

Оригінальна методика / Original method

Journal of PERIOPERATIVE MEDICINE

Медицина Періопераційна • Периоперационная Медицина

• www.perioperative.org.ua •

Математичне моделювання прогнозування ускладненого перебігу гострого апендициту у дітей

Коноплицький В. С., Мотигін В. В., Якименко О. Г., Дмитрієв Д. В., Михальчук Т. І.

Вінницький національний медичний університет ім. М. І. Пирогова

Резюме. Ендогенна інтоксикація, що супроводжує ускладнений перебіг гострого апендициту у дітей, є одним із важких ускладнень, яке слугує основною причиною розвитку ускладнень в післяопераційному періоді. Всебічні клінічні спостереження клініцистів свідчать, що, незважаючи на застосування новітніх технологій в операційній техніці, анестезіологічному забезпеченні, кількості ускладнень у післяопераційному періоді не має тенденції до зниження, а основною причиною летальних наслідків при апендикулярному перитоніті є блискавичний розвиток важкого ступеня ендотоксикозу, який виникає внаслідок резорбції токсичних речовин, що утворилися з численних компонентів, у тому числі з патогенної мікрофлори та зруйнованого метаболізму.

Мета. Створення математичної моделі прогнозування ускладненого перебігу гострого апендициту у дітей за допомогою використання функцій регресійного аналізу електронної таблиці EXCEL шляхом апроксимації експериментальних даних.

Матеріали та методи. В роботі використана інформація із 59 МКСХ, які перебували на лікуванні в клініці дитячої хірургії ВНМУ ім. М. І. Пирогова з діагнозом “гострий апендицит”. Дані про хворих були зведені у дві групи: хворі з неускладненим перебігом патології – 24 пацієнти і хворі з ускладненим перебігом гострого апендициту – 35 дітей. Нормою служили дані обстеження 37 здорових дітей. У кожній з трьох груп вивчалися наступні критерії: стать, вік, ліжко-день, тривалість захворювання, пульс, температура тіла, наявність симптомів подразнення очеревини, характер ексудату в черевній порожнині, кількість еритроцитів, гемоглобін, лейкоцитоз, кількісна характеристика формули білої крові, ШОЕ, рівні клітинних критеріїв ендогенної інтоксикації: лейкоцитарного індексу інтоксикації та гематологічного показника інтоксикації. Усі дані визначалися безпосередньо на момент госпіталізації хворого в стаціонар.

Результати. Обговорення. На основі множинної кореляції $x = f(x_1, x_2)$ створена двофакторна математична модель ймовірності ускладненого перебігу гострого апендициту у дітей. Згідно з отриманими результатами дослідження системи трьох рівнянь з трьома невідомими з'ясовано, що якщо показник форми захворювання складає $y = 2,67$ і вище, що в цифровому обчисленні результат потрапляє в зону ускладненого перебігу гострого апендициту і повністю збігається з даними аналізу трьох груп хворих.

Висновки. До чинників, що визначають ускладнений перебіг гострого апендициту у дітей, згідно з розробленою двофакторною моделлю, відносяться величини ЛПІ і ГПІ, які перевершують значення певних визначуваних інтервалів на момент госпіталізації. Значення аналізованих індексів визначалися ступенем вираженості патологічного процесу, тривалістю захворювання, вірулентністю мікрофлори і віком пацієнтів на тлі сформованого синдрому ендогенної інтоксикації.

Ключові слова: діти, гострий апендицит, ускладнений перебіг, модель.

Вступ

Ендогенна інтоксикація, яка супроводжує ускладнений перебіг гострого апендициту (ГА) у дітей, є одним із важких ускладнень, яке служить основною причиною розвитку важких ускладнень у післяопераційному періоді. Всебічні клінічні спостереження клініцистів свідчать, що, незважаючи на застосування новітніх технологій в операційній техніці, анестезіологічному забезпеченні, кількість ускладнень в післяопераційному періоді не має тенденції до суттєвого зниження, а основною причиною летальних наслідків при апендикулярному перитоніті є блискавичний розвиток важкого ступеня ендотоксикозу, який виникає внаслідок резорбції токсичних речовин, що утворилися з численних компонентів, у тому числі з патогенної мікрофлори та зруйнованого метаболізму. Отже, боротьба з ендотоксикозом – одна з головних задач дитячих хірургів та анестезіологів, а прогнозування перебігу захворювання у хворих з гострою хірургічною патологією є важливою складовою частиною їхнього лікування [1, 2]. Від правильної інтерпретації клінічних і лабораторних показників, результатів інструментальних досліджень залежить як тактика лікування хворих з ускладненим перебігом гострого апендициту, так і прогноз [3]. Останнім часом з цією метою все ширше використовують комп'ютерний моніторинг [4]. В останні роки запропоновано багато методів для прогнозування наслідків захворювання, серед яких найчастіше застосовуються наступні прогностичні статистичні методи: оцінювальний метод, емпіричний метод, логістичний регресійний аналіз (система оцінки питомої ваги захворювання APACHE II), багатофакторна логістична регресія та оцінка якості моделі [5, 6].

Найбільше поширення в сучасних медичних дослідженнях знайшло використання кореляційного і регресійного аналізів, розрахунку лінійного та квадратичного рівнянь регресії для кожного індексу, а також коефіцієнта кореляції r між клінічними і лабораторними показниками [1]. Статистичні методи активного математичного прогнозування слугують одним з емпіричних способів одержання математичного опису статистики складних об'єктів дослідження, тобто рівняння зв'язку відгуку об'єкту і незалежних, нормованих вхідних змінних (факторів). При цьому математичний опис подається у вигляді деякого полінома – відрізка ряду Тейлора, у який розкладається невідома залежність в межах основної точки.

Метою роботи було створення математичної моделі прогнозування ускладненого перебігу гострого апендициту у дітей за допомогою використання функцій регресійного аналізу електронної таблиці EXCEL шляхом апроксимації експериментальних даних.

Матеріали та методи

В роботі використана інформація з 59 МКСХ, які перебували на лікуванні в клініці дитячої хірургії ВНМУ ім. М.І.Пирогова з діагнозом ГА. Дані про хворих були зведені у дві групи: хворі з неускладненим перебігом ГА – 24 пацієнти і хворі з ускладненим перебігом ГА – 35 дітей. Усі хворі цих груп підбиралися методом випадкових ви-

бірок для збільшення незалежності результатів. Нормою служили дані обстеження 37 дітей, у яких були відсутні будь які прояви гнійно-запальних захворювань. У кожній з трьох груп було зведено найістотніші стандартні критерії: стать, вік, ліжко-день, тривалість захворювання, пульс, температура тіла, наявність симптомів подразнення очеревини, характер ексудату в черевній порожнині, кількість еритроцитів, гемоглобін, лейкоцитоз, кількісна характеристика формули білої крові, ШОЕ, рівні клітинних критеріїв EI, а саме: лейкоцитарного індексу інтоксикації (ЛІІ) та гематологічного показника інтоксикації (ГПІ). Усі дані визначалися безпосередньо на момент госпіталізації хворого в стаціонар.

