

УДК [616.12] https://doi.org/10.17816/MAJ19327-36

SAS ENTERPRISE GUIDE 6.1 ДЛЯ ВРАЧЕЙ: НАЧАЛО РАБОТЫ

Н.С. Буненков, Г.Ф. Буненкова, С.А. Белый, В.В. Комок, О.А. Гриненко, А.С. Немков

ФГБНУ «Первый Санкт-Петербургский медицинский университет им. акад. И.П. Павлова» Минздрава России, Санкт-Петербург

Для цитирования: Буненков Н.С., Буненкова Г.Ф., Белый С.А., и др. SAS Enterprise Guide 6.1 для врачей: начало работы // Медицинский академический журнал. – 2019. – Т. 19. – № 3. – С. 27–36. https://doi.org/10.17816/MAJ19327-36

Поступила: 25.07.2019

Одобрена: 12.08.2019

Принята: 29.08.2019

Цель — разработать алгоритм обработки базы данных проспективного нерандомизированного исследования AMIRI–CABG Trial (ClinicalTrials.gov Identifier: NCT03050489) в программном пакете SAS Enterprise Guide 6.1.

Материалы и методы. В проспективное нерандомизированное исследование AMIRI–CABG, проведенное на базе ПСПбГМУ им. И.П. Павлова с 2016 по 2019 г., включено 336 пациентов. Создана база данных с клиническими данными, результатами лабораторных и инструментальных исследований. Статистическая обработка данных выполнена с помощью лицензионного программного обеспечения SAS Enterprise Guide 6.1.

Результаты. Реализован алгоритм обработки данных проспективного нерандомизированного исследования AMIRI–CABG. Данный алгоритм может быть использован врачами различных специальностей, научными сотрудниками для обработки результатов научных или клинических исследований.

Заключение. Реализованный алгоритм обработки базы данных научного исследования позволит врачам и научным сотрудникам упростить и ускорить анализ результатов исследований. SAS Enterprise Guide 6.1 дает возможность качественно и быстро обработать большие массивы данных.

Ключевые слова: SAS Enterprise Guide; статистическая обработка данных; статистика; клинические исследования; научные исследования.

SAS ENTERPRISE GUIDE 6.1 FOR PHYSICIANS: GETTING STARTED

N.S. Bunenkov, G.F. Bunenkova, S.A. Beliy, V.V. Komok, A.A. Grinenko, A.S. Nemkov

Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia

For citation: Bunenkov NS, Bunenkova GF, Beliy SA, et al. SAS Enterprise Guide 6.1 for physicians: Getting started. *Medical Academic Journal.* 2019;19(3):27-36. https://doi.org/10.17816/MAJ19327-36

Received: July 25, 2019

Revised: August 12, 2019

Accepted: August 29, 2019

Objective. To develop algorithm of data analysis of prospective non-randomized clinical trial AMIRI–CABG (ClinicalTrials.gov Identifier: NCT03050489) using SAS Enterprise Guide 6.1.

Materials and methods. Data collection was performed according prospective non-randomized clinical trial AMIRI–CABG in Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia between 2016–2019 years with 336 patients. There is database with clinical, laboratory and instrumental data. Statistical analysis was performed with SAS Enterprise Guide 6.1.

Results. There was developed algorithm of data analysis of prospective non-randomized clinical trial AMIRI–CABG. This algorithm could be useful for physicians and researchers for data analysis.

Conclusion. Presented algorithm of data analysis could make easier and improve efficient data analysis. SAS Enterprise Guide 6.1 allows fast and accurate process big data.

Keywords: SAS Enterprise Guide 6.1; statistical analysis; statistics; clinical trials; scientific research.

Список сокращений

ИК — искусственное кровообращение; КШ — коронарное шунтирование; GCP — надлежащая клиническая практика (от англ. good clinical practice); TnI1 — переменная, содержащая значения концентрации тропонина I на первые сутки после операции; LGTnI1 — логарифм TnI1.

