої точності моделювання [4, 7, 8]. При цьому однією з основних переваг нечітких моделей являється те, що вони використовують експортно-експериментальну інформацію малих вибірок. При цьому експертна частина – це вислів типу "ЯКЩО–ТОДІ". Це природньо-мовний вислів, побачений експертом в наборі експериментальних даних, які використовуються для адаптації нечіткої меоделі [4, 7, 8].

Аналізуючи 6183 історії хвороби дітей з апендикулярним перитонітом, що лікувалися в Житомирській та Вінницькій [9] обласних лікарнях на протязі 1982–1997 років, нами помічені певні закономірності – наявність чотирирічних періодів, в яких на третьому місті стоїть високосний рік.

Апендикулярний перитоніт вибраний в якості моделі захврювання, як найхарактерніша і найпоширеніша патологія дитячого віку [1, 6].

Ці періоди і роки умовно позначаємо так:

Закономірності, що вдається спостерігати серед експериментальних даних (рис , табл. 1) легко записати у вигляді чотирьох експертних висловлювань природньою мовою. Ці висловлювання є правилами “ЯКЩО–ТОДІ”, які пов’язують кількості захворювань в і-ому та (і+1) – му періодах:

$$
F_1: \left|
\begin{array}{l}
\text{ЯКЩО } x_1^i = \textit{середня} \\
\text{І } x_2^i = \textit{нижче середньої}, \\
\text{ТОДІ } x_3^i = \textit{середня} \\
\text{ЯКЩО } x_1^i = \textit{вище середньої} \\
\text{І } x_2^i = \textit{середня}, \\
\text{ТОДІ } x_3^i = \textit{висока} \\
\text{ЯКЩО } x_1^i = \textit{вище середньої} \\
\text{І } x_2^i = \textit{вище середньої}, \\
\text{ТОДІ } x_3^i = \textit{висока} \\
\text{ЯКЩО } x_1^i = \textit{нижче середньої} \\
\text{І } x_2^i = \textit{середня}, \\
\text{ТОДІ } x_3^i = \textit{нижче середньої}
\end{array}
\right.
\qquad
F_2: \left|
\begin{array}{l}
\text{ЯКЩО } x_1^i = \textit{середня} \\
\text{І } x_2^i = \textit{нижче середньої}, \\
\text{ТОДІ } x_4^i = \textit{нижче середньої} \\
\text{ЯКЩО } x_1^i = \textit{вище середньої} \\
\text{І } x_2^i = \textit{середня}, \\
\text{ТОДІ } x_4^i = \textit{середня} \\
\text{ЯКЩО } x_1^i = \textit{вище середньої} \\
\text{І } x_2^i = \textit{вище середньої}, \\
\text{ТОДІ } x_4^i = \textit{низька} \\
\text{ЯКЩО } x_1^i = \textit{нижче середньої} \\
\text{І } x_2^i = \textit{середня}, \\
\text{ТОДІ } x_4^i = \textit{висока}
\end{array}
\right.
$$

$$
F_3: \left|
\begin{array}{l}
\text{ЯКЩО } x_4^i = \textit{нижче середньої}, \\
\text{ТОДІ } x_1^{i+1} = \textit{вище середньої} \\
\text{ЯКЩО } x_4^i = \textit{середня}, \\
\text{ТОДІ } x_1^{i+1} = \textit{вище середньої} \\
\text{ЯКЩО } x_4^i = \textit{низька}, \\
\text{ТОДІ } x_1^{i+1} = \textit{нижче середньої}
\end{array}
\right.
\qquad
F_4: \left|
\begin{array}{l}
\text{ЯКЩО } x_4^i = \textit{нижче середньої} \\
\text{І } x_1^{i+1} = \textit{вище середньої}, \\
\text{ТОДІ } x_2^{i+1} = \textit{середня} \\
\text{ЯКЩО } x_4^i = \textit{середня} \\
\text{І } x_1^{i+1} = \textit{вище середньої}, \\
\text{ТОДІ } x_2^{i+1} = \textit{вище середньої} \\
\text{ЯКЩО } x_4^i = \textit{низька} \\
\text{І } x_1^{i+1} = \textit{нижче середньої}, \\
\text{ТОДІ } x_2^{i+1} = \textit{середня}
\end{array}
\right.
$$

