

Майнинг данных по выживаемости онкологических больных. Сравнение существующих подходов.

Корытова А.О. Масленникова М.С.
Кузнецов С.О.

Национальный исследовательский университет «Высшая школа экономики»
Факультет компьютерных наук
Департамент анализа данных и искусственного интеллекта

Москва, 2015

Содержание

1 Введение

- Цели исследования
- Задачи исследования

2 Заключение

- Текущий прогресс
- Дальнейшие перспективы

3 Список литературы

Аннотация

Сегодня острый лимфобластный лейкоз(ОЛЛ) занимает одно из ведущих мест в структуре онкологических заболеваний у детей и подростков. Существующие на данный момент протоколы лечения ОЛЛ у детей дают высокие показатели уровня выживаемости, и в большинстве случаев речь идет о полном излечении пациента. Таким образом, перед исследователями встают новые задачи: достижение еще большей эффективности терапии путем модификации протокола по всем пациентам становится проблематично, и наиболее перспективным и дающим результаты способом решения этой ситуации является анализ в подгруппах. На сегодняшний день предложены некоторые подходы по выделению таких подгрупп пациентов. В данной работе будет представлено сравнение существующих подходов к выделению подгрупп, а также результаты тестирования новых гипотез, получаемых в результате модификаций предложенных подходов на имеющихся данных по протоколу MB-2002.

Abstract

Nowadays acute lymphoblastic leukemia (ALL) is among top oncologies of children and adolescents. Existing treatment protocols of ALL among children give high survival rate, and in most cases lead to complete cure of a patient. Thus, researchers are faced with new tasks: to achieve even greater efficiency for all patients we need to use subgroup analysis. There are some approaches to allocate these subgroups of patients. This work contains a comparison of the existing approaches to the subgroup isolation, as well as the results of testing new hypotheses resulted from modifications of existing approaches suggested by protocol data MB-2002.

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%
- В 1990 году был введен в использование модифицированный под условия российских клиник немецкий протокол лечения BFM-90

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%
- В 1990 году был введен в использование модифицированный под условия российских клиник немецкий протокол лечения BFM-90
- В 1991 году совместно с клиникой Шарите (Берлин, Германия) был разработан отечественный протокол лечения MB-91

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%
- В 1990 году был введен в использование модифицированный под условия российских клиник немецкий протокол лечения BFM-90
- В 1991 году совместно с клиникой Шарите (Берлин, Германия) был разработан отечественный протокол лечения MB-91
- Первое клиническое рандомизированное исследование MB-91 против BFM - 90. Выживаемость составила 74% и 71% соответственно

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%
- В 1990 году был введен в использование модифицированный под условия российских клиник немецкий протокол лечения BFM-90
- В 1991 году совместно с клиникой Шарите (Берлин, Германия) был разработан отечественный протокол лечения MB-91
- Первое клиническое рандомизированное исследование MB-91 против BFM - 90. Выживаемость составила 74% и 71% соответственно
- В 2002 году был запущен новый вариант протокола MB-2002. Исследование было закончено в 2008 году, и выживаемость составила 78%

Введение

- До начала 90-х годов выживаемость при детском ОЛЛ составляла меньше 10%
- В 1990 году был введен в использование модифицированный под условия российских клиник немецкий протокол лечения BFM-90
- В 1991 году совместно с клиникой Шарите (Берлин, Германия) был разработан отечественный протокол лечения MB-91
- Первое клиническое рандомизированное исследование MB-91 против BFM - 90. Выживаемость составила 74% и 71% соответственно
- В 2002 году был запущен новый вариант протокола MB-2002. Исследование было закончено в 2008 году, и выживаемость составила 78%
- В этом же году было запущено новое исследование, и началось внедрение анализа в подгруппах с целью повышения эффективности существующих протоколов лечения

Введение

Цели исследования

- Тестирование новых гипотез, получаемых в результате применения модификаций предложенного подхода к выделению подгрупп на имеющихся данных по протоколу MB-2002
- Сравнение подходов по выделению подгрупп