Введение

В повседневной работе врачи и ученые сталкиваются с задачами анализа данных. С развитием науки и техники требования, предъявляемые к методологии и принципам проведения исследований с участием пашиентов, становятся все более строгими. Так, с целью повышения объективности результатов исследования должны быть выполнены в соответствии с принципами доказательной медицины [1]. Следование стандартам надлежащей клинической практики (good clinical practice — GCP) общепринято среди авторов публикаций в ведущих международных журналах [2]. Повышаются требования не только к качеству проведения исследований, но и к обработке полученных данных, а для этого необходимы определенные знания статистики. Для врача, а тем более хирурга, приоритетной областью приложения усилий является клиническая работа, что подразумевает проведение значительного времени в операционной. Глубокое погружение врача-хирурга в математическую статистику на уровне профильных специалистов, имеющих математическое образование, вряд ли целесообразно и осуществимо с учетом объема клинической работы. Таким образом, врач оказывается в ситуации, когда, с одной стороны, результаты его работы необходимо обрабатывать с использованием все более сложного математического аппарата, а с другой — изучение методов обработки данных может начать конкурировать по временным затратам с основной деятельностью.

К счастью, современные программы статистической обработки данных позволяют решить дилемму постоянно возрастающих требований к методам обработки данных и оптимальным

Файл	Изменить Вид Задачи Программа Сервис	Cr	Справка 🛛 🎽 🖌 🊰 🖓 🖓	
<u></u>	Создать	•	🕨 🤯 Проект	1
	Эткрыть Закрыть проект Сохранить Проект как Ампорт данных Выполнить Проект Планировать Проект Планировать Проект Планировать Проект Экранетры страницы для Схема процесса Тредарительный просмотр печати для Схема процесса	• 5 •	Данные Данные Программа Отчет Хранимый процесс Примечание Георадоченный список	Ma
C C C C C C C C C C C C C C C C C C C	сойства проекта Тоследние проекты Тоследние программы Ънкод	• •	•	



погружением врача в статистический анализ. Так, например, в программном пакете SAS Enterprise Guide 6.1 для статистического анализа данных достаточно знать границы применимости критерия и название процедуры, что дает возможность врачу избегать использования сложного математического аппарата с большим количеством формул и в то же время получать качественные расчеты.

Цель данной работы состояла в написании алгоритма обработки данных проспективного нерандомизированного исследования Assessment of Myocardial Ischemic-Reperfusion Injury during Off- and On- Pump CABG AMIRI–CABG (ClinicalTrials.gov Identifier: NCT03050489) в программном пакете SAS Enterprise Guide 6.1.

Материалы и методы

В проспективное нерандомизированное исследование AMIRI-CABG. проведенное на базе Научно-исследовательского центра серлечно-сосулистой хирургии ПСПбГМУ им. И.П. Павлова с 2016 по 2019 г., было включено 336 пациентов с ишемической болезнью сердца, имевших показания к операции коронарной реваскуляризации. Создана база данных с клиническими данными, результатами лабораторных и инструментальных исследований. Статистическая обработка базы данных выполнена с использованием лицензионного программного обеспечения SAS Enterprise Guide 6.1. Все анализируемые переменные проверяли на нормальность распределения следующими способами:

- визуально анализировали график (типичный колокол);
- применяли критерий Колмогорова Смирнова и Шапиро – Уилка.

Для переменных, имеющих нормальное распределение, применяли параметрические методы статистики. Если переменная имела распределение, отличное от нормального и являлась лабораторным показателем (например, уровень тропонина I), проводили логарифмирование (натуральный логарифм или десятичный). Проверяли на нормальность логарифмированную переменную. Если нормальность подтверждалась, то использовали методы параметрической статистики, если нет — непараметрической. Если распределение отличалось от нормального, то такую статистику по переменной представляли в виде медианы (верхний и нижний квартили). Если распределение было нормальное, то статистику представляли в виде среднего (стандартное отклонение или ошибка среднего) [3]. Создание файла проекта представлено на рис. 1.

Результаты и их обсуждение

Загрузку базы данных из файла Excel в SAS Enterprise guide 6.1 (все переменные должны быть на латинице и не содержать конфликтных символов) выполняли через меню «Файл» — «Импорт данных» (рис. 2, 3).

В открывшемся окне выбирали файл Excel (см. рис. 3).

