

## Современное программное обеспечение для статистической обработки биомедицинских исследований



Герасевич Виталий Александрович кандидат медицинских наук, кафедра кардиологии и ревматологии Белорусской медицинской академии последипломного образования;



Аветисов Арам Рубенович кандидат медицинских наук, доцент кафедры радиационной медицины и экологии БГМУ

Одним из обязательных этапов любого научного исследования является статистический анализ данных. Продолжительное время анализ медицинских данных был уделом специалистов, так как это требовало серьезной предварительной подготовки. С появлением и совершенствованием современных программ обработки данных статистическая обработка поднялась на новый уровень. Теперь исследователь-медик может и не иметь математической подготовки. Достаточно оперировать статистическими понятиями и, самое главное, правильно выбрать метод анализа. Все осуществимо благодаря компьютеру и новейшим программам.

Все программы статистической обработки данных можно разделить на профессиональные, полупрофессиональные (популярные) и специализированные. Статистические программы относятся к наукоемкому программному обеспечению, цена их часто недоступна индивидуальному пользователю. Профессиональные пакеты имеют большое количество методов анализа, популярные пакеты - количество функций, достаточное для универсального применения. Специализированные же пакеты ориентированы на какую-либо узкую область анализа данных. Создатели программных статистических пакетов заявляют, что их продукт превосходит аналоги. Отсутствие у

большинства исследователей времени для освоения нескольких программ, делает непростым ее выбор. В данной статье приведена базовая информация о присутствующих на рынке основных полупрофессиональных программных пакетах пригодных для статистической обработки биомедицинских данных.

**MS Excel.** Самой часто упоминаемой (и используемой) в отечественных статьях является приложение MS Excel из пакета офисных программ компании Microsoft – MS Office. Причины этого кроются в широком распространении этого программного обеспечения, наличии русскоязычной версии, тесной интеграцией с MS Word и PowerPoint. Однако, MS Excel - это электронная таблица с достаточно мощными математическими возможностями, где некоторые статистические функции являются просто дополнительными встроенными формулами. Расчеты сделанные при ее помощи не признаются авторитетными биомедицинскими журналами. Также в MS Excel невозможно построить качественные научные графики. Безусловно, MS Excel хорошо подходит для накопления данных, промежуточного преобразования, предварительных статистических прикидок, для построения некоторых видов диаграмм. Однако окончательный статистический анализ необходимо делать в программах, которые специально созданы для этих целей. Существует макрос-дополнение XLSTAT-Pro (<http://www.xlstat.com>) для MS Excel который, включает в себя более 50 статистических функций, включая анализ выживаемости, которых в основных случаях достаточно для обычного применения. Пробную версию макроса можно взять на сайте производителя.

**STADIA.** Программа отечественной разработки с 16-и летней историей. Включает в себя все необходимые статистические функции. Она прекрасно справляется со своей задачей - статистическим анализом. Но. Программа внешне фактически не изменяется с 1996 года. Графики и диаграммы, построенные при помощи STADIA, выглядят в современных презентациях архаично. Цветовая гамма программы (красный шрифт на зеленом) очень утомляет в работе. К положительным качествам программы можно отнести русскоязычный интерфейс и наличие книг описывающих работу. Например: Кулаичев А.П. Методы и средства анализа данных в среде Windows. - М: ИнКо, 2002. - 341 с. Со страницы <http://www.protein.bio.msu.su/~akula/index.htm> можно взять демо-версию STADIA.

**SPSS (Statistical Package for Social Science).** Самый часто используемый пакет статистической обработки данных с более чем 30-и летней историей (<http://www.spss.com/>). Отличается гибкостью, мощностью применим для всех видов статистических расчетов применяемых в биомедицине. Недавно вышла 13-я англоязычная версия. Существует русскоязычное представительство компании (<http://www.spss.ru/>) которое предлагает полностью русифицированную версию SPSS 12.0.2 для Windows. Появился учебник на русском языке, позволяющий шаг за шагом освоить возможности SPSS, репетитор по статистике на русском языке, помогающий в выборе нужной статистической или графической процедуры для конкретных данных и задач, а также справка по SPSS Base и SPSS Tables. Российский офис SPSS регулярно проводит учебные курсы по анализу данных при помощи программного обеспечения SPSS. На русский язык переведена книга по SPSS, которая вышла в свет в 2002 году в Киевском издательстве «Диасофт» под названием «SPSS 10: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей».

**STATA.** Профессиональный статистический программный пакет с data-management system, который может применяться для биомедицинских целей. Один из

самых популярных в образовательных и научных учреждениях США наряду с SPSS. Официальный сайт <http://www.stata.com>. Программа хорошо документирована, издается специальный журнал для пользователей системы. Однако возможности предварительного ознакомления с демо-версией нет.

**STATISTICA.** Производителем программы является фирма StatSoft Inc. (США) (<http://www.statsoft.com/>), которая выпускает статистические приложения, начиная с 1985 года. STATISTICA включает большое количество методов статистического анализа (более 250 встроенных функций) объединенных следующими специализированными статистическими модулями: Основные статистики и таблицы, Непараметрическая статистика, Дисперсионный анализ, Множественная регрессия, Нелинейное оценивание, Анализ временных рядов и прогнозирование, Кластерный анализ, Факторный анализ, Дискриминантный функциональный анализ, Анализ длительностей жизни, Каноническая корреляция, Многомерное шкалирование, Моделирование структурными уравнениями и др. Несложный в освоении этот статистический пакет может быть рекомендован для биомедицинских исследований любой сложности.

