

**РЕЛЯЦІЙНА БАЗА ДАНИХ ІЗ ФАРМАЦЕВТИЧНОЮ ІНФОРМАЦІЄЮ:
ПРОБЛЕМИ СТВОРЕННЯ ТА ПЕРВИННОГО НАПОВНЕННЯ ДЛЯ
ЗАБЕЗПЕЧЕННЯ ЯКІСНОЇ СТАТИСТИЧНОЇ ОБРОБКИ****Ключові слова:** фармацевтична інформація, лікарська форма, доза лікарського засобу, структура реляційної бази данихI. Yu. RERVYATSKYY (<https://orcid.org/0000-0002-4726-3639>)*Danylo Halytsky Lviv National Medical University***A RELATIONAL DATABASE WITH PHARMACEUTICAL INFORMATION:
PROBLEMS OF CREATION AND PRIMARY FILLING FOR PROVISION OF A
QUALITATIVE STATISTICAL PROCESSING****Key words:** pharmacy information, drug form, drug dose, a structure of a relational database

Інтернет став основним джерелом інформації в сучасному суспільстві та займає ключову роль у його розвитку. Деякі положення законодавства (у відповідності щодо забезпечення доступності населення інформацією) зобов'язують установи, організації, компанії тощо здійснювати оприлюднення інформації з допомогою інтернет ресурсів [1, 2]. Також оприлюднення інформації здійснюється за власної ініціативи, в тому числі з використанням рерайтингу та копірайтингу [3]. Як форма інтернет-бізнесу – інтернет-сервіси, використовуючи різні методи, дозволяють пошуковим системам ідентифікувати незначну частину наявної інформації, проте для інтернет-користувачів доступ до повної інформації на них платний.

Інформація в інтернеті дублюється та нагромаджується, що призводить до зниження якості одержаних користувачами результатів від пошукових систем.

Це можна розглядати як пониження рівня доступності та швидкості доступу до якісної (корисної) для користувача інформації. Це призводить до розробки та удосконалень методологій оцінки якості мережевої інформації [4].

Загальнодоступна фармацевтична інформація структурується під певні стандарти з метою задовольнити певні потреби (які визначаються оприлюднювачем) та може бути неповною [2]. Оскільки оприлюднювачі не ставлять за мету забезпечити можливість незалежним спеціалістам здійснення подальшої автоматизованої обробки поданої інформації, то й обирають методи її представлення, які спрямовані на оптимальне візуальне сприйняття користувачами.

Дослідження з опрацювання фармацевтичних баз даних в Україні проводилися Б. Л. Парновським (1981), Л. П. Смирноюю (1986), Л. С. Шелепетень (1988), О. М. Заліською (2004), А. І. Бойком (2009), за кордоном – Л. Н. Інаповим (2015), Г. Г. Данагулян (2016) та ін. Вищевказані дослідження було присвячено опрацюванню загальної методології фармацевтичних баз даних, створенню баз даних про конкретні лікарські засоби, однак проблематика первинного наповнення баз даних фармацевтичною інформацією із забезпеченням її подальшої статистичної обробки в Україні не розглядалася.

Матеріали та методи дослідження

Об'єктами дослідження стали (чинні станом на лютий 2019 р.): міжнародний анатомо-терапевтично-хімічний (АТХ) класифікатор, міжнародний класифікатор хвороб 10-го видання (МКХ-10); зміст Європейської фармакопеї 9.8-го видання; класифікатори лікарських форм (ЛФ) за FDA (Food and Drug Administration, USA), за ЕМА (European Medicines Agency), за European Pharmaceutical Market Research

Association (EphMRA), за МОЗ України. У процесі роботи використано такі методи: пошук фармацевтичної інформації в інтернеті з допомогою пошукової системи Google із використанням можливостей розширеного пошуку; візуальне порівняння знайденої інформації на відповідність її структури до можливості завантаження у реляційну базу даних (РБД); обробка інформації для її стандартизації з допомогою формування комп'ютерного коду мовою VBA.