Примечание. Название каталогов и файлов должно быть только на латинице. Если в пути к файлу будет хотя бы один каталог на кириллице, программа выдаст ошибку. Кроме того, файл Excel должен быть в версии 98-2003, более поздние версии могут не читаться. В Excel нужно выбрать сохранение таблицы в версии 98-2003.



Рис. 2. Импорт файла с базой данных Fig. 2. Excel database import

После выбора файла в открывшемся окне нажимали «Вперед» (рис. 4).

Ставили галочку напротив «Переименовать столбцы в соответствии с требованиями SAS Enterprise Guide 6.1», отмечали лист Excel, с которым будем работать (в нашем случае In hospital), нажимали «Вперед» (рис. 5).

В появившемся окне проверяли, что числовым переменным соответствует число, а строковым — строка (рис. 6).

Иногда SAS Enterprise Guide 6.1 определял числа как строки, в таком случае в колонке «Тип» нажимали на «Строка», выбирали «Число» (рис. 7).

Проверив и исправив, если требуется, типы переменных, нажимали **«Вперед»** (рис. 8).

Искать в:	WorkStat		~ 4	- 🖻 🛛 🗶 🍃	🔲 = \$5		
 Рабочки стол Мои документы Мой констьютер Спересе Сорржене Серееры 	Visa Visa 2019 June 2019 June 2019 Vork Stat (2 2 Vork Stat (2 2	Дата изменения) log10 vork ABG Macros afterCABG 2018 zo18 дена Tnl (конец оп	Тип Ц ерацин)	Размер Индексы Юден исходные значе	адля МПО ния		
	Имя файла:	(Tells			dah Carda Carad	and the state of the	~
	initi wannoa:	чаилы импортир	уемых данн	eer (los; lodsx;"lodsm;"	xeep; .mdb;".accdl	5; .Dd; .CsV; .as	0; ~

Рис. 3. Выбор файла с базой данных Fig. 3. File Excel import

Дерево проекта 🔹 🤉	с Схема процесса -				
E Reg Otema npouecca 과 3월 20:06:2019 work.xls 내 관련 Minnopt gameaix (20:06-2019 wo	BuinonHum CortaHOBITE BeinonHum CortaHOBITE BeinonHum CortaHOBITE BeinonHum CortaHOBITE BeinonHum CortaHOBITE BeinonHum CortaHOBITE BeinonHum Second Second	орт • Расписание • Масштаб •	🎲 Журнал проекта 📄 Свойства 🔹		
		Импортданны 1 из 4 Зад	их из 20-06-2019 work.xts ать данные		sas.
		Мастер импорта д необходимые для	E DIVENE SAS,		
		Фала данных источ Расположение: Путь к файлу:	ника Покальная файловая система C:\WorkStat\Ume 2019D\20-06-2019 work.xls		
		Тип данных: Выходной набор да	RHHR'S EXCel		
		Сервер SAS: Библиотека:	Local WORK	Ofdsop	
		Набор данных:	20_06_2019 work		
				chianan y Rosean For	

Рис. 4. Выбор файла с базой данных (продолжение) **Fig. 4.** File Excel import (next step)

Copaexa 🛛 🎦 • 🚰 • 🍕 | 🏯 🤟 🖄 🎊 🗡 | 19 🕫 | 🗖 • | kg Crewa opouecca 6 • | 🔂 🛪 кта | 🔚 Сво * Серенс Справка 🎼 - 🚔 - 🍕 🖂 🤟 🖄 🖄 🗡 💷 СК 🔲 - 🔩 Скема пр 6 - 172.8 кта 📔 Свойства 🔸 <u>*</u>---. 20-06-2019 Импорт данных work.xls (20-06-2019 work.xls) a Ma Ssas 2 из 4 Выбрать источник дан Первая строка диапазона содержит имена полей Cfpoc ar сНазад ▼ Вперед> Готово Отинена Справка