Таблиця 1

Розподілення кількості захворювань по рокам (Житомир)

Рік	1982	1983	1984	1985	1986	1987	1988	1989
Кількість випадків	180	165	190	165	245	183	252	203
Рік	1990	1991	1992	1993	1994	1995	1996	1997
Кількість випадків	229	215	252	153	164	186	164	253

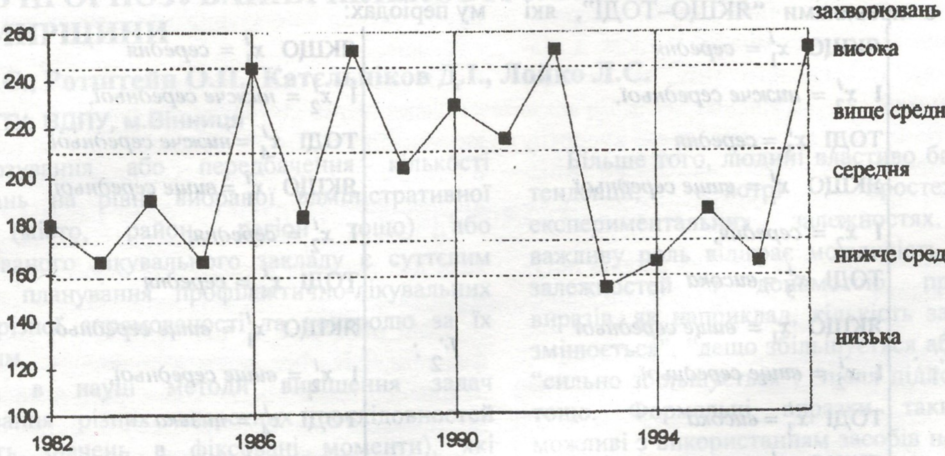


Рис. 1. Динаміка зміни кількості захворювань (Житомир)

Використовуючи сформовані вище правила, по двом першим рокам *i*-го періоду можна прогнозувати на чотири роки вперед: на два останні роки *i*-го періоду і на два перші роки наступного (*i*+1)-го періоду.

Для використання експертно-лінгвістичних висловлювань F1 / F4 скористаємося апаратом теорії нечітких множин [4]. Згідно з цією теорією лінгвістичні оцінки "низький", "нижче середнього" та інші формалізуються за допомогою функцій належності. Задамо ці функції у вигляді [8]:

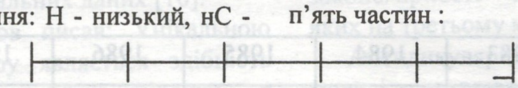
$$\mu^T(x) = \frac{1}{1+\left(\frac{x-b}{c}\right)^2},$$

де b і c – параметри, які спочатку обираються експертно, а потім настроюються на експериментальні дані,

$\mu^T(x)$ - число в діапазоні [0,1], яке характеризує суб'єктивну міру відповідності значення *x* лінгвістичній оцінці *T*.

Експертно обрані параметри *b* і *c* для різних лінгвістичних оцінок, що використані в правилах $F_1 \div F_4$, представлені в табл. 2. Отримані при цьому функції належності показані на рис. 2. На цьому рисунку використані наступні скорочення: Н - низький, нС - нижче середнього, С - середній, вС - вище середгього, В - високий.

Позначимо $[\underline{x}, \overline{x}]$ діапазон можливих значень кількості захворювань. Розіб'ємо цей діапазон на п'ять частин :



Таблиця 2

Параметри функцій належності до настройки (Житомир)

Лінгвістичні оцінки змінних $x_1^i \div x_4^i$	Параметри до настройки		Параметри після настройки	
	b	*c*	*b*	*c*
низька (Н)	154	11	150	0,616
нижче середнього (нС)	166	19	161,293	0,949
середня (С)	184,5	30	184,961	0,394
вище середньої (вС)	236	30	246,787	2,16
висока (В)	250	15	254,311	15