Результати дослідження та обговорення

Першим етапом забезпечення автоматизованого аналізу інформації є формування структури бази даних та її наповнення необхідною інформацією: збір інформації та її попередня обробка. Повнота і структурованість інформації – ключові фактори, які впливають на якість одержаної результуючої інформації. Тому перед формуванням РБД необхідно визначити, які поставлені цілі вона має задовольнити. Це впливає на вибір інформації, яка буде використовуватися для її первинного наповнення. Процес створення і наповнення бази даних – завдання, що вимагає істотних часових і трудових витрат. Такі роботи носять одиничний характер, закриті, прив'язані до певних локальних завдань, що призводить до обмеженості методичних розробок в цьому напрямі [5].

Для електронної онлайн-системи (ЕОС) «Ліквіпедія» (<http://www.likypedia.zzz.com.ua/>) було поставлено такі базові цілі:

- аналіз лікарських засобів (ЛЗ) за лікарськими формами (ЛФ), способами введення, а також за їхнім якісним та кількісним складом;
- аналіз ЛЗ за АТХ-класифікацією із передбаченням можливості використання інших регіональних та міжнародних класифікаторів;
- аналіз ЛЗ за їх регіональним виробництвом та споживанням;
- аналіз нормативних документів різних країн, які включають переліки торгових назв (ТН) ЛЗ, міжнародних назв ЛЗ чи їхніх інгредієнтів;
- аналіз супроводжуючої фармацевтичної інформації при відпуску ЛЗ;
- забезпечення уніфікованості та загальнодоступності наявної інформації;
- по можливості, адаптація інформації для її сприймання непрофільними (медичними, фармацевтичними) спеціалістами.

Після аналізу поставлених цілей (з метою їх досягнення) було сформовано перелік необхідної інформації для первинного завантаження до ЕОС «Ліквіпедія»:

- класифікатори АТХ та МКХ;
- перелік інгредієнтів ЛЗ;
- уніфіковане дозування та уніфікований (стандартизований певним чином) класифікатор ЛФ;
- перелік назв країн із їх регіональним розподіленням;
- перелік компаній – виробників ЛЗ та власників брендів торгових марок ЛЗ.

Існують деякі особливості структуризації РБД, що ускладнюють процеси підготовки інформації та її автоматизованого завантаження. Тому ця інформація теж повинна мати певну структуру, а при її полікомпонентності структура має також забезпечувати комп'ютерному коду доступ до кожного окремого компонента інформації. Полікомпонентність інформації має визначатися на основі стандартів полів таблиць РБД. Оптимальними файлами, які містять структуровану інформацію є, зокрема, текстовий файл – *.txt (може містити інформацію у стандарті XML, послідовний рядковий запис, табличну структуру із відповідним розділовим символом) чи файли електронних таблиць (МО Excel, OpenOffice Calc тощо). Наприклад, запис «лідокаїн розчин для ін'єкцій 20 мг/мл; 2 мл № 10, ПАТ Лекхім-Харків, Україна» є полікомпонентною інформацією, яка, відповідно до стандартів РБД, складається з наступних компонентів: торгової назви «лідокаїн»; лікарської форми «розчин для ін'єкцій»; ве-

личини дози «20»; одиниці виміру дози «мг/мл»; кількості ЛЗ в одиниці продукції «2»; одиниці виміру кількості ЛЗ в одиниці продукції «мл»; кількості одиниць продукції в упаковці (номером) «10»; назви виробника «ПАТ Лекхім-Харків»; назви країни розташування виробника «Україна».

Оскільки поставлені цілі передбачають функціонування ЕОС «Лікепедія» з опрацюванням інформації із різних країн, то для поєднання та уніфікації такої інформації було прийнято рішення вносити значну її частину англійською мовою. До винятків можна віднести торгову назву ЛЗ, яка буде вноситися зазначеною на упаковці мовою.

Інформацію, одержану з допомогою пошукової системи Google із різних онлайн-джерел було скопійовано (із файлів формату: *.html, *.exl, *.pdf) до текстового файлу (*.txt), структуровано, фільтровано (відібрано або вилучено) та завантажено до РБД.