Рис. 5. Выбор листа в файле Excel Fig. 5. List Excel selection

30



Рис. 6. Проверка типа переменных **Fig. 6.** Variables type checking

церево проекта 🔹 👻	beens appugeces +											
Eng Clerva ripsuecca ⊕ ∰ 20-06-2019 work.xls ⊢∰ Инпорт данных (20-06-2015 work ⊢∰ Инпорт данных (20-06-2015 work	Выполнить • Ш Остановить Экспорт •	Расписание •	Масштаб •	Журнал проек	та 📑 Свойства	•						
	workuls (20-6-2019 workuls)											
		3 v	мпорт данных из из 4 Опреде	20-06-2019 worl лить атрибут	uds пы поля					<u>S</u> .Si	as	
		Buttep	ите столбцы и опр	еделите атрибут	ы:							
		Bkn	Имя источника	Имя	Ярлык	Tim	Информат источника	Дл.	Формат вывода	Информат вывода		
			AI	AI	AI	4wono -	BEST12.	8	BEST12.	BEST12.		
			LCx	LCx	LCx	Ywcno	BEST12	8	BEST12	BEST12.		
			OM	OM	OM	Yucno	BEST12.	8	BEST12	BEST12.		
			RCA	RCA	RCA	Yercho	BEST12.	8	BEST12	BEST12.		
			PDA	PDA	PDA	Yucno	BEST12.	8	BEST12.	BEST12.		
			EI	EII	EII	Строка	\$CHAR4.	4	\$CHAR4.	\$CHAR4.		
			\$\$1	\$\$1	\$\$1	Hincho	BEST12.	8	8EST12.	BEST12.		
			\$\$2	\$\$2	\$\$2	Ywono	BEST12.	8	BEST12	BEST12.		
						Строка 🗸						
			ICD	ICD	ICD.	Строка	BEST12.	8	BEST12	BEST12.		
			MDChart	MDChart_0003	MDChart	Валота	\$CHAR23.	23	\$CHAR23.	SCHAR23.		
			Name	Name_0001	Name	Время	\$CHAR37.	37	\$CHAR37.	SCHAR37.		
			CP8Tme	CPBTime	CPSTime	Дата и время	BEST12.	8	BEST12.	BEST12.		
			AoClamp	AcClamp	AcClamp	Число	BEST12.	8	BEST12.	BEST12.		
			OpType	OpType_0003	OpType	Чиспо	BEST12.	8	BEST12.	BEST12.		
		Lon		17						E Marine	-	

Рис. 7. Исправление типа переменных **Fig. 7.** Variables type correction

Fig. 8. Completing database import



7.3	к Импо	рт данных (2	20-06-20)19 work.xls[In	hospita	• ([]•														
		ioa 🔟 Xio	нал 🖉	Выходные д	531910-10															
0.06.2019	0	Изменить	задачу	Фильтр	и сорти	ировка 🖓 Р	онструктор :	sanpocos	s Данные	Описательная статист	ика - График	- Анализировать	- Экспорт - Отпр	авить - 🔢						
		Strai	n0 6	Strain1	1	Strain2	0 1	V K	Strain	Agreement	Call	A Note	Call_date	Dep	A Name	MDChart	PreopDays	IOAdrenalin 😡	IONor G	HDyn
	1		1					1						2		10672/C2019	6	0.05	0	
	2							2				0		1		10696/C2016	15	0	0.03	
	3							3				0		2		11933/C2016	7	0.003	0.1	
	4							4				0		1		12082/C2016	15	0	0	
	5							5				0		2		12254/C2016	13	0.01	0.07	
	6							6						1		12774/C2017	7	0	0	
	7							7				0		2		13656/C2016	6	0.001	0.02	
	8						1	8				<u>Î</u>		2		14702/C2018	15	0.02	0.4	
	9				-		1	9		+				2		17896/C2017	2	0.007	0.07	
	10		1		1		1	10		1				1		18213/02018	28	0	0.02	
	11							11				0		2		18633/C2016	14	0.003	0.01	
	12						1	12					42767	2		19432/02016	6	0	0.1	
	12				-		-	13						1		20156/02017		0	0.1	
	14						1	15				-	-	2		20552(02019	7	0.01	0.2	
					-		-	16					-	1		21114/02017		0.005	0.26	
	10				-			17				-	1			22244/02017		0.000	040	
	10				-		+	10				*	-			22244/02017	0	0.003	0.07	
	10				-			10		-			1			225220002017	24	0.003	0.07	
	10						-	20		. 0.10		-				25522/02017		0	0.5	
	13							21		. 0.10						20004102017	10	0	0	
	20				-		<u>.</u>	21		. ects		-		1		20038/02017	16	0.01	01	
	21		-					22						2		2/8/202017	16	0.01	0.1	
	22		+		-		-	23				-		1		2885/C2018	14	0.007	0.2	
	23				1			24		. есть				1		29242/0.2017	13	0	U	
	24				-			25		. •		0		2		29399/C2016	8	0.005	0.3	
	25		-		-		-4	26				0		2		30041/C2016	6	0.01	0.3	
	26							27		. •		1 ошибка номера	43137	1		30858/C2016	8	0	0	
	27				-			28		. •		0 нет номера		2		31713/C2016	9	0.02	0.2	
	28						-	29								31715/C2018	7	0	0	
	29							30		. •			1	2		32792/C2016	9	0.005	0.05	
	30				-			34		. •				2		33947/C2016	9	0.05	0.05	
	31		+				4	35		. •			43143	1		34031/C2016	15	0	0	
	32						3	36		- C.			A			34113/C2018	17	0	0	
	33				-			39						1		34881/C2017	24	0	0	
	34		1		14		3	41		. +			1	2		35160/C2016	10	0.01	0.3	
	35						2	48						2		37825/C2017	8	0.003	0.05	
	36		+				1	50				2	1	1		39165/C2017	10	0.13	0	
	37							52						2		42207/C2018	8	0.01	0.1	
	> 38				14		2	53						2		43017/C2018	16	0.1	0.2	
-	39							54						2		43098/C2018	8	0	0.4	
1	× 40		-		-			55						1		43180/C2018	23	0	0.6	
	41							58						1		43482/C2018	30	0	0.8	
1	42				-		+	58		. +		-		1		4560/C2017	15	0	0	
2	43							60		. •			43143	1		5070/C2017	6	0	0	
ŝ	44							61		. +				2		5550/C2017	6	0.05	0.05	
	45						2	62					1	2		6043/C2017	9	0.01	0.1	
	40						4	63						2		6098/C2019	26	0.01	0.1	