Тоді, користуючись методом нечітких логічних рівнянь [7] і висловлюваннями $F_1 \div F_4$, можливо записати модель прогнозування в явному вигляді:

$$F_1:\begin{cases} x_3^i=\dfrac{x_1\mu^{нC}\left(x_3^i\right)+x_2\mu^{C}\left(x_3^i\right)+x_4\mu^{B}\left(x_3^i\right)}{\mu^{нC}\left(x_3^i\right)+\mu^{C}\left(x_3^i\right)+\mu^{B}\left(x_3^i\right)}, \\ \mu^{нC}\left(x_3^i\right)=\min\left(\mu^{нC}\left(x_1^i\right),\mu^{C}\left(x_2^i\right)\right), \\ \mu^{C}\left(x_3^i\right)=\min\left(\mu^{C}\left(x_1^i\right),\mu^{нC}\left(x_2^i\right)\right), \\ \mu^{B}\left(x_3^i\right)=\max\begin{pmatrix}\min\left(\mu^{вC}\left(x_1^i\right),\mu^{C}\left(x_2^i\right)\right), \\ \min\left(\mu^{вC}\left(x_1^i\right),\mu^{вC}\left(x_2^i\right)\right)\end{pmatrix} \end{cases}$$

$$F_2:\begin{cases} x_4^i=\dfrac{\underline{x}\mu^{H}\left(x_4^i\right)+x_1\mu^{нC}\left(x_4^i\right)+x_2\mu^{C}\left(x_4^i\right)+x_4\mu^{B}\left(x_4^i\right)}{\mu^{H}\left(x_4^i\right)+\mu^{нC}\left(x_4^i\right)+\mu^{C}\left(x_4^i\right)+\mu^{B}\left(x_4^i\right)} \\ \mu^{H}\left(x_4^i\right)=min\left(\mu^{вC}\left(x_1^i\right),\mu^{вC}\left(x_2^i\right)\right) \\ \mu^{нC}\left(x_4^i\right)=min\left(\mu^{C}\left(x_1^i\right),\mu^{нC}\left(x_2^i\right)\right) \\ \mu^{C}\left(x_4^i\right)=min\left(\mu^{вC}\left(x_1^i\right),\mu^{C}\left(x_2^i\right)\right) \\ \mu^{B}\left(x_4^i\right)=min\left(\mu^{нC}\left(x_1^i\right),\mu^{C}\left(x_2^i\right)\right) \end{cases}$$

$$F_3:\begin{cases} x_1^{i+1}=\dfrac{x_1\mu^{нC}\left(x_1^{i+1}\right)+x_3\mu^{вC}\left(x_1^{i+1}\right)}{\mu^{нC}\left(x_1^{i+1}\right)+\mu^{вC}\left(x_1^{i+1}\right)} \\ \mu^{нC}\left(x_1^{i+1}\right)=\mu^{H}\left(x_4^i\right) \\ \mu^{вC}\left(x_1^{i+1}\right)=max\left(\mu^{нC}\left(x_4^i\right),\mu^{C}\left(x_4^i\right)\right) \end{cases},$$

$$F_4:\begin{cases} x_2^{i+1}=\dfrac{x_2\mu^{C}\left(x_2^{i+1}\right)+x_3\mu^{вC}\left(x_2^{i+1}\right)}{\mu^{C}\left(x_2^{i+1}\right)+\mu^{вC}\left(x_2^{i+1}\right)}. \\ \mu^{C}\left(x_2^{i+1}\right)=max\begin{pmatrix}min\left(\mu^{нC}\left(x_4^i\right),\mu^{вC}\left(x_1^{i+1}\right)\right) \\ min\left(\mu^{H}\left(x_4^i\right),\mu^{нC}\left(x_1^{i+1}\right)\right)\end{pmatrix} \\ \mu^{вC}\left(x_2^{i+1}\right)=min\left(\mu^{C}\left(x_4^i\right),\mu^{вC}\left(x_1^{i+1}\right)\right) \end{cases}$$

За допомогою отриманої моделі можна грубо прогнозувати кількість захворювань, як показано на рис. 3. Для підвищення точності прогнозу необхідно перейти до настройки моделі.