Основне призначення АТХ-класифікатора – представлення статистичних даних про споживання ЛЗ. При пошуку в інтернеті АТХ-класифікатор не був знайдений у вигляді єдиного повного розкритого переліку і був одержаний із вікіпедії шляхом поетапного копіювання переліку кожної окремої групи (https://en.wikipedia.org/wiki/Anatomical_Therapeutic_Chemical_Classification_System). При фільтруванні одержаної інформації із одержаного об'єднаного переліку були вилучені коди, препарати яких використовуються у ветеринарії (із позначенням «Q»). Загальна кількість відібраних записів – 6 147.

Відповідно до стандартів структуризації інформації РБД будь-яка класифікаційна система має бути в окремій таблиці. Взаємозв'язок (класифікація) об'єктів здійснюється через використання ідентифікаторів записів із таблиці класифікатора. Проте в Україні більшість інформаційних систем використовують код АТХ як текстову величину (ручне введення класифікаційної інформації) для кожного окремого ЛЗ, що призводить до утруднення одержання коректної інформації за критеріями відбору («Передумови стандартизації електронної номенклатури лікарських засобів в Україні», Запорізький медичний журнал, 2013, № 1).

З метою одержання в майбутньому статистичної інформації МКХ-10 був одержаний із сайту «Centers for Disease Control and Prevention» (<https://www.cdc.gov/nchs/icd/icd10cm.htm>), який забезпечує якісну різнопланово структуровану подачу даної інформації. Перелік МКХ-10 був легко опрацьований та завантажений до РБД. На даний момент триває розробка 11-ї версії (<https://icd.who.int/>), тому в ОЕС «Лікепедія» МКХ-10 буде використана як основа для побудови взаємозв'язків у РБД і, згодом, замінена на МКХ-11.

Одержати перелік інгредієнтів (субстанції, активні речовини, допоміжні речовини, лікарська рослинна сировина), з яких виготовляються ЛЗ, при пошуку в інтернеті – не вдалося. В його результатах здебільшого відображалися законодавчі переліки рекомендованих (для певної конкретної мети) або заборонених до використання речовин.

Базовий перелік інгредієнтів був отриманий із англomовного змісту європейської фармакопеї (ЄФ) 9.8-го видання (<https://www.edqm.eu/en/european-pharmacopoeia-ph-eur-9th-edition>). Інформація із змісту (сторінки 7040–7072) була скопійована у текстовий файл. При її опрацюванні було враховано особливість змісту, який включав не лише назви статей на інгредієнти, але й назви інших статей. Алфавітний зміст також включає інформацію із доповнень (9.1–9.8). Тому, при обробці інформації візуально було визначено та враховано перші та останні сторінки, які містили інформацію про інгредієнти. Межі відбору інформації наведено у табл. 1.

**Інтервали сторінок в Європейській фармакопеї 9.8,
у межах яких було відібрано інформацію про назви інгредієнтів**

ЄФ	Початок інтервалу		Завершення інтервалу	
	інгредієнт	стор.	інгредієнт	стор.
9.0	Abacavir sulfate	1619	Zuclopenthixol decanoate	3978
9.1	Alfuzosin hydrochloride	4123	Water for injections	4215
9.2	Acanthopanax bark	4447	Xanthan gum	4627
9.3	Aceclofenac	4835	Vitamin A concentrate (solubilisate/emulsion), synthetic	5038
9.4	Arsenic	5103	Water, purified	5475
9.5	Acitretin	5627	Zolmitriptan	5759
9.6	Akebia stem	5985	Zoledronic acid monohydrate	6225
9.7	Acetone	6561	Terpin monohydrate	6688
9.8	Acacia	6841	Xylazine hydrochloride for veterinary use	7033

Після вилучення записів із зазначенням «for veterinary use» та сортування алфавітного переліку за сторінками, було залишено лише одну із продубльованих назв інгредієнта (наприклад: Ethyl parahydroxybenzoate sodium; Parahydroxybenzoate, sodium ethyl; Sodium ethyl parahydroxybenzoate). Загальна кількість відібраних записів – 1 766.

Перелік лікарської рослинної сировини, як інгредієнтів, довелось шукати та завантажувати окремо. Для одержання в майбутньому оптимального статистичного аналізу інформація би мала складатися з назви рослини та назви частини рослини або її лікарська форма, що входить в склад ЛЗ. Запит пошукової системи Google не привів до бажаного результату. Інформація, яка була завантажена до ЕОС «Лікепедія», була одержана з сайту European Medicines Agency (science medicines health) в файлі формату *.xls (<https://www.ema.europa.eu/en/medicines/download-medicine-data#herbal-medicines-section>). Загальна кількість відібраних записів – 218.