Отмечали галочку напротив «Удалить символы, которые могут привести к ошибкам», нажимали «Готово» (рис. 9).

Загруженные в SAS Enterprise Guide 6.1 данные обрабатывали при помощи методов статистики.

Для статистической обработки можно использовать интуитивно понятный графический интерфейс или текстовый ввод процедур. Каждый из способов имеет свои преимущества и недостатки. Мы выбрали текстовый ввод процедур, так как этот метод является более гибким. Для этого выбирали **«Файл»** → **«Создать»** → **«Программа»** (рис. 10).

Теперь в открывшемся окне можно записать процедуры для статистических расчетов. Возьмем группу пациентов после коронарного шунтирования (КШ) без искусственного кро-

Φai	л Изменить Вид Задачи Г	рограмма Сервис	Справка 📋 🖬 🚰 🦋	-0	≦ X 5) (8 ⊡ • 648 C	хема проц
	Создать		,	3	Проект	
2	Открыть		,	12	Данные	
	Закрыть проект			A.	Программа	писат
3	Сохранить Проект		Ctrl+S	10	Отчет	
	Сохранить Проект как			581	Хранимый процесс	
				-	-	•
2	Импорт данных			13	Примечание	•
•	Выполнить Проект			840	Схема процесса	•
	Планировать Проект			冟	Упорядоченный список	•
	Экспорт		,	1	7	
	Отправить		,	-	8	
41	Публиковать			-	9	+
				1	10	1
ĉa -	Параметры страницы для Данные, и	імпортированные из 2	0-06-2019 work.xls		11	. +
	Предварительный просмотр печати	для Данные, импорти	рованные из 20-06-2019 work.xls		12	. +
a	Печать Данные, импортированные	as 20-06-2019 work.xls	Ctrl+P		13	. +
-	C-12			ŀ	15	
•	Своиства проекта			·	16	
	Последние проекты		,	<u>-</u>	12	
	Последние программы		,	-	10	
	Выход			i-	20	ecth
-		20			21	есть
		21			22	
		22			23	
		23		-	24	. есть
		24			25	. +
		25			26	. +
		26			27	. +
		27		-	25	. *
		28		-	20	
		30		-	34	
		21		-	25	