Задача настройки полягає в підборі таких параметрів b і c функцій належності лінгвістичних оцінок (рис.2), які забезпечують мінімум розбіжностей між теоретичними і експериментальними кількостями захворювань. Згідно з методом найменших квадратів ця задача формулюється так:

Найти параметри b і c, що

$$\sum_{i=1}^{N}\left(x_3^i-\hat{x}_3^i\right)^2+\sum_{i=1}^{N}\left(x_4^i-\hat{x}_4^i\right)^2+\sum_{i=1}^{N-1}\left(x_1^{i+1}-\hat{x}_1^{i+1}\right)^2+\sum_{i=1}^{N-1}\left(x_2^{i+1}-\hat{x}_2^{i+1}\right)^2=\min_{b,c},$$

де $x_3^i, x_4^i, x_1^{i+1}, x_2^{i+1}$ - прогнозні кількості захворювань, які залежать від параметрів b і c функцій належності; $\hat{x}_3^{i}, \hat{x}_4^{i}, \hat{x}_1^{i+1}, \hat{x}_2^{i+1}$ - експериментальні кількості захворювань; N - число періодів, які використовуються для настройки моделі.

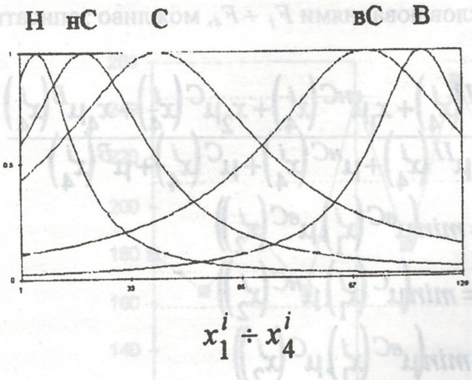


Рис. 2. Функції належності лінгвистичних оцінок до настройки (Житомир)

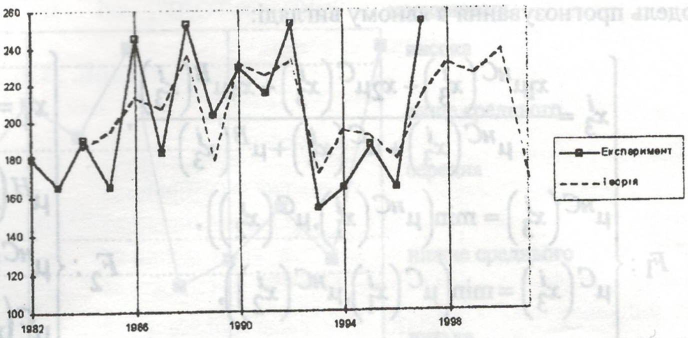


Рис. 3. Співставлення експериментальних даних і моделі прогнозу до настройки (Житомир)

Після настройки моделі прогнозування при $N = 3$, що відповідає використанню даних за 1982-1993 роки, отримані параметри функцій належності, показані в табл. 2. Функції належності після настройки приведені на рис. 4. В даному випадку: $x = 153$, $x_1 = 158$, $x_2 = 173$, $x_3 = 210$, $x_4 = 246$, $x = 253$.

Використання настроєних функцій належності дає модель прогнозування, що досить близька до експериментальних даних (рис. 5).

Оскількі експериментальні значення кількості захворювань в 1994 - 1997 роках не використовувались при настройці моделі, то досить близькі результати теорії і експерименту в цих роках свідчать про високу якість побудованої моделі прогнозування.

Інформацію про похибку моделі прогнозу, а також прогноз кількості захворювань апендикулярним перитонітом до 2001 представлено в табл. 3.

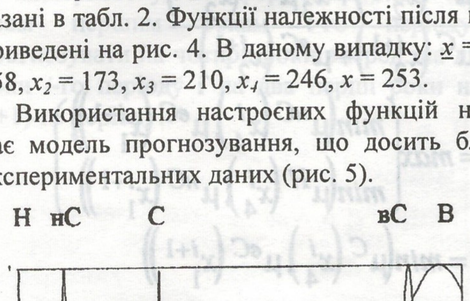


Рис. 4. Функції належності лінгвістичних оцінок після настройки (Житомир)

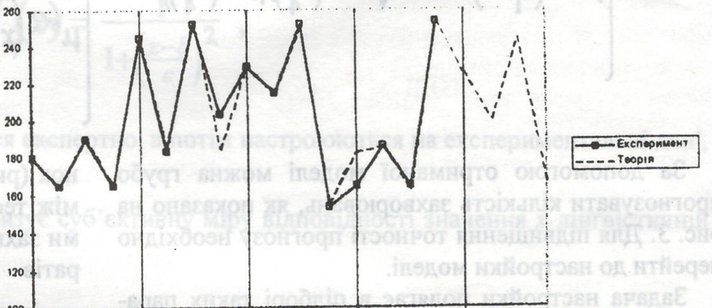


Рис. 5. Співставлення експериментальних даних і моделі прогнозу після настройки (Житомир)

Таблиця 3

Експериментальні та теоретичні кількості захворювань (Житомир)

Рік	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
Експеримент	180	165	190	165	245	183	252	203	229	215
Теорія			190	166	245	185	254	184	229	215
Похибка			0	1	0	2	2	19	0	0
Рік	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Експеримент	252	153	164	186	164	253				
Теорія	250	155	182	185	164	253	225	200	243	166
Похибка	2	2	18	1	2	0				

ВИСНОВКИ.