Підрахунок ЛІІ (Я. Я. Кальф-Каліф, 1947), який являє собою розрахункову величину із лейкограми, визначався за формулою:

$$ЛІІ = (4Mi + 3Ю + 2П + С) \times (Пл + 1) \\ (Мон + Лім) \times (E + 1),$$

де Mi – міелоцити; $Ю$ – юні; $П$ – паличкоядерні; $С$ – сегментоядерні; $Пл$ – плазматичні; $Мон$ – моноцити; $Лім$ – лімфоцити; E – еозинофіли.

Метод простий, не потребує для свого проведення коштовного обладнання або спеціально підготовлених фахівців, вирізняється високою інформативністю при різних патологічних станах, доступний для кожної клінічної лабораторії [7]. Нормальним вважають рівень ЛІІ, який складає 0,8–1,0. Субнормальні цифри індексу характеризують млявий перебіг хронічної інфекції, високі – гостре запалення [8].

ГПІ – розрахунковий інтегральний показник, який розраховується за формулою:

$$ГПІ = ЛІІ \times Кл \times Кш,$$

де ЛІІ – лейкоцитарний індекс інтоксикації, Кл – коефіцієнт лейкоцитозу; Кш – коефіцієнт швидкості осідання еритроцитів. Обидва коефіцієнти знаходять за розрахунковими таблицями авторів (табл. 1, 2) [9].

ГПІ враховує порушення в лейкограмі, кількість лейкоцитів та швидкість осідання еритроцитів, що дозволяє проводити комплексну оцінку загальної картини крові на різних етапах захворювання.

За нормальний рівень ГПІ показник індексу автори пропонують вважати рівень $0,62 \pm 0,068$, який був отриманий ними при дослідженні здорових донорів. При наявності в організмі патологічного процесу відбувається збільшення величини індексу, а при зниженні EI його рівень поступово нормалізувався, але порівняно з ЛІІ даний індекс нормалізувався значно пізніше, що свідчить про його більш виражену специфічність відповідно до рівня EI організму.

Результати. Обговорення

В процесі роботи були створені одно- та двофакторні математичні моделі. За даними аналізу МКСХ виявлено, що форма перебігу ГА залежить від чотирьох параметрів,

Таблиця 1. Поправочний коефіцієнт на лейкоцитоз (Кл)

Кількість лейкоцитів в 1 мл	Кл	Кількість лейкоцитів в 1 мл	Кл
1–2×10 ⁹	0,2	20–21×10 ⁹	2,4
2–3×10 ⁹	0,4	21–22×10 ⁹	2,6
3–4×10 ⁹	0,6	22–23×10 ⁹	2,8
4–5×10 ⁹	0,8	23–24×10 ⁹	3,0
5–8×10 ⁹	1,0	24–25×10 ⁹	3,2
8–9×10 ⁹	1,1	25–26×10 ⁹	3,4
9–10×10 ⁹	1,2	27–27×10 ⁹	3,6
10–11×10 ⁹	1,3	27–28×10 ⁹	3,8
11–12×10 ⁹	1,4	28–29×10 ⁹	4,0
12–13×10 ⁹	1,5	29–30×10 ⁹	4,2
13–14×10 ⁹	1,6	30–31×10 ⁹	4,4
14–15×10 ⁹	1,7	31–32×10 ⁹	4,6
15–16×10 ⁹	1,8	32–33×10 ⁹	4,8
16–17×10 ⁹	1,9	33–34×10 ⁹	5,0
17–18×10 ⁹	2,0	34–35×10 ⁹	5,2
18–19×10 ⁹	2,1	35–36×10 ⁹	5,4
19–20×10 ⁹	2,2	36–37×10 ⁹	5,6

Таблиця 2. Поправочний коефіцієнт на ШОЕ (Кш)

Швидкість осідання еритроцитів (мм/год)	Кш	Швидкість осідання еритроцитів (мм/год)	Кш
До 5	0,9	36–40	1,7
6–15	1,0	41–45	1,9
16–20	1,1	46–50	2,1
21–25	1,2	51–55	2,3
26–30	1,3	56–60	2,5
31–35	1,5	51–65	2,7

а саме: пульсу, температури тіла, ЛПІ та ГПІ на момент госпіталізації. На нашу думку, величина зміни середніх значень пульсу для трьох груп не є чинником, що впливає на форму захворювання. Це пояснюється тим, що визначення частоти пульсу на момент госпіталізації відбувається без урахування загального стану дитини, тривалості захворювання і нормальних значень показника у віковому аспекті. Крім того, підрахунок пульсу відбувається так само без обліку компенсаторних можливостей серцево-судинної системи дитячого організму, яка ще не обтяжена хронічними серцево-судинними захворюваннями.

Елементи регресійного аналізу

Застосування статистичних методів активного математичного прогнозування є одним із емпіричних шляхів отримання математичного опису статистики складних об'єктів дослідження, тобто рівняння зв'язку відгуку об'єкта і незалежних керованих нормованих вхідних змінних (факторів) $\bar{z}^T = (z_1, z_2, \dots, z_n)$. При цьому математичний опис представляється у вигляді деякого повного відрізка ряду Тейлора, в який розкладається невідома залежність в межах основної точки $z^* \rightarrow_0$:

$$M\{y\} = \varphi(z_1, z_2, \dots, z_n) = \beta_0 + \sum_{i=1}^n \beta_i z_i + \sum_{\substack{i,j=1 \\ i < j}}^n \beta_{ij} z_i z_j + \sum_{i=1}^n \beta_{ii} z_i^2 + \dots, \quad 1.1$$

де $\beta_i = \frac{\partial \varphi}{\partial z_i} \Big|_{\bar{z}=\bar{z}_0}$; $\beta_{ij} = \frac{\partial^2 \varphi}{\partial z_i \partial z_j} \Big|_{\bar{z}=\bar{z}_0}$; $\beta_{ii} = \frac{1}{2} \frac{\partial^2 \varphi}{\partial z_i^2} \Big|_{\bar{z}=\bar{z}_0}$ – теоретичні коефіцієнти. Внаслідок присутності некерованих і навіть неконтрольованих чинників зміна величини y носить випадковий характер, тому функціональна залежність $\varphi(\bar{z})$ не дає точного зв'язку між керованими чинниками і математичним очікуванням випадкової величини y

$$M\{y_g\} = \varphi(\bar{z}_g) \quad 1.2$$

де $\bar{z}_g^T = (z_{1g}, z_{2g}, \dots, z_{ng})$ – g -та точка простору незалежних керованих факторів (факторного простору). В такому випадку за результатами прогнозування можна відшукати рівняння регресії у формі деякого повного:

$$y = b_0 + \sum_{i=1}^n b_i z_i + \sum_{\substack{i,j=1 \\ i < j}}^n b_{ij} z_i z_j + \sum_{i=1}^n b_{ii} z_i^2 + \dots, \quad 1.3$$