Рис. 10. Создание программы **Fig. 10.** Program list creating



Рис. 11. Указатель на базу данных **Fig. 11.** DATASET pointer

Файл Изменить Вид Задачи [рограмма 🛛 Сервис 🖉 правка 🛛 🗃 🕶 🥳 🖓 🛛 📇 🌮 🖄 🖾 🛪 🖬 🦳 🖉 🖛 Слав процесс
Дерево проекта • >	Программа •
Breg Ceeka populacia → 32 2042 (501 work.ds → 43 (204 control and	A Foressent > Degressent B Cappeners. > Degressent B Cappeners. > Degressent B Prool UNIVARIAT Bunchmers Floorpanneselocit 9 Prool UNIVARIAT Bunchmers Floorpanneselocit 9 Work*n normaltest plots. VAR Thil Age; run; ods graphics off;

Рис. 12. Запуск программы Fig. 12. Run program

вообращения (ИК), обработаем для примера переменную TnI1, которая содержит значения тропонина I на первые сутки после операции КШ, а также переменную Age (возраст).

В нашей таблице содержатся данные по трем группам: КШ без ИК, КШ с ИК и КШ на параллельном ИК. При использовании процедур необходимо указывать название таблицы, с которой работаешь. Чтобы узнать название таблицы, нужно щелкнуть мышью два раза на вкладку «Импорт данных» (голубая рамка), выбрать «Код» (красная рамка), скопировать имя таблицы после слова DATA (зеленая рамка) (рис. 11).

Затем следует нажать два раза вкладку «Программа» и записать код, позволяющий оценить характер распределения:

```
ods graphics on;
Proc UNIVARIATE DATA=WORK.'20_06_2019
work'n normaltest plots;
where CPBType=1;
VAR TnI1 Age;
run;
ods graphics off;
```

Примечание. СРВТуре — переменная, обозначающая тип КШ. Может принимать значения от 1 до 3 (1 — без ИК, 2 — с ИК, 3 — параллельное ИК).

normaltest — сообщает программе о том, что нужно проверить нормальность.

plots — сообщает программе о необходимости построить графики (можно не указывать, если достаточно числовых данных).

where — сообщает программе, по какому признаку осуществлять отбор подгруппы из совокупности данных. Если убрать where, будет обработано все количество пациентов.

VAR — сообщает программе о том, какие переменные анализировать, здесь можно перечислить большое количество переменных. В нашем случае для наглядности указано только две.

ods graphics on — включить графический режим, чтобы были красивые графики.

ods graphics off — выключить графический режим.



Рис. 13. Анализ распределения **Fig. 13.** Distribution analysis

После этого следует нажать кнопку «Выполнить» (рис. 12).

Получили результат в виде таблиц и графиков, при помощи которых можно установить характер распределения, медиану, квартили, среднее, моду, дисперсию и ряд других параметров. Наиболее быстрый способ определить нормальность распределения — визуально оценить графики (рис. 13).

В случае нормального распределения вид гистограммы приближается к форме колокола, а точки укладываются в прямую линию. Хорошо видно, что переменная Age (возраст) имеет нормальное распределение (точки легли на прямую линию, гистограмма похожа на колокол). Кроме того, признаком нормальности распределения является симметричный «ящик с усами» и небольшое количество выбросов (outliers). Переменная TnI1, напротив, характеризуется распределением, отличным от нормального, так как гистограмма не имеет формы колокола, несимметрична, много выбросов, точки легли в виде кривой.

Из множества табличных данных интерес представляют критерии нормальности для TnI1 (рис. 14).

Если p (Value) меньше 0,05, то распределение отличается от нормального, если больше 0,05, то распределение соответствует нормальному. Переменная TnI1 имеет распределение, отличное от нормального.

Распределение переменной Аge соответствовало нормальному (рис. 15).

Все значения *р* больше 0,05, что говорит о нормальности распределения.

После определения характера распределения переменных можно выбрать соответствующий статистический метод и сравнить группы. Так, зная, что уровень тропонина I на первые сутки после КШ без ИК имеет характер распределения, отличный от нормального, мы не будем использовать для сравнения групп критерий Стьюдента (*t*-критерий). Применение критерия Стьюдента в такой ситуации может ложно показать отсутствие различия.

