**Информационные технологии и интеллектуальный анализ данных в диагностике гипопитуитаризма у детей и подростков**

***Русаков Михаил Олегович***

*Студент*

*Алтайский государственный университет,*

*Институт математики и информационных технологий, Барнаул, Россия*

*E-mail: mor1999@mail.ru*

Исследование посвящено изучению проблем диагностики и прогнозирования гипопитуитаризма у детей и подростков. Гипопитуитаризм – эндокринное заболевание, связанное со сниженной или отсутствующей секрецией одного или более гормонов гипофиза. Характерными симптомами заболевания являются снижение скорости роста, а также задержка общего развития [1].

Целью настоящего исследования является применение современных информационных технологий и интеллектуального анализа данных для решения актуальных задач диагностики и прогнозирования гипопитуитаризма. К основным задачам исследования относятся: формирование базы данных, предварительная обработка данных, прогнозирование SDS роста, применение методов машинного обучения для коррекции лечения, оценка рисков развития осложнений.

Актуальность и практическая значимость проводимого исследования обусловлены потребностью в применении средств интеллектуального анализа данных для изучения структуры заболевания, необходимостью достижения ускоренных темпов роста в первые годы лечения и их нормализации в последующем, а также снижением факторов риска развития сердечно-сосудистых заболеваний [2].

Важнейшей задачей исследования является структурирование информации и разработка базы данных. Данные для исследования представлены в виде обезличенных медицинских выписок, хранящихся в формате текстового редактора Microsoft Word.

Для структурирования информации и создания базы данных из медицинских выписок необходимо решить несколько задач: извлечение числовых и строковых данных из таблиц, нахождение в тексте упоминаний медицинских концептов (диагноз, сопутствующие заболевания, жалобы и симптомы и др.), извлечение числовых характеристик из текста [4]. Для решения поставленных задач выбран язык программирования Python (в частности, модуль re – предназначен для работы с регулярными выражениями; библиотеки: docx – для обработки файлов, созданных в текстовом редакторе Microsoft Word и NLTK – для обработки естественного языка).

Информация в базе данных структурирована на девять таблиц и содержит полную информацию о состоянии здоровья пациента на момент нахождения в стационаре, результаты анализов и проведенное лечение (рисунок 1). В качестве системы управления базами данных выбрана свободно распространяемая и имеющая открытый исходный код СУБД PostgreSQL.

В медицинских исследованиях, как и в практической медицине, спектр решаемых задач очень широк, что позволяет использовать различные методы интеллектуального анализа данных [3]. Результаты, полученные в ходе исследования, послужат основой для разработки системы поддержки принятия врачебных решений при диагностике и лечении гипопитуитаризма у детей и подростков.

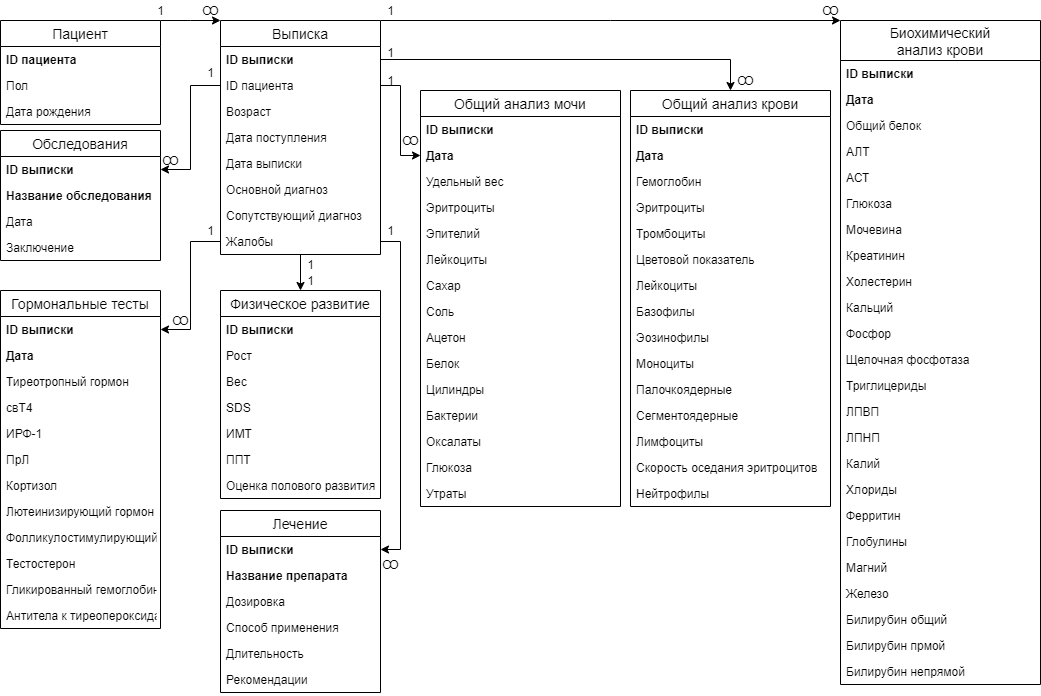


Рисунок 1. Логическая схема базы данных

**Литература**

1. Воронцова М.В. Гипопитуитаризм у детей и подростков // Медицинский совет. 2019. №2. – C. 250–258. DOI: https://doi.org/10.21518/2079-701X-2019-2-250-258.

2. Нагаева Е.В. Федеральные клинические рекомендации по диагностике и лечению гипопитуитаризма у детей и подростков // Проблемы эндокринологии. 2013. № 59(6). – С. 27–43.

3. Krotova O.S., Moskalev I.V., Nazarkina O.M., Khvorova L.A. Diagnostics of diabetic polyneuropathy in children and adolescents using data mining methods // Journal of Physics: Conference Series. 2020. Vol. 1615. Article ID 012015. doi:10.1088/1742-6596/1615/1/012015.

4. Moskalev I.V., Krotova O.S., Khvorova L.A., Bobkova D.G. Extraction of structured data from unstructured medical records using text data mining technologies: process automation // Journal of Physics: Conference Series. 2020. Vol. 1615. Article ID 012031. doi:10.1088/1742-6596/1615/1/012031.