Для оптимізації одержання статистичних вибірок в подальшому інгредієнти будуть класифіковані за походженням одержання: рослинні; біологічні; синтетичні [6] із передбаченням можливості якісного та кількісного редагування класифікаційних груп.

Пошук переліку ЛФ був досить успішним. Проте, було одержано кілька переліків відповідно до різних класифікацій ЛФ. Критерій ЛФ є одним з базових факторів, які впливають на ефективність лікування, що підтверджується появою нових ЛФ чи їх характеристик [6]. Тому, з врахуванням всіх цих критеріїв, перелік ЛФ для ЕОС «Лікепедія» формувався на основі власного практичного досвіду, тенденцій зазначення ЛФ на упаковках ЛЗ виробниками, електронних базах даних фармацевтичних дистриб'юторів та на основі назв статей, наведених у ЄФ 9.8 видання. За основу, при формуванні адаптованого для використання в РБД переліку, було взято чотири англійські переліки ЛФ, затверджені: МОЗ України (чинні у лютому 2019 року) (<https://zakon.rada.gov.ua/go/v0500282-06>); FDA (<https://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/ucm162038.htm>), ЕМА (https://www.ema.europa.eu/documents/other/list-pharmaceutical-dosage-forms_en.xls), EphMRA (https://www.ephmra.org/media/1473/nfc_version_2018_guidelines.pdf).

В процесі формування переліку інформація із різних джерел про ЛФ була проаналізована та розділена на п'ять інформаційних блоків:

№ 1. Виготовлена виробником ЛФ.

№ 2. Характеристики виготовленої виробником ЛФ.

№ 3. ЛФ для застосування, яка буде приготовлена із виготовленої виробником ЛФ.

№ 4. Характеристика приготовленої для застосування ЛФ.

№ 5. Спосіб застосування ЛФ.

Приклад результату такого поділу інформації наведено у табл. 2. Для формування прикладу нами були обрані наступні препарати (відповідно до державного реєстру ЛЗ): № 1 оспексин (гранули оральної суспензії); № 2 пульмікорт (суспензія для розпилення); № 3 пульмікорт турбухалер (порошок для інгаляцій); № 4 омепразол-тева (капсули гастрорезистентні тверді); № 5 дексаметазон ВФЗ (краплі очні, суспензія); № 6 отипакс (краплі вушні); № 7 проспан форте (таблетки шипучі); № 8 респолепт конста (порошок для суспензії для ін'єкцій пролонгованої дії); № 9 ліпін-біолік (ліофілізат для емульсії).

Т а б л и ц я 2

Приклад розподілення інформації про лікарські форми за інформаційними блоками

№ ЛЗ	№ 1	№ 2	№ 3	№ 4	№ 5
1	granules		suspension		oral
2	suspension				sprayin
3	powder				inhalation
4	capsule	gastro-resistant, solid			oral
5	emulsion	drops			eye
6	solution	drops			ear
7	tablet	effervescent	solution		oral
8	powder		suspension	extended release	injection
9	powder	lyophilized	emulsion		intravenous, inhalation, internally, rectally

Поєднання інформації про ЛФ як єдиного запису реалізовано наступним алгоритмом у зазначеній послідовності:

1. Інформація в ІБ № 1 завжди наявна, подається першою.
2. Інформація з ІБ № 2, якщо вона наявна, подається в дужках.
3. При наявності інформації в ІБ № 3, перед нею доставляється «for preparation of».
4. Інформація з ІБ № 4, якщо вона наявна, подається в дужках.
5. Інформація в ІБ № 5 завжди наявна, подається між словами «for» та «use».

Наприклад, деякі записи із таблиці № 2 матимуть наступний вигляд:

№ 1: granules for preparation of suspension for oral use;

№ 5: emulsion (drops) for eye use;

№ 6: solution (drops) for ear use;

№ 7: tablet (effervescent) for preparation of solution for oral use;

№ 9: powder for preparation of suspension (extended release) for injection use.