Выбор критериев и тестов для сравнения групп заслуживает отдельной статьи, здесь же ограничимся описанием практического использования программы SAS Enterprise Guide 6.1.

Следует отметить, что SAS Enterprise Guide 6.1 сохраняет схему процесса, но не сохраняет результаты всех процедур, поэтому при закрытии и новом открытии SAS Enterprise Guide 6.1 следует выполнять всю схему процесса. Для этого, открыв проект, нужно правой кнопкой мыши нажать на схему процесса, затем выбрать **«Выполнить»** (рис. 16).

Tests for Normality										
Test Statistic p Value										
Shapiro-Wilk	W	0.255611	Pr < W	< 0.0001						
Kolmogorov-Smirnov	D	0.391175	Pr > D	< 0.0100						
Cramer-von Mises	W-Sq	8.797194	$Pr > W_Sq$	< 0.0050						
Anderson-Darling	A-Sq	41.72411	Pr > A-Sq	< 0.0050						

Рис. 14. Критерии нормальности для переменной TnI1 Fig. 14. Normality tests for TnI1

Tests for Normality										
Test	St	atistic	p Value							
Shapiro-Wilk	W	0.990338	Pr < W	0.2615						
Kolmogorov-Smirnov	D	0.052546	Pr > D	>0.1500						
Cramer-von Mises	W-Sq	0.053161	Pr > W-Sq	>0.2500						
Anderson-Darling	A-Sq	0.34096	Pr > A-Sq	>0.2500						

Рис. 15. Критерии нормальности для переменной Age **Fig. 15**. Normality tests for Age



Рис. 16. Выполнение схемы процесса **Fig. 16.** Schema running

Приведем к нормальному виду распределение TnI1, чтобы можно было использовать методы параметрической статистики:

```
ods graphics on;
Proc UNIVARIATE DATA=WORK.'20 06 2019
work'n normaltest plots;
where CPBType=1;
  VAR TnI1 Age;
run;
DATA LG; /* СОЗДАЕМ НОВУЮ ТАБЛИЦУ LG */
  SET WORK.'20 06 2019 work'n;
  /* КОПИРУЕМ ВСЮ НАШУ ТАБЛИЦУ
                                B LG
                                      */
  LGTnI1=LOG10 (TnI1);
  /*
      ВЫЧИСЛЯЕМ ДЕСЯТИЧНЫЙ ЛОГАРИФМ
                                      */
RUN;
Proc UNIVARIATE DATA=LG normaltest plots;
/* ТЕПЕРЬ МЫ РАБОТАЕМ С ТАБЛИЦЕЙ LG */
  VAR LGTnI1;
run;
ods graphics off;
```



Рис. 17. Распределение логарифмированной переменной Fig. 17. Lognormal distribution

Tests for Normality										
Test Statistic p Va										
Shapiro-Wilk	W	0.995578	Pr < W	0.5340						
Kolmogorov-Smirnov	D	0.034077	Pr > D	>0.1500						
Cramer-von Mises	W-Sq	0.055329	Pr > W-Sq	>0.2500						
Anderson-Darling	A-Sq	0.397005	Pr > A-Sq	>0.2500						

Рис. 18. Критерии нормальности для логарифмированной переменной

Fig. 18. Normality tests for lognormal distribution

После нажатия кнопки «Выполнить» получаем результат (рис. 17).

Нормальность распределения подтвердилась не только графиками, но и четырьмя критериями, p > 0.05 (рис. 18).

Таким образом, для переменной LGTnI1 можно использовать методы параметрической статистики, например критерий Стьюдента.