де вибіркові коефіцієнти регресії є лише оцінками для теоретичних коефіцієнтів, відповідно $\beta_0, \beta_i, \beta_{ij}, \dots, a_y$ – оцінкою для $M\{y\}$. Нехай $\bar{z}_g = (g = 1.2 \dots N)$ – точки факторного простору, в яких проводиться моделювання. Тоді задача пошуку оцінок коефіцієнтів рівняння регресії (1.3) за результатами досвіду в N точках факторного простору є типовим завданням множинного регресійного аналізу в тому випадку, якщо виконуються наступні рекомендації:

1. Результати спостережень y_1, y_2, \dots, y_N відгуку в N точках факторного простору являють собою незалежні, нормально розподілені випадкові величини, тобто на них діють нормально розподілені випадкові перешкоди ε з нульовим математичним очікуванням $M\{\varepsilon\} = 0$.

2. Дисперсії $\sigma^2\{y_g\}$ ($g = 1, 2, \dots, N$) рівні. Це означає, що отримані при проведенні багаторазових повторних спостережень над величиною y_g в точках \bar{z}_g вибіркові оцінки $S_g^2\{y\}$ однорідні, дисперсія ж $\sigma^2\{y_g\}$ не залежить від математичного очікування $M\{y_g\}$, тобто не відрізняється від дисперсії $\sigma^2\{y_g\}$, отриманої при повторних спостереженнях в будь-якій точці \bar{z}_g факторного простору (відтворення з рівною точністю).

3. Незалежні керовані чинники z_1, z_2, \dots, z_n вимірюються з несуттєво маленькими помилками порівняно з помилкою у визначенні y (мається на увазі вплив їх помилок на

величину у порівняно з впливом некерованих і неконтрольованих чинників ε).

Розглянемо задачу знаходження коефіцієнтів рівняння регресії (1.3) на прикладі рівняння другого порядку з чотирма незалежними факторами; при цьому, звичайно, вся процедура і зроблені висновки можуть бути поширені на рівняння будь-якого ступеня з n незалежними факторами. Перш за все спростимо систему позначень: введемо фіктивну змінну $z_0 = 1$ і позначимо $z_0 = f_0$; $z_1 = f_1$; ...; $z_4 = f_4$; $z_{12} = f_5$; ...; $x_4^2 = f_8$; $z_1 z_2 = f_9$; $z_1 z_2 = f_9$; ...; $z_3 z_4 = f_4$. У новій системі позначень повне число другого ступеня записується як лінійне однорідне рівняння:

$$y = \sum_{j=0}^{14} b_j f_j, \quad 1.4$$

Нехай проводяться досліди в N точках 4-факторного простору x , причому y_g – значення відгуку при досліді в точці z_g ($g = 1, 2, \dots, N$). Коефіцієнти рівняння (5.4) знаходяться на основі методу найменших квадратів, тобто з умови мінімуму суми квадратів відхилень значень відгуку y_g , передбаченим рівнянням (1.4) для умов дослідів в точках z_g , від значень y_g , що спостерігаються, які виходять при досліді в цих точках:

$$\sum_{g=1}^N (y_g - y_g)^2 = \sum_{g=1}^N \left[y_g - \sum_{j=0}^{14} b_j f_{jg} \right]^2, \quad 1.5$$

Оскільки завдання полягає в пошуку значень коефіцієнтів b_j , що мінімізують вираз (1.5), воно вирішується за допомогою системи так званих нормальних рівнянь, отриманих прирівнянням до нуля часткових похідних від квадратичної форми (1.5) за змінними параметрами b_j ($j = 0, 1, 2, \dots, 14$). Система нормальних рівнянь має вигляд

$$\left. \begin{aligned} c_{0'0} b_0 + c_{0'1} b_1 + \dots + c_{0'14} b_{14} &= \alpha_0 \\ c_{1'0} b_0 + c_{1'1} b_1 + \dots + c_{1'14} b_{14} &= \alpha_1 \\ \dots & \\ c_{14'0} b_0 + c_{14'1} b_1 + \dots + c_{14'14} b_{14} &= \alpha_{14} \end{aligned} \right\}, \quad 1.6$$

де

$$C = \{c_{jl}\} = \begin{pmatrix} c_{0'0} & c_{0'1} & \dots & c_{0'14} \\ c_{1'0} & c_{1'1} & \dots & c_{1'14} \\ \dots & \dots & \dots & \dots \\ c_{14'0} & c_{14'1} & \dots & c_{14'14} \end{pmatrix}, \quad 1.7$$

- матриця коефіцієнтів системи (1.6), елементи якої знаходяться в такий спосіб:

$$C_{jl} = \sum_{g=1}^N f_{jg} f_{lg}, \quad 1.8$$

Матриця $F = \{f_{jg}\}$ називається матрицею незалежних змінних, тоді $F^T = \{f_{jg}\}$ – матриця, отримана транспортуванням матриці F , причому:

$$F = \{f_{jg}\} = \begin{pmatrix} f_{1,0} & f_{1,1} & \dots & f_{1,14} \\ f_{2,0} & f_{2,1} & \dots & f_{2,14} \\ \dots & \dots & \dots & \dots \\ f_{N,0} & f_{N,1} & \dots & f_{N,14} \end{pmatrix}, \quad 1.9$$

Можна показати, що $C = F^T F$.

Вільні члени α_j системи нормальних рівнянь визначаються за допомогою рівності:

$$\alpha_{jl} = \sum_{g=1}^N f_{jg} y_g, \quad 1.10$$

Для того щоб система (1.6) мала єдине рішення, необхідно і достатньо, щоб матриця C була незникаючою, тобто її визначник повинен бути відмінним від нуля $|C| \neq 0$. Неважко показати, що ця умова зводиться до умови лінійної незалежності векторів-стовпців матриці F , тобто для того, щоб система (5.6) мала єдине рішення, необхідно і достатньо, щоб вектори-стовпці матриці F були лінійно незалежні. Аналіз рішення системи (1.6), отриманого, наприклад за формулою Крамера:

$$b_l = |C|^{-1} |C_l| \quad (l = 0, 1, 2, \dots, 14), \quad 1.11$$

де $|C_l|$ – визначник, який виходить з $|C|$ при заміні елементів c_{jl} –го стовпчика відповідними вільними членами α_j , показує, що значення коефіцієнтів b_j залежать від кількості членів рівняння регресії, тобто збільшення або зменшення числа членів рівняння впливає на значення коефіцієнтів усіх включених у рівняння членів. Така невизначеність в оцінці коефіцієнтів регресії вкрай ускладнює їх фізичну інтерпретацію. У тому випадку, якщо матриця C діагональна, тобто виконується умова:

$$C_{jl} = \sum_{g=1}^N f_{jg} f_{lg} = 0 \quad \text{при } j \neq l, \quad 1.12$$