В настоящее время выпущена версия 7. Российское представительство компании (<http://www.statsoft.ru/>) предлагает полностью русифицированную 6-ю версию программы. Сайт компании содержит много информации по статистической обработке медицинских данных, учебник по статистике на русском языке. Сам пакет STATISTICA описан в нескольких книгах, одна из которых, для медицинских работников: О.Ю. Реброва «Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA.» Москва, МедиаСфера, 2002. 312 с. Ознакомиться с содержанием книги можно по адресу: [http://mediasphera.aha.ru/book/rebr\\_sod.htm](http://mediasphera.aha.ru/book/rebr_sod.htm)

**JMR.** Один из мировых лидеров в анализе данных. Развивает этот статистический пакет SAS Institute (<http://www.jmp.com/>), который выкупил в конце 2002 года известную статистическую программу StatView. Однако особых преимуществ для медико-биологической статистики этот программный продукт не имеет.

**SYSTAT.** Статистическая система для персональных компьютеров (<http://systat.com>). Последняя 11 версия обладает неплохим интуитивно понятным интерфейсом. Компания Systat Software также разрабатывает популярные у отечественных исследователей SigmaStat и SigmaPlot, которые являются соответственно, программой статистической обработки и программой построения диаграмм. При совместной работе становятся единым пакетом для статистической обработки и визуализации данных.

**NCSS.** Программа развивается с 1981 года и рассчитана на непрофессионалов в области статистической обработки. Интерфейс системы многооконный и как следствие этого явления - немного непривычный в использовании. Все действия пользователя сопровождаются подсказками. Сейчас доступна версия 2004 г. С сайта (<http://www.ncss.com>) можно переписать полнофункциональную пробную версию работающую 30 дней.

**MINITAB 14.** Статистический пакет MINITAB в настоящее время выпускается в версии 14. С сайта производителя (<http://www.minitab.com>) можно взять полнофункциональный пробный вариант программы, которая работает 30 дней. Это достаточно удобный в работе программный пакет, имеющий хороший интерфейс пользователя, хорошие возможности по визуализации результатов работы. Имеет подробную справку.

**STATGRAPHICS PLUS.** Довольно мощная статистическая программа. Содержит

более 250 статистических функций, генерирует понятные, настраиваемые отчеты. Последняя доступная версия - 5.1. Ее можно получить на сайте <http://www.statgraphics.com>. Есть возможность скачать демо-версию. Следует отметить, что ранние версии этой программы были весьма популярны у отечественных исследователей.

PRISM. Эта программа создавалась специально для биомедицинских целей. Интуитивно понятный интерфейс позволяет в считанные минуты проанализировать данные и построить качественные графики. Программа содержит основные часто применяемые статистические функции, которых в большинстве исследований будет достаточно. Однако, как отмечают сами разработчики, программа не может полностью заменить серьезных статистических пакетов. На сайте <http://www.graphpad.com> помимо возможности ознакомления с демо-версией Prism можно получить справочник в формате PDF по биомедицинской статистике.

Дополнительная информация. В настоящее время в Интернет доступны многие ресурсы, посвященные статистической обработке данных. Один из них - это статистический портал, созданный при содействии В. П. Боровикова, автора книг по программному пакету STATISTICA (<http://www.statsoft.ru/home/portal/>). Российское представительство StatSoft Inc. предлагает на своем сайте бесплатный электронный учебник по статистике, который призван помочь разобраться с основными понятиями статистики и более полно представить диапазон применения статистических методов (<http://www.statsoft.ru/home/download/textbook/default.htm>). На этом же сайте существует «Статистический медицинский советник», который поможет правильно выбрать нужный статистический метод (<http://www.statsoft.ru/home/portal/applications/medicine/medadvisor.htm>). Из ресурсов Интернет заслуживает внимания сайт с пятилетней историей – Биометрика (<http://www.biometrika.tomsk.ru/>). Из книг заслуживает внимания Primer of Biostatistics, S.A.Glantz, 4-е издание. Она была переведена на русский язык издательским домом «Практика» под названием «Медико-биологическая статистика». В книге содержатся объяснения статистических методов на примерах, применяемых в биомедицине. При этом она не содержит тяжелых выкладок формул. Автор - С.Гланц, перевод с английского под ред. Н. Е. Бузикашвили и Д. В. Самойлова. В издательстве РУДН в 2002 году вышла книга Лукьяновой Е.А., адаптированная для нетехнической аудитории под названием «Медицинская статистика». Могут быть полезны и другие книги: Боровиков В. Statistica. Искусство анализа данных на компьютере. Для профессионалов. 2-е изд. Изд-во «Питер», 2003, 688 стр.; Лапач С.Н. и др. Статистические методы в медико-биологических исследованиях с использованием Excel. Издательство «Морион Лтд», 2000, 320 стр.

На какой программе остановить свой выбор? Безусловно, дороговизна программ не позволяет их менять. Поэтому имеет смысл посмотреть демо-версии, разобраться с работой и потом делать окончательный вывод. Русскоязычные версии (с документацией) имеют только SPSS и STATISTICA. Что касается возможных рекомендаций, то они следующие:

- Если нужен мощный, общепризнанный пакет с простым и понятным даже начинающим пользователям интерфейсом, то лучше воспользоваться SPSS.

- Для начинающих и профессионалов, которым нужна подсказка и развитая документация на русском языке, можно рекомендовать STATISTICA. Это мощное приложение с профессиональными возможностями.

- Для непритязательных пользователей, которые ограничиваются в своих исследованиях стандартными статистическими методами можно рекомендовать англоязычную программу Prism.