Сформована за наведеним алгоритмом інформація зберігається в окремому полі (текстового формату) таблиці. В подальшому заплановано реалізацію такого самого алгоритму для представлення записів із використанням скорочених позначень.

Такий підхід до структуризації даної інформації дозволяє з легкістю відсортувати чи зробити статистичний аналіз за будь-яким критерієм чи комбінацією критеріїв з різних ІБ.

Загальна кількість сформованих записів ЛФ – 214.

У більшості інформаційних джерел доза вказується у поєднанні з торговою назвою або у окремій колонці разом з одиницею виміру дозування. Недоліком такої системи є подання доз у різних одиницях вимірювання (наприклад, одна доза може бути подана як 500 мг або як 0,5 г). Такий підхід до зазначення доз призводить до складнощів з автоматизації процесів одержання статистичних вибірок чи здійснення математичних розрахунків. Для усунення цього недоліку в ОЕС «Лікипедія» був реалізований наступний підхід: дози зберігаються в окремій таблиці; величина дози та одиниці виміру дози записуються у різні поля; кожна доза (на даний момент)

записується трьома варіантами – в грамах (г, г/мл), в міліграмах (мг, мг/мл, мг/г), в зручному для сприйняття користувачем варіанті (г, мг, мкг, ОД, ...).

У разі, коли дози вимірюються іншими одиницями (доз/мл, ммоль/мл, різні типи одиниць дій, крапель та ін.) – ці одинці виміру вказуються у всіх трьох варіантах запису. Приклад запису доз наведено у табл. 3. Для позначення мікрограму в дозах оптимально використовувати скорочення «mcg», оскільки скорочення «µg» утруднено вводити з клавіатури.

Т а б л и ц я 3

Приклад збереження інформації про дози лікарських засобів у електронній онлайн-системі «Лікепедія»

Ідентифікатор запису	Величина дози в грамах	Одиниця виміру дозування	Величина дози в міліграмах	Одиниця виміру дозування	Величина дози в оптимальному типі дозування	Одиниця виміру дозування
12	0,5	g	500	mg	500	mg
26	0,0000275	g/dose	0,0275	mg/dose	27,5	mcg/dose
17	1	g	1000	mg	1	g
32	0,009	g/ml	9	mg/ml	9	mg/ml
47	50	IU	50	IU	50	IU

Доза пов'язується із інгредієнтом з допомогою ідентифікаторів їх записів. Ця структура також передбачає можливе розширення кількості варіантів представлення доз. Такий підхід дає змогу:

- відображати дозу у зручнішому для сприйняття користувачем варіанті;
- при необхідності здійснення автоматизованих математичних розрахунків та статистичних вибірок – використовувати єдиний варіант виміру доз.

Отримання переліків первинних та вторинних упаковок не склало труднощів через їх малий об'єм інформації. Вони були сформовані в ручному режимі на основі презентації Primary And Second Packaging (<https://www.slideshare.net/vamsikrishnareddy57/primary-and-second-packaging>).

Для відображення інформації про виробників ЛЗ, локалізацію препаратів під торговими назвами тощо, а також для можливості одержання статистичних даних за регіональним розподілом БД має містити перелік назв країн. Такий перелік був одержаний у придатному для автоматичного завантаження форматі (<https://www.countries-ofthe-world.com/all-countries.html>).

Перелік компаній (General index on Marketing Authorisation Holders and Sponsors) був одержаний з сайту European commission розділу Public information (http://ec.europa.eu/health/documents/community-register/html/mh_index.htm). Даний перелік містить прив'язку «назва компанії» – «країна». Це дало змогу завантажити дану інформацію зі вже готовим зв'язком.

Інформація в ЕОС «Лікепедія» надається користувачам у повному вигляді, адже вони самі можуть бути авторами наявної інформації. З цією метою був розроблений спеціальний фільтр, який дає змогу користувачам самостійно створювати умови відбору та сортування інформації. Його функціональність побудована на напіваавтоматизованому конструкторі частин SQL-запиту: структур «WHERE» і «ORDER». Такий підхід дав змогу користувачам самостійно створити будь-які власні фільтри для наявної інформації. Вивід одержаної інформації здійснюється у вигляді окремої таблиці, яка структурована відповідно до стандартів РБД. Це дає змогу реалізувати її коректне переміщення у табличний редактор, в якому користувачі з легкістю зможуть редагувати, видаляти, поєднувати, сортувати та структурувати інформацію під власні потреби.