система (1.6) розпадається на незалежні нормальні рівняння:

$$\left. \begin{aligned} c_{0'0} b_0 &= \alpha_0 \\ c_{1'1} b_1 &= \alpha_1 \\ \dots & \\ c_{14'14} b_{14} &= \alpha_{14} \end{aligned} \right\}, \quad 1.13$$

рішення яких знаходять зі співвідношень:

$$b_j = \alpha_j / c_{jj} \quad (j = 0, 1, 2, \dots, 14), \quad 1.14$$

При цьому вдається позбутися невизначеності, пов'язаної з неоднозначною оцінкою коефіцієнтів регресії. Таким чином, для отримання незалежних одна від одної оцінок коефіцієнтів регресії потрібно проводити прогнозування так, щоб виконувалися умови лінійної незалежності і ортогональності векторів-стовпців матриці F незалежних змінних.

Використання функцій регресійного аналізу електронної таблиці EXCEL

Однією з поширених завдань в науці і техніці є апроксимація експериментальних даних аналітичними виразами.

Існують три способи використання Excel для апроксимації даних: використання вбудованих функцій регресії, перетворення нелінійних рівнянь у лінійні з наступним використанням вбудованих функцій регресії, використання надбудови Пошук рішення або програми Visual Basic для точного підбору параметрів нелінійного рівняння. Найчастіше дані можуть бути описані вбудованими функціями лінійної і поліноміальної регресії. Наявність лінійної залежності, в крайньому ступені, на обстежуваному інтервалі дозволяє успішно використовувати формулу лінійного наближення.

У разі нелінійних рівнянь, які допускають лінеаризацію за допомогою заміни змінних, функції лінійної регресії можна використовувати до перетворених рівнянь. Якщо ж рівняння не можуть бути лінеаризовані, можна скористатися надбудовою "Пошук рішення" або програмою на мові Visual Basic і підібрати коефіцієнти таким чином, щоб досягти мінімальної залишкової помилки (суми квадратів різниці між фактичними і прогнозованими значеннями) або максимального коефіцієнта кореляції.

Excel володіє вбудованими можливостями регресійного аналізу, який дозволяє апроксимувати дані як прямою лінією, так і складними поліномами. Велика частина завдань опису даних може бути вирішена засобами лінійної регресії.

При використанні функцій регресії для апроксимації даних відбувається мінімізація залишкової квадратичної помилки між фактичними і прогнозованими значеннями (метод найменших квадратів). Залишкова помилка обчислюється за наступною формулою:

$$E = \sum_{i=1}^N (y(x_i) - y_i)^2$$

де $y(x_i)$ – значення, яке прогнозується, n – число точок, x_i і y_i – фактичні значення.

В Excel використовується модель багатовимірної лінійної регресії. Це означає, що прогнозовані значення $y(x_i)$ мають вигляд:

$$y(x_{1,i}, x_{2,i}, \dots) = A + Bx_{1,i} + Cx_{2,i} + \dots$$

де A, B, C – шукані коефіцієнти, які визначаються підставкою функції $y(x_{1,i}; x_{2,i}; \dots)$ у вираз для залишкової помилки і приведенням до нуля часткових похідних по всіх коефіцієнтах. Це приводить до системи лінійних рівнянь для коефіцієнтів. Усі обчислення виконуються Excel автоматично.

Крім коефіцієнтів рівняння регресії, Excel також обчислює додаткову регресійну статистику:

- стандартна помилка для оцінки y ;
- коефіцієнт детермінованості (r^2);
- стандартні помилки для коефіцієнтів;
- F – статистика;
- число ступенів свободи;

- регресійна сума квадратів і остаточною сума квадратів.

Стандартна помилка для оцінки Y

Стандартна помилка для оцінки y є оцінкою помилки для одиночного значення y , обчисленого на основі рівняння регресії. Ця оцінка використовується разом з критерієм Стьюдента при визначенні довірчого інтервалу для обчислення кривої. Довірчий інтервал – це область навколо прогнозованої кривої, в якій з певною вірогідністю (y медицині не $< 95\%$) міститься істинна крива. Стандартна помилка для оцінки y обчислюється за наступною формулою:

$$S_{y,x} = \sqrt{\frac{\sum_{i=1}^n (y_i - y(x_i))^2}{p}}$$

де p – число ступенів свободи ($p = n - 2$ для прямої лінії).

Коефіцієнт детермінованості

Коефіцієнт детермінованості є квадратом коефіцієнта кореляції (r^2), який визначає, наскільки правильне рівняння, яке отримано за допомогою регресійного аналізу, та описує фактичні дані. Коефіцієнт детермінованості може приймати значення від 0 до 1, причому 1 відповідає повному збігу прогнозованих і фактичних даних. Правильним наближенням вважається таке, при якому значення коефіцієнта детермінованості більше 0,9. Коефіцієнт обчислюється за формулою:

$$r^2 = 1 - \frac{\sum_{i=1}^n (y_i - y(x_i))^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

де через \bar{y} позначено середнє значення

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}.$$

Стандартне значення помилок для коефіцієнтів

Стандартні значення помилок для коефіцієнтів є заходами точності будь-якого з коефіцієнтів регресії (A, B, \dots). Стандартне значення помилки першого коефіцієнта S_A обчислюється по допомогою стандартної помилки для оцінки y :

$$S_A = \sqrt{\frac{1}{n} + \frac{(x)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}, \text{ де } \bar{x} = \frac{\sum_{i=1}^n x_i}{n}.$$

Основне використання стандартних значень помилок для коефіцієнтів знаходять при перевірці статистичної значимості коефіцієнтів (статистичної рівності нулю). Оскільки всі коефіцієнти входять у рівняння регресії ліній-

но, рівність нулю будь-якого коефіцієнта означає відсутність кореляції між відповідним членом x і даними y . Для перевірки статистичної значущості коефіцієнта необхідно взяти значення t -розподілу Стьюдента для потрібного довірчого інтервалу ($1 - \alpha$) і числа ступенів свободи (p) і перевірити виконання нерівності:

$$|B| > \frac{t_{\alpha}}{2} S_B$$

Якщо нерівність виконується, то коефіцієнт є значущим і значення y залежать від значень x , пов'язаних з даним коефіцієнтом. Якщо нерівність не виконується, значення y не залежить від зазначених значень x і коефіцієнт вважається рівним нулю. Інші коефіцієнти перевіряються таким же способом. Таблиці значень t -розподілу функції (СТЬЮДРАСПОБР), яка повертає значення t -розподілу за заданим значенням α і певним числом ступенів свободи.