Введение

Задачи исследования

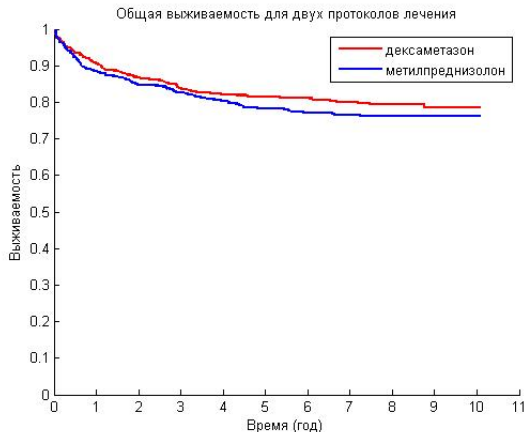
- Изучение литературы по введению в предметную область
- Изучение методов медико-биологической статистики
- Обзор и сравнение существующих методов оценки выживаемости
- Построение кривых выживаемости по имеющимся данным по протоколу MB-2002
- Проверка новых гипотез
- Анализ полученных результатов

Текущий прогресс

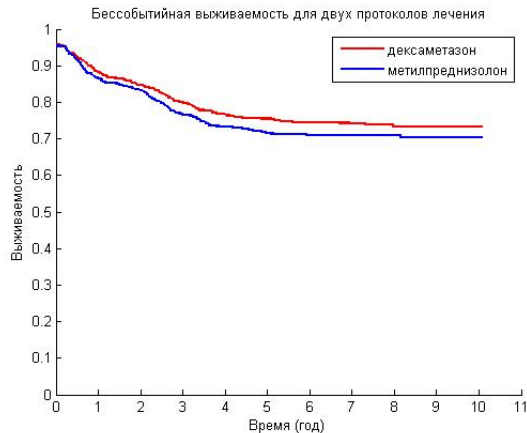
- Проведен обзор литературы по предметной области
- Изучен метод построения кривых выживаемости Каплана - Майера

Текущий прогресс

- Построены графики общей выживаемости и произведен анализ двух разных протоколов лечения:
 - пациенты, получающие дексаметазон
 - пациенты, получающие метилпреднизолон








- Построены графики бессобытийной выживаемости



Дальнейшие перспективы

- Применение к данным и сравнение результатов различных методов отбора признаков в программной среде для анализа данных Weka
- Проверка новых гипотез по протоколу MB-2002
- Изучение различных методов оценки выживаемости и сравнение их на данных протокола MB-2002

Список литературы

-  Karachunskiy, A., Herold, R., Stackelberg, A. et al. Results of the first randomized multicenter trial on childhood acute lymphoblastic leukaemia in Russia. Leukemia. 00, 1-10 (2008).
-  E. L. Kaplan and Paul Meier, Journal of the American Statistical Association, Vol. 53, No. 282 (Jun., 1958), pp. 457- 481
-  Румянцева Ю.В. Риск-адаптированная терапия острого лимфобластного лейкоза у детей и подростков в исследовании ALL-MB-2002: автореф. дис. д.м.н.: 14.01.08, 14.01.21/Юлия Васильевна Румянцева; Москва, ФНКЦ ДГОИ – М., 2011. – 61с.
-  Стентон Гланц, Медико-биологическая статистика, 1999г.
-  Алексеев Е.Р., Чеснокова О.В. MATLAB 7. Самоучитель. ISBN: 5-477-00283-2. Издательство "НТ Пресс"2006г. 464 стр.

Список литературы

-  Pintilie M. Competing risks., A practical perspective. John Wiley Sons, Ltd, (2011).
-  Armitage P., Berry G., Matthews J. N. S., Statistical methods in medical research, Blackwell Science Ltd (2002).
-  Esteve J., Benhamou E., Raymond L., Descriptive epidemiology, Lyon, International Agency for Research on Cancer, (1994).
-  Слинин А.С., Конкурирующие риски. Практическое применение в медицине., статья 1999г.
-  Maimon O., Rokach L., Data mining and knowledge discovery handbook, Springer Science+Business Media, LLC, (2010).

Спасибо за внимание!