Висновки

1. Встановлено, що представлена в мережі інтернет структура фармацевтичної інформації є адаптована до оптимального візуального сприйняття та не відповідає стандартам використання для первинного наповнення інформаційної бази даних, що унеможливує її подальшу автоматизовану статистичну обробку.

2. Опрацьовано і представлено варіанти структуризації інформації про лікарські форми зі врахуванням та уніфікацією інформації з різних класифікаторів відповідно до стандартів РБД, що дають змогу отримувати якісну різнопланову статистичну інформацію.

3. Опрацьовано і представлено варіант структуризації запису інформації про дозу лікарського засобу в РБД кількома варіантами, що забезпечує можливість подальшого здійснення автоматизованих розрахунків та оптимізацію відбору і сортування даних.

Список використаної літератури

1. Янішевська К. Д., Баленко Є. О. Адміністративно-правове регулювання доступу до публічної інформації // Правові горизонти. – 2018. – Вип. 13 (26). – С. 86–91. <https://doi.org/10.21272/legalhorizons.2018.i13.p86>

2. Павлов М. М. Особливості реалізації концепту відкритих даних у системі публічного управління: вітчизняний та зарубіжний досвід // Державне управління: теорія та практика. – 2018. – № 1. – С. 34–45.

3. Блинова Н. М., Кирилова О. В. Копірайтинг як діяльність зі створення онлайн-контенту // Communications and Communicative Technologies (CCT). – 2018. – Вип. 18. <https://doi.org/10.15421/291802>

4. Зубець В. В., Ільїна І. В. О качестве сетевой информации // Вестн. Тамбовского университета. – 2016. – Т. 21, Вип. 12 (164). – С. 168–172. [https://doi.org/10.20310/1810-0201-2016-21-12\(164\)-168-172](https://doi.org/10.20310/1810-0201-2016-21-12(164)-168-172)

5. Казимиров И. А. Проблемы создания и первичного наполнения базы данных сделок и предложений на рынке недвижимости // Вестн. ИрГТУ. – 2014. – № 12. – С. 283–286.

6. Макух Х. І., Чайківський Р. А., Зіменковський А. Б. Історія лікодїєзнавства: лікарські форми, еволюція їх технології та термінології // Клінічна фармація, фармакотерапія та медична стандартизація. – 2010. – № 3–4. – С. 108–117.

References

1. Yanishevska K. D., Balenko Ye. O. Administratyvno-pravove rehuliuвання доступу до publichnoi informatsii // Pravovi horyzonty. – 2018. – Vyp. 13 (26). – S. 86–91. <https://doi.org/10.21272/legalhorizons.2018.i13.p86>

2. Pavlov M. M. Osoblyvosti realizatsii kontseptu vidkrytykh danykh u systemi publichnoho upravlinnia: vitchyzniani ta zarubizhnyi dosvid // Derzhavne upravlinnia: teoriia ta praktyka. – 2018. – № 1. – S. 34–45.

3. Blynova N. M., Kyrylova O. V. Kopiraitynh yak diialnist zi stvorennia onlainovoho kontentu // Communications and Communicative Technologies (CCT). – 2018. – Vyp. 18. <https://doi.org/10.15421/291802>

4. Zubets V. V., Ylyna Y. V. O kachestve setevoi ynformatsyy // Vestnyk Tambovskoho unyversyteta. – 2016. – T. 21, Vyp. 12 (164). – S. 168–172. [https://doi.org/10.20310/1810-0201-2016-21-12\(164\)-168-172](https://doi.org/10.20310/1810-0201-2016-21-12(164)-168-172)

5. Kazymyrov Y. A. Problemy sozdanyia y pervychnoho napolnenyia bazy dannykh sdelok y predlozheny na rynke nedvyzhymosti // Vestnyk YrHTU. – 2014. – № 12. – S. 283–286.