Якщо абсолютне значення коефіцієнта по порядку величини більше, ніж стандартне значення помилки, то коефіцієнт є значущою. У разі якнайменше чотирьох ступенів свободи (наприклад, при побудові прямої лінії за шістьма точками), значення t -розподілу Стьюдента для 95 % довірчого інтервалу становить лише 2,1 і зменшується зі збільшенням числа ступенів свободи. Тому можна вважати, що коефіцієнт є значущим, якщо його абсолютне значення перевищує стандартні помилки більше ніж у 2,5 рази. У протилежному випадку необхідно підставляти у вищезазначену формулу точне значення t -розподілу і перевіряти виконання нерівності.

F -статистика використовується разом із F -значеннями для визначення достовірності того, що дані дійсно описуються зазначеним виразом або збіг викликано випадковими флуктуаціями. Як і у випадку t -теста Стьюдента, для використання F -статистики необхідно використовувати значення F , узяті із таблиць або обчислені за допомогою функції (FРАСПОБР). Табличне або повернене функцією значення F , більше ніж табличне, є лише схожістю, яка зумовлена реальною кореляцією, а не випадковими флуктуаціями.

Для визначення значень F необхідні два числа ступенів свободи. Перше – n_f – яке дорівнює кількості коефіцієнтів у рівнянні регресії мінус один. Друге – p – стандартне число ступенів свободи, рівне різниці числа точок і числа коефіцієнтів у рівнянні регресії. Значення числа ступенів свободи p повертається функцією ЛИНЕЙН та використовується в t -тесті Стьюдента.

Число ступенів свободи

Число ступенів свободи p дорівнює різниці числа точок і числа коефіцієнтів у рівнянні регресії. Наприклад, рівняння прямої лінії має два коефіцієнти, які відповідають куту нахилу і зсуву по осі y . Якщо є десять точок, число ступенів свободи дорівнює 8 ($= 10 - 2$). Це число використовується в багатьох статистичних таблицях при визначенні довірчих інтервалів.

Регресійна сума квадратів і остаточної сума квадратів

Для оцінки помилки апроксимації використовуються дві суми квадратів. Регресійна сума дорівнює сумі квадратів різниці між значеннями y і середнім значенням y :

$$\sum_{i=1}^n (y_i - \bar{y})^2$$

Таким чином, ця величина є мірою розкиду даних щодо середнього значення.

Залишкова сума квадратів рівна сумі квадратів відмінностей між реальними і прогнозованими значеннями y :

$$\sum_{i=1}^n (y_i - y(x_i))^2$$

Таким чином, ця величина є мірою розкиду даних щодо лінії регресії. Розділивши зазначені величини на число ступенів свободи, отримуємо дисперсію даних щодо середнього значення і щодо лінії регресії. Обчисливши квадратний корінь дисперсії, отримуємо середньоквадратичне відхилення. Отримані значення показують, чи дійсно дані описуються рівнянням регресії, чи є постійними.

Обчислення лінійної регресії

Лінійний регресійний аналіз в Excel виконується або ж за допомогою вбудованої функції ЛИНЕЙН, ЛГРФПРИБЛ, ТЕНДЕНЦИЯ і РОСТ, або ж за допомогою інструменту Регресія пакету Аналіз даних. Лінійна регресія виконується обома способами практично однаково. Основна відмінність полягає в тому, що значення функцій автоматично оновлюються при кожній зміні даних, в той час як інструмент аналізу необхідно запускати вручну. Крім того, в пакет аналізу входить набагато більша кількість різноманітних статистик.

Обчислення регресії за допомогою функції

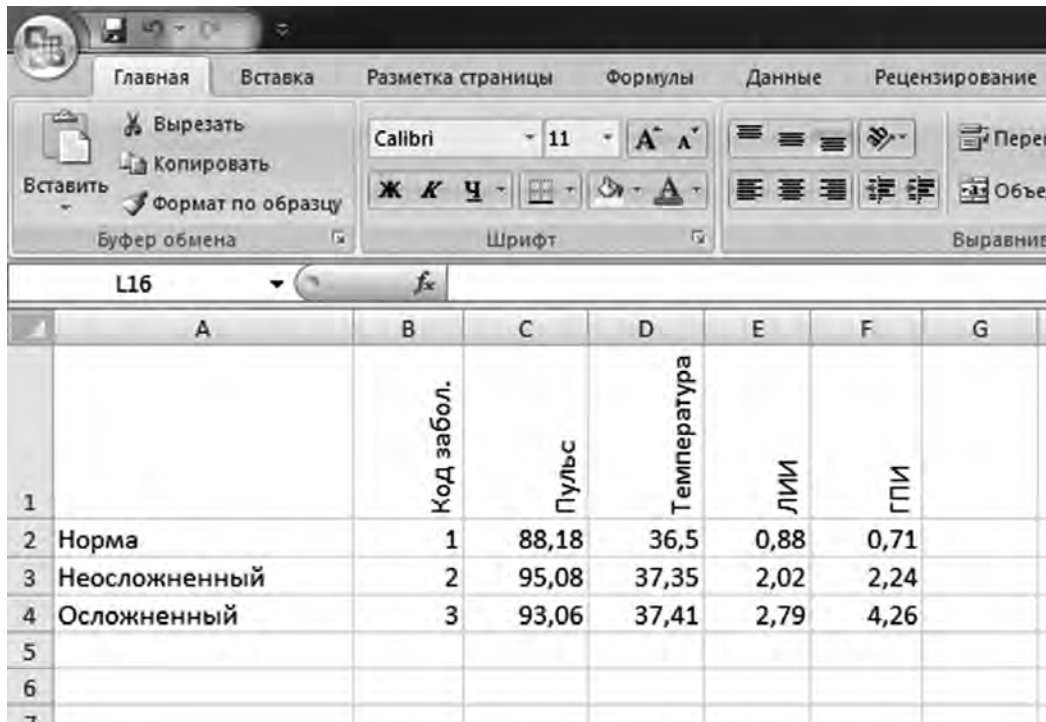
Функції ЛИНЕЙН, ЛГРФПРИБЛ, ТЕНДЕНЦИЯ і РОСТ обчислюють регресію даних робочого листа. Всі вони повертають масив, який містить регресивну криву, коефіцієнти або рівняння регресії. Функція ЛИНЕЙН виконує безпосередню лінійну регресію і повертає коефіцієнти, які відповідають тангенсу кута нахилу і зсуву по осі y . Функція ЛГРФПРИБЛ є варіацією лінійної регресії, апроксимуючи дані наступним рівнянням:

$$y = A(B^{x1})(C^{x2})...$$

і повертають коефіцієнти (A , B , C , ...). Функції ТЕНДЕНЦИЯ і РОСТ використовують ті ж формули для обчислення регресії, що і функції ЛИНЕЙН і ЛГРФПРИБЛ, але повертають безліч точок апроксимуючої кривої. Оскільки всі чотири функції повертають масиви, їх необхідно використовувати з діапазонами комірок або ж визначати окремі елементи масиву за допомогою функції ИНДЕКС.