Обсуждение

Представленные нами результаты являются итогом предварительного этапа обработки данных, обязательного для любого исследования, который позволяет выбрать адекватные статистические тесты. Следует отметить, что концентрация маркеров и ферментов очень часто имеет логнормальное распределение, поэтому с практической точки зрения переменные, содержащие концентрации, можно сразу логарифмировать, после чего проверять нормальность распределения. Несмотря на то что в случае распределения, отличающегося от нормального, можно использовать непараметрические тесты, предпочтительнее перейти к нормальному виду, взяв десятичный или натуральный логарифм, так как многие статические тесты были разработаны для нормального распределения. Проверка на нормальность является простой процедурой, которую можно выполнить в любой программе для статистических расчетов, однако приведенный пошаговый и подробный алгоритм может стать первым шагом к освоению программы SAS Enterprise Guide 6.1, имеющей мощные инструменты обработки данных, которые гораздо удобнее по сравнению с инструментами, предлагаемыми в других пакетах статистических расчетов. Реализованный алгоритм представлен с таким расчетом, чтобы читатели, никогда не работавшие с данной программой, могли сразу же использовать приведенный выше код для своей базы данных, поменяв названия переменных и файла в соответствии с базой данных. SAS — один из самых популярных программных пакетов, используемых для обработки данных медицинских исследований в США, и снабжен набором инструментов для работы с большими объемами данных [4]. По сравнению с другими статистическими программами SAS обладает следующими преимуществами:

- построение схемы расчетов позволяет проверить правильность выбора и применения статистических методов. Таким образом, видны не только конечный результат расчетов, но и промежуточные результаты, что уменьшает вероятность ошибки в процессе обработки данных;
- возможность использования не только графического интерфейса, но и программного кода, благодаря чему можно проводить более гибкие расчеты;
- популярные статистические тесты выполняются автоматически, отсутствует необходимость в специальном вызове. Так, например, при выполнении дисперсионного анализа SAS автоматически проверяет равенство дисперсий, в случае их неравенства вносит необходимую поправку.

Выводы

- 1. SAS Enterprise Guide 6.1 обладает гибкими и мощными инструментами статистической обработки данных с минимальными требованиями к владению математическим аппаратом, что делает этот программный пакет интересным для врачей.
- 2. SAS Enterprise Guide 6.1 позволяет быстро и легко провести статистические расчеты.

3. Алгоритм проверки на нормальность может быть применен для обработки различных баз данных научных и клинических исследований.

Дополнительная информация

Финансирование. Работа выполнена в рамках государственного задания по теме: «Оценка регенеративного потенциала пациента при операциях на сердце».

Соблюдение этических норм. Исследование одобрено локальным этическим комитетом ФГБНУ ПСПбГМУ им. И.П. Павлова.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Литература

- Koretz RL. Assessing the evidence in evidence-based medicine. *Nutr Clin Pract.* 2019;34(1):60-72. https://doi. org/10.1002/ncp.10227.
- May M. Twenty-five ways clinical trials have changed in the last 25 years. *Nat Med.* 2019;25(1):2-5. https://doi. org/10.1038/s41591-018-0314-1.
- Гржибовский А. М. Типы данных, проверка распределения и описательная статистика // Экология человека. – 2008. – № 1. – С. 52–58. [Grjibovski AM. Data types, control of distribution and descriptive statistics. *Ecology*, *human.* 2008;(1):52-58. (In Russ.)]
- Dolezel D, McLeod A. Big data analytics in healthcare: investigating the diffusion of innovation. *Perspect Health Inf Manag.* 2019;16(Summer):1a.

Сведения об авторах / Information about the authors

Николай Сергеевич Буненков — аспирант кафедры факультетской хирургии Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. https://orcid.org/0000-0003-4331-028X. E-mail: bunenkov2006@gmail.com.

Гульнара Физулиевна Буненкова — клинический ординатор кафедры госпитальной терапии Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. E-mail: gulnara533@ gmail.com.

Сергей Алексеевич Белый — канд. мед. наук, кардиохирург отделения кардиохирургии № 2 Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. https://orcid.org/ 0000-0002-6415-4389. E-mail: sabel1968@mail.ru. *Nikolay S. Bunenkov* – Aspirant, Department of Faculty Surgery, Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. https://orcid.org/0000-0003-4331-028X. E-mail: bunenkov2006@gmail.com.

Gulnara F. Bunenkova — Resident, Department of Hospital Therapy, Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. E-mail: gulnara533@ gmail.com.

Sergey A. Beliy — PhD, Cardiac Surgeon, Department of Cardiac Surgery No. 2, Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. https://orcid.org/0000-0002-6415-4389. E-mail: sabel1968@mail.ru.

Владимир Владимирович Комок — канд. мед. наук, кардиохирург отделения кардиохирургии № 2 Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. https://orcid.org/ 