1.На прикладі апендикулярного перитоніту у дітей Житомирської області опробований новий спосіб прогнозування кількості захворювань для організації профілактично-лікувальних заходів різної спрямованості.

2.В порівнянні з статистичними методами проведений спосіб прогнозування на основі експертної інформації про закономірності, які вдається побачити в наявних експериментальних даних у вигляді природньо-мовних правил "ЯКЩО-ТОДІ", що формалізуються за допомогою нечіткої логіки, дозволяє будувати моделі прогнозу при відносно малих вибірках експериментальних даних з урахуванням невідомого закону про явище, що лежить в основі процесу.

3. Даний спосіб прогнозування на прикладі кількості захворювань характеризується тим, що на основі показника часу (чотирирічного періоду, в якому на третьому місці стоїть високосний рік) діє можливість по дво першим рокам першого чотирирічного періоду прогнозувати кількість захворювань на чотири роки вперед: на два останні (третій і четвертий) роки першого періоду і на два перші роки наступного чотирирічного періоду.

4. В разі змін в межах чотирирічного періоду, як структурної основи моделі прогнозування, принципи прогнозування збурігаються, оскільки вони визначаються нечіткими правилами "ЯКЩО-ТОДІ", а параметри періоду можуть уточнюватися (настроюватися) по мірі накопичення нових експериментальних даних.

ЛІТЕРАТУРА

1. Баиров Г,А, Срочная хирургия детей.- С.-Петербург: Питер-Пресс, 1997.- С. 302, 310.
2. Глушков В.М. Введение в АСУ.- Киев: Техника.- 1974.- 320 с.
3. Глушков В.М. Основы бкзбумажных технологий.- М.: Наука.- 1982.- 278 с.
4. Заде Л.А. Понятие лингвистической переменной и её применению к принятию приближённых решений.- М.: Мир, 1976.- 167 с.
5. Ивахненко А.Г., Лапа В.Г. Предсказание случайных процессов.- Киев: "Наукова думка", 1971.- 416 с.
6. Исаков Ю.Ф., Степанов Э.П., Красовская Т.В. Абдоминальная хирургия у детей.- М.: Медицина, 1988.- с.222, 240.
7. Ротштейн А.П. Медицинская диагностика на нечёткой логике.- Винница: Континент-ПРИМ, 1996.- 132 с.
8. Ротштейн О.П., Катєльніков Д.І. Ідентифікація нелінійних об'єктів нечіткими базами знань//Вісник ВПІ.- 1997.- №4.- С. 98-103.
9. Ротштейн О.П., Лойко ЄЄ, Катєльніков Д.І., Лойко Л.С. Прогнозування кількості захворювань на базі нечіткої логіки//Вінк ВДМУ.- 1998.- №2.- С. 486-489.
10. Mayoraz Frederic, Thierry Cornu, Laurent Vulliet. Prediction of Slope Movements Using Neural Networks.//Fuzzy Systems&A.I.- 1997.- Vol.4- №1-р. 9-17.

РЕЗЮМЕ

Способ прогнозирования количества заболеваемости у детей житомирщині

Лойко С.Е., Ротштейн О.П., Кательников Д.И., Лойко Л.С.

На примере аппендикулярного перитонита у детей Житомирской области опробован новій способ прогнозирования количества заболеваний для организации профилактически-лечебніх мероприятий разной направленности.

SUMMARY

Method of prognostication of the quantity of diseases in children of Zhytomir region

S.E.Loiko, O.P.Rotshtein, D.I.Katelnykov, L.S.Loiko

On the example of appendicular peritonitis in children of Zhytomir region a new method of prediction of the number of diseases is approbated for the organisation of preventive-managment measures.