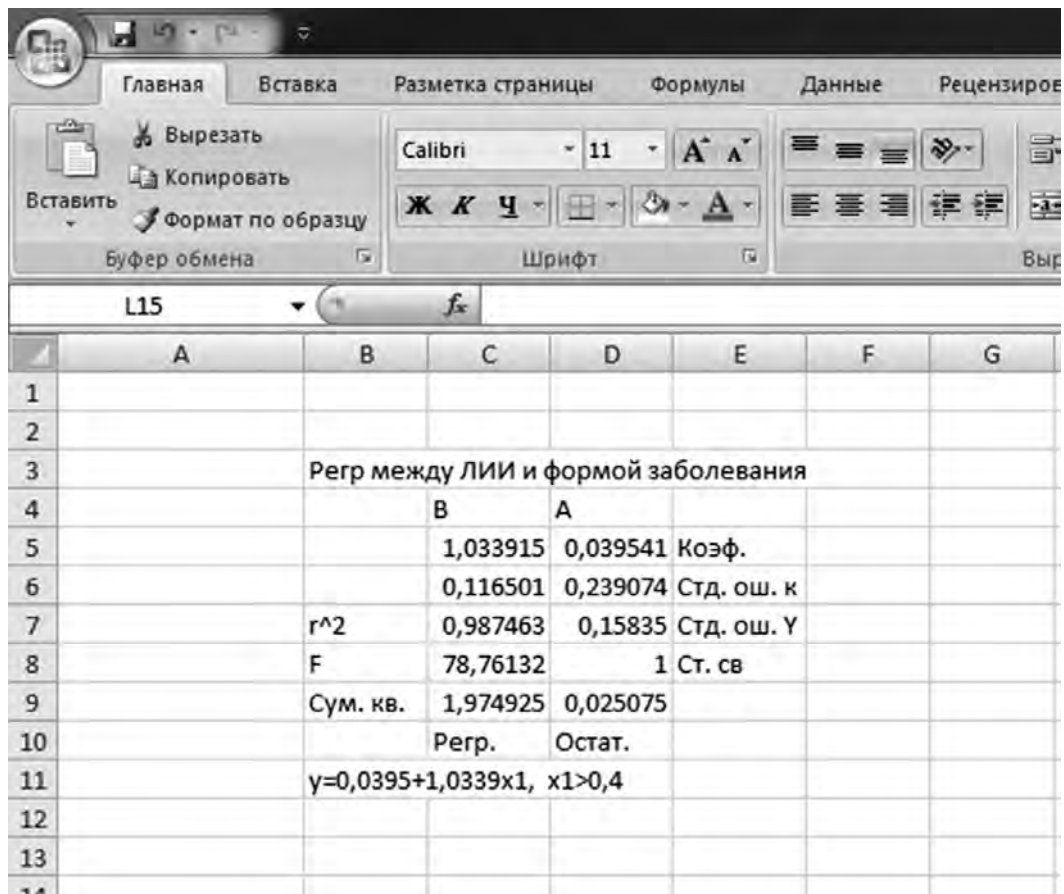
Функції ЛИНЕЙН і ЛГРФПРИБЛ мають наступний синтаксис:

ЛИНЕЙН (y -масив; x -масив; конст; статистика),
ЛГРФПРИБЛ (y -масив; x -масив; конст; статистика),



	A	B	C	D	E	F	G
1		Код забол.	Пультс	Температура	ЛИИ	ГПИ	
2	Норма	1	88,18	36,5	0,88	0,71	
3	Неосложненный	2	95,08	37,35	2,02	2,24	
4	Осложненный	3	93,06	37,41	2,79	4,26	
5							
6							
7							

Рис. 1. Середні значення параметрів пульсу, температури тіла, ЛІІ і ГПІ



	A	B	C	D	E	F	G
1							
2							
3		Регр между ЛИИ и формой заболевания					
4			B	A			
5			1,033915	0,039541	Коэф.		
6			0,116501	0,239074	Стд. ош. к		
7		r^2	0,987463	0,15835	Стд. ош. Y		
8		F	78,76132	1	Ст. св		
9		Сум. кв.	1,974925	0,025075			
10			Регр.	Остат.			
11			y=0,0395+1,0339x1, x1>0,4				
12							
13							
14							

Рис. 2. Зв'язок форми захворювання (y) з ЛІІ (x₁)

	A	B	C	D	E	F	G
1							
2							
3		Регр. между ГПИ и формой захворювання					
4			B	A			
5			0,560113	0,655576	Коеф.		
6			0,045081	0,126488	Стд. ош. к		
7		r^2	0,993564	0,113456	Стд. ош. Y		
8		F	154,3728	1	Ст. св		
9		Сум. кв.	1,987128	0,012872			
10			Регр.	Остат.			
11		y=0,6556+0,5601x2, x2>0,4					
12							

Рис. 3. Зв'язок форми захворювання (y) з ГПІ (x₂)

	A	B	C	D	E	F	G
1							
2							
3		Регр. между температурой и формой захворювання					
4			B	A			
5			1,73658983	-62,4069	Коеф.		
6			0,89153996	33,06769	Стд. ош. к		
7		r^2	0,791411836	0,645892	Стд. ош. Y		
8		F	3,794135872	1	Ст. св		
9		Сум. кв.	1,582823672	0,417176			
10			Регр.	Остат.			
11		y=-62,4069+1,7366x3, 36<x3<41					
12							
13							
14							

Рис. 4. Зв'язок форми захворювання (y) з температурою тіла (x₃)

	A	B	C	D	E	F	G
1		Код забол.	ЛИИ	ГПИ			
2	Норма	1	0,88	0,71			
3	Неосложненный	2	2,02	2,24			
4	Осложненный	3	2,79	4,26			
5							
6	Сумма	6	5,69	7,2			
7							
8							
9	Вычисление коэффициентов уравнения регрессии						
10		2,000	5,690	7,200		6,000	
11		5,690	12,634	17,011		13,287	
12		7,200	17,010	23,618		17,949	
13							
14		-1,160	1,536	-0,752		-0,062	
15		1,536	0,594	-0,896		1,024	
16		-0,752	-1,896	0,917		0,041	
17							
18		y = -0,062 + 1,024x1 + 0,04x2; x1 > 0,4; x2 > 0,4					
19							

Рис. 5. Система трьох рівнянь з трьома невідомими

де y -масив служить посиланням на масив даних y , x -масив служить посиланням на один або кілька масивів даних x , конст. – це логічне значення, яке визначає константу зсуву, і статистика – це логічне значення, яке вказує, чи потрібно повернути додаткову статистику по регресії.

Якщо член x -масив пропущений, то замість нього використовують безліч натуральних чисел {1, 2, 3, ...}. Якщо член конст. пропущений або ИСТИНА, константа зсуву (A) обчислюється звичайним способом. Якщо член конст. дорівнює ЛОЖЬ, константа зсуву вважається рівною 1 для функції ЛГРФПРИБЛ. Якщо член статистика дорівнює ИСТИНА, разом з коефіцієнтами рівняння регресії, повертається таблиця з вісьмома або більше статистичними значеннями.

Однофакторна математична модель

Клінічні дані про хворих з ГА було зведено в три таблиці, у кожній з яких зведені найбільш істотні стандартні критерії, аналіз даних яких виявив, що форма перебігу ГА залежить від чотирьох параметрів, а саме: пульс, температура тіла, ЛІІ і ГПІ на момент госпіталізації.

Середні значення даних цих параметрів були зведені в таблицю (рис. 1)

На нашу думку, величина зміни середніх значень пульсу для трьох груп не є фактором, що впливає на форму захворювання. На наш погляд, це пояснюється тим, що облік пульсу на момент госпіталізації відбувається без урахування загального стану дитини, тривалості захворювання і нормальних значень показника у віковому аспекті. Підрахунок пульсу відбувається так само без урахування компенсаторних можливостей серцево-судинної системи організму, яка ще не обтяжена хронічними захворюваннями.

За описаною вище методикою була розрахована однофакторна модель, що зв'язує форму захворювання (y) з ЛІІ (x_1) (рис. 2).

За описаною вище методикою була розрахована однофакторна модель, яка об'єднала форму захворювання (y) з ЛІІ (x_1). Подібним чином були розраховані однофакторні математичні моделі, які зв'язують форму захворювання (y) з ГПІ (x_2) (рис. 3) і температурою тіла (x_3) (рис. 4).

Аналіз однофакторних математичних моделей $y = a_1x_1$, $y = a_2x_2$ та $y = a_3x_3$ показав, що вони мають наступні коефіцієнти детермінуючої важливості: $a_1 = 0,987$; $a_2 = 0,994$; $a_3 = 0,791$. З цього стає зрозумілим, що математична модель $y = a_3x_3$ недостатньо добре змальовує фактичні дані.

Тому ми вважаємо, що температура тіла не є важливим фактором, який впливає на форму захворювання (в даному випадку на ускладнений перебіг ГА), а є тільки окремим показником загального стану організму дитини на момент госпіталізації її в стаціонар, без урахування всієї сукупності супутніх критеріїв (вік, тривалість захворювання, супутня патологія і т.д.).

Таким чином, форму ускладненого перебігу гострого апендициту описує наступна система рівнянь:

$$\begin{cases} y = 0.0395 + 1.0339x_1, & x_1 > 0.4 \\ y = 0.6556 + 0.5601x_2, & x_2 > 0.4 \end{cases}$$

Звідси випливає, що вплив показника ЛПІ приблизно в 2 рази більший, ніж у ГПІ. Тому ми вважаємо за можливе зробити висновок про те, що обидва ці показники, ЛПІ та ГПІ, для діагностики імовірнісного ускладненого перебігу ГА у дітей закономірні вже на момент госпіталізації дітей в стаціонар.

Система рівнянь, наведена вище, не впливає на взаємодію факторів x_1 і x_2 на форму перебігу захворювання.

Двофакторна математична модель

Для створення двофакторної математичної моделі скористалися множинною кореляцією:

$$x = f(x_1, x_2)$$

Підготовлені дані для побудови системи трьох рівнянь із трьома невідомими представлені на рис. 5.

Отримана двофакторна математична модель:

$$y = -0,062 + 1,024x_1 + 0,041x_2; \quad x_1 > 0,4; \quad x_2 > 0,4,$$

дозволяє при госпіталізації визначати ймовірність ускладненого перебігу гострого апендициту.

Наприклад, якщо при надходженні у хворого ЛПІ дорівнює 2,2, а ГПІ 2,5, то показник форми захворювання $y = 2,29$, що перебуває в цифровому обчисленні в межах неускладненого перебігу захворювання і повністю збігається з даними аналізу трьох груп хворих (рис. 5).

Наприклад, для показника при надходженні ЛПІ дорівнює 2,5, а ГПІ відповідно 4,1. Показник форми захворювання $y = 2,67$, що в цифровому обчисленні потрапляє в зону ускладненого перебігу гострого апендициту і повністю збігається з даними аналізу трьох груп хворих (рис. 5).

Висновки

1. До чинників, що визначають ускладнений перебіг гострого апендициту у дітей, згідно з розробленою двофакторною моделлю, відносяться величини ЛПІ і ГПІ, які перевершують значення певних визначуваних нами інтервалів на момент госпіталізації.
2. Значення аналізованих індексів визначалися ступенем вираженості патологічного процесу, тривалістю захворювання, вірулентністю мікрофлори і віком пацієнтів на тлі сформованого синдрому ендогенної інтоксикації.

Література

1. Chemich ND. Ispolzovanie innovatsionnykh kompiuterizirovannykh sistem v praktike infektsionista (The use of innovative computerized systems in the practice of an infectious disease specialist). Klinicheskaia infektologiya i parazitologiya. 2014;8(1):11–8. (In Russian)
2. Matveev SB. Sistemnaia otsenka endogennoi intoksikatsii pri razlichnykh neotlozhnykh sostoianiiakh (Systemic assessment of endogenous intoxication in various emergency conditions). Neotlozhnaia meditsinskaia pomoshch. 2012;(4):17–21. (In Russian)
3. Uspenskaya ID, Shabunina EI, Zhukov EA. Endogennaia intoksikatsiia i biokhimicheskaia protektsiia u detei pri tseliakii (Endogenous intoxication and biochemical protection in children with celiac disease). Vestnik RAMN. 2014;(7-8):93–9. (In Russian)
4. Shmoilov DK. Patogeneticheskaia rol endogennoi intoksikatsii (Pathogenetic role of endogenous intoxication). Laboratorna diahnozyka. 2012;60(2):65–9. (In Russian)
5. Gasanov MD. Algoritm otsenki stepeni tiazhesti endotoksikoza pri peritonitakh (Algorithm for assessing the severity of endotoxemia with peritonitis). Vestnik Rossiiskoi voenno-meditsinskoi akademii. 2014;48(4):75–8. (In Russian)
6. Sadchikov DV, Suburbs MV, Vartanyan TS. Perioperatsionnye oslozhneniia u patsientov vysokogo anesteziologo-operatsionnogo riska (Perioperative complications in patients with high anesthetic operative risk). Saratovskii nauchno-meditsinskii zhurnal. 2012;8(1):51–7. (In Russian)
7. Fayzieva DB, Shukhrat Buritoshevich S, Holidbek Mahammadjanovich K. Integrated leukocyte index in diagnosis of endogenous uveitis. European Journal of Research [Internet]. Tadqiqot; 2017 May 5;5(5):71–82. Available from: <https://doi.org/10.26739/2521-3253-2017-5-5-7>
8. Kalf-Kalif YY. O leukotsitarnom indekse intoksikatsii i ego prakticheskome znachenii (About leukocyte index of intoxication and its practical significance). Vrachebnoe delo. 1941;(1):31–5. (In Russian)
9. Vasiliev VS, Komar VI, Sheiko MI. Otsenka tiazhesti intoksikatsii pri ostrykh infektsionnykh bolezniakh (Assessment of the severity of intoxication in acute infectious diseases). Zdravookhranenie Belorussii. 1984;(3):46–50. (In Russian)