6. Makukh Kh. I., Chaikivskiy R. A., Zimenkovskiy A. B. Istoriia likodїєznavstva: likarski formy, evoliutsiia yikh tekhnolohii ta terminolohii // Klinichna farmatsiia, farmakoterapiia ta medychna standartyzatsiia. – 2010. – № 3–4. – S. 108–117.

Надійшла до редакції 20 квітня 2019 р.

Прийнято до друку 10 червня 2019 р.

І. Ю. Рев'яцький (<https://orcid.org/0000-0002-4726-3639>)

Львівський національний медичний університет імені Данила Галицького

РЕЛЯЦІЙНА БАЗА ДАНИХ ІЗ ФАРМАЦЕВТИЧНОЮ ІНФОРМАЦІЄЮ: ПРОБЛЕМИ СТВОРЕННЯ ТА ПЕРВИННОГО НАПОВНЕННЯ ДЛЯ ЗАБЕЗПЕЧЕННЯ ЯКІСНОЇ СТАТИСТИЧНОЇ ОБРОБКИ

Ключові слова: фармацевтична інформація, лікарська форма, доза лікарського засобу, структура реляційної бази даних

АНОТАЦІЯ

Оприлюдновачі фармацевтичної інформації не ставлять за мету забезпечити можливість подальшої автоматизованої обробки представлених даних незалежними спеціалістами, і, відповідно, обирають методи її представлення, які спрямовані на оптимальне візуальне сприйняття користувачами.

Метою роботи було проаналізувати в інтернеті наявність загальнодоступної інформації у форматі, відповідному (підходящому) для автоматичного наповнення реляційної бази даних. Об'єктами

дослідження були: АТХ, МКХ-10, зміст ЄФ 9.8, класифікатори лікарських форм за МОЗ України, FDA (Food and Drug Administration, USA), ЕМА (European Medicines Agency) та EphMRA (European Pharmaceutical Market Research Association). Методи, що використовували: пошук інформації з відповідною структурою в інтернеті, обробка знайденої інформації з допомогою комп'ютерного коду.

Апробація проходила на базі електронної онлайн-системи «Ліквіпедія» (<http://likypedia.zzz.com.ua>; <http://facebook.com/likypedia>).

Сформовано базові цілі щодо якості результатів статистичної обробки фармацевтичної інформації. Для їх досягнення визначено перелік необхідної інформації для первинного завантаження до реляційної бази даних. Наведено джерела отриманої інформації, яка була використана для первинного завантаження до реляційної бази даних. Охарактеризовано полікомпонентність запису фармацевтичної інформації.

Перелік лікарських форм формувався на основі власного практичного досвіду, тенденцій зазначення лікарських форм на упаковках лікарських засобів виробниками, електронних базах даних фармацевтичних дистриб'юторів та на основі назв статей, наведених у ЄФ 9.8 видання. Інформацію із різних джерел про лікарські форми було проаналізовано та розділено на п'ять інформаційних блоків: виготовлена виробником лікарська форма; її характеристика; лікарська форма для застосування, яка буде приготовлена із виготовленої виробником лікарської форми; її характеристика; спосіб застосування лікарської форми.

Опрацьовано і представлено варіант структуризації запису інформації про дозу лікарського засобу в реляційній базі даних кількома варіантами, що забезпечує можливість подальшого здійснення автоматизованих розрахунків та оптимізацію відбору і сортування даних.

И. Ю. Ревяцкий (<https://orcid.org/0000-0002-4726-3639>)

Львовский национальный медицинский университет имени Данила Галицкого

РЕЛЯЦИОННАЯ БАЗА ДАННЫХ С ФАРМАЦЕВТИЧЕСКОЙ ИНФОРМАЦИЕЙ: ПРОБЛЕМЫ СОЗДАНИЯ И ПЕРВИЧНОГО НАПОЛНЕНИЯ ДЛЯ ОБЕСПЕЧЕНИЯ КАЧЕСТВЕННОЙ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ

Ключевые слова: фармацевтическая информация, лекарственная форма, доза лекарственного средства, структура реляционной базы данных

АННОТАЦИЯ

Обнародователи фармацевтической информации не ставят цели обеспечить возможность дальнейшей автоматизированной обработки представленных данных независимыми экспертами, и, соответственно, выбирают методы ее представления, направленные на оптимальное визуальное восприятие пользователями.