Математическое моделирование прогнозирования осложненного течения острого аппендицита у детей

Коноплицкий В. С., Мотыгин В. В., Якименко А. Г., Дмитриев Д. В., Михальчук Т. И.

Винницкий национальный медицинский университет им. Н. И. Пирогова

Резюме. Эндогенная интоксикация, сопровождающая осложненное течение острого аппендицита у детей, является одним из тяжелых состояний, которые служат основной причиной развития осложнений в послеоперационном периоде. Разносторонние клинические наблюдения клиницистов свидетельствуют, что, несмотря на использование современных технологий в операционной технике, анестезиологическом обеспечении, количество осложнений в послеоперационном периоде не имеет тенденции к снижению, а основной причиной летальных исходов при аппендикулярном перитоните служит молниеносное развитие тяжелой степени эндотоксикоза, который развивается вследствие резорбции токсических веществ, образовавшихся из многих компонентов, в том числе из патогенной микрофлоры и разрушенного метаболизма.

Цель. Создание математической модели прогнозирования осложненного течения острого аппендицита у детей с использованием функций регрессионного анализа электронной таблицы EXCEL путем аппроксимации экспериментальных данных.

Материалы и методы. В работе использована информация из 59 МКСБ, находившихся на лечении в клинике детской хирургии ВНМУ им. Н. И. Пирогова по поводу острого аппендицита. Данные о больных были сведены в две группы: больные с неосложненным течением патологии – 24 пациента и больные с осложненным течением острого аппендицита – 35 детей. Группой контроля служили данные обследования 37 здоровых детей. В каждой из трех групп изучались следующие показатели: пол, возраст, койко-день, длительность заболевания, наличие симптомов раздражения брюшины, характер экссудата в брюшной полости, количество эритроцитов, гемоглобин, лейкоцитоз, количественная характеристика формулы белой крови, СОЭ, уровни клеточных критериев эндогенной интоксикации: лейкоцитарного индекса интоксикации и гематологического показателя интоксикации. Все данные определялись непосредственно на момент госпитализации больного в стационар.

Результаты. Обсуждение. На основе множественной корреляции $x = f(x_1, x_2)$ создана двухфакторная математическая модель вероятности осложненного течения острого аппендицита у детей. Согласно полученным результатам исследования системы трёх уравнений с тремя неизвестными установлено, что если показатель формы заболевания составляет $y = 2,67$ и выше, при котором в цифровом исчислении результат попадает в зону осложненного течения острого аппендицита и полностью совпадает с данными анализа трёх групп больных.

Выводы. К составляющим, которые определяют неосложненное течение острого аппендицита у детей, согласно разработанной двухфакторной модели, относятся показатели ЛИИ и ГПИ, которые превышают значения определенных изучаемых интервалов на момент госпитализации. Значения анализируемых индексов определялись степенью выраженности патологического процесса, длительностью заболевания, вирулентностью микрофлоры и возрастом пациентов на фоне сформировавшегося синдрома эндогенной интоксикации.

Ключевые слова: дети, острый аппендицит, осложненное течение, модель.

Mathematical modeling for prediction of complicated acute appendicitis in children

Konoplitsky V. S., Motygin V. V., Yakimenko A. G., Dmytriiev D. V., Mikhachuk T. I.

National Pirogov Memorial Medical University, Vinnytsia

Abstract. Endogenous intoxication, which accompanies the complicated course of acute appendicitis in children, is one of the serious conditions, which are the main cause of complications in the postoperative period. Diverse clinical observations done by clinicians suggest that despite the use of modern technologies in operating equipment, anesthetic support, the number of complications in the postoperative period has no tendency to decrease, and the main cause of lethal outcomes with appendicular peritonitis is lightning development of severe degree of endotoxiosis, which develops as a result of resorbing toxic substances formed of many components, including pathogenic microflora and destruction metabolism.

Purpose. Creation of a mathematical model for prediction of complicated acute appendicitis in children using the functions of regression analysis of EXCEL spreadsheet by approximating experimental data.

Materials and methods. In the work, there is information about 59 patients with acute appendicitis treated at the clinic of pediatric surgery at the National Pirogov Memorial Medical University. The data on patients were divided into two groups: patients with not complicated course of pathology – 24 patients, and patients with complicated course of acute appendicitis – 35 children. The control group served the data of 37 healthy children. In each of the three groups, the following parameters were studied: gender, age, bed day, duration of the disease, presence of symptoms of peritoneal irritation, the nature of exudate in the abdominal cavity, the number of red blood cells, hemoglobin, leukocytosis, quantitative characteristics of white blood formula, ESR, cellogenic endogenous intoxication: leukocyte index of intoxication and hematological index of intoxication. All the data were determined directly at the time of hospitalization of a patient to the hospital.

Results. Discussion. Based on multiple correlation $x = f(x_1, x_2)$ a two-factor mathematical model of probability of the complicated course of acute appendicitis in children was created. According to the obtained results of the study of the system of three equations with three unknowns, it was established that if the indicator of the form of the disease is $y = 2.67$ and above, when, in numerical terms, the result falls into the zone of complicated flow of acute appendicitis and completely coincides with data analysis of the three groups of patients.

Conclusions. The components that determine the uncomplicated course of acute appendicitis in children, according to the developed two-factor model, include LII and GPI, which exceed the values intoxication of certain intervals studied at the time of hospitalization. The analyzed indices were determined by the degree of severity of the pathological process, duration of disease, virulence of microflora and age of patients against the background of the formed endogenous intoxication syndrome.

Key words: children, acute appendicitis, complicated flow, model.