Цель работы – проанализировать в интернете наличие информации в формате, подходящем для автоматического наполнения реляционной базы данных. Объектами исследования были: АТХ, МКБ-10, содержание ЕФ 9.8, классификаторы лекарственных форм по МОЗ Украины, FDA (Food and Drug Administration, USA), по ЕМА (European Medicines Agency), по EphMRA (European Pharmaceutical Market Research Association). Использованы методы: поиск информации с соответствующей структурой в интернете, обработка полученной информации с помощью компьютерного кода.

Апробация была проведена на базе электронной онлайн-системы «Ликипедия» (<http://likypedia.zzz.com.ua>; <http://facebook.com/likypedia>).

Сформированы базовые цели по качеству результатов статистической обработки фармацевтической информации. Для их достижения определен перечень необходимой информации для первоначальной загрузки в реляционную базу данных. Приведены источники полученной информации, которая была использована для первоначальной загрузки в реляционную базу данных. Охарактеризована поликомпонентность записи фармацевтической информации.

Перечень лекарственных форм формировался на основе собственного практического опыта, тенденций указания лекарственных форм на упаковках лекарственных средств производителями, электронных базах данных фармацевтических дистрибьюторов и на основе названий статей, приведенных в ЕФ 9.8 издания. Информация из разных источников о лекарственных формах была проанализирована и разделена на пять информационных блоков: изготовленная производителем лекарственная форма; ее характеристика; лекарственная форма для применения, которая будет приготовлена из изготовленной производителем лекарственной формы; ее характеристика; способ применения лекарственной формы.

Разработан и представлен вариант структуризации записи информации о дозе лекарственного средства в реляционной базе данных несколькими вариантами, что обеспечивает возможность дальнейшего осуществления автоматизированных расчетов, оптимизации отбора и сортировки данных.

I. Yu. Revyatsky (https://orcid.org/0000-0002-4726-3639)

Danylo Halytsky Lviv National Medical University

A RELATIONAL DATABASE WITH PHARMACEUTICAL INFORMATION: PROBLEMS OF CREATION AND PRIMARY FILLING FOR PROVISION OF A QUALITATIVE STATISTICAL PROCESSING

Key words: pharmacy information, drug form, drug dose, a structure of a relational database

A B S T R A C T

The disclosure of pharmaceutical information is not intended to provide the possibility of further automated processing of the data provided by independent experts, and, accordingly, the choice of the methods of presentation is aimed to optimize visual perception by users.

The aim of the work was to analyze in the internet-sources the availability of information in the appropriate format for automatic filling of the relational database. The subjects of the study were: ATC: ICD-10 (International Classification of Diseases); content of EF 9.8; classification of dosage forms by the Ministry of Health of Ukraine, FDA (Food and Drug Administration, USA), EMA (European Medicines Agency) and EphMRA (European Pharmaceutical Market Research Association). Methods used are review of information with the corresponding structure in the internet, processing of the information found using computer code.

Approbation was carried out on the basis of the digital online system «Likypedia» (<http://likypedia.zzz.com.ua>; <http://facebook.com/likypedia>).

The basic goals for the quality of the results of statistical processing of pharmaceutical information are formed. To achieve them, a list of required information for the initial download to the relational database is defined. The sources of the received information which was used for initial loading to the relational database are presented. The multicomponent record of pharmaceutical information is described.

The list of dosage forms was formed on the basis of own practical experience, trends in the indication on packaging of medicines by manufacturers, electronic databases of pharmaceutical wholesalers, the titles of articles given in the EF 9.8 edition. Information from different sources about pharmaceutical dosage forms was analyzed and divided into five information blocks: made by manufacturer; its characteristic; for DF, which will be prepared with DF produced by the manufacturer; its characteristic; way of using DF.

A variant has been developed and presented for structuring the information record of the drug dose in the relational database in several variants, which makes it possible to further carry out automated calculations, optimize the selection and sorting of data.

Електронна адреса для листування з автором: iwan.revyatsky@gmail.com

(Рев'яцький І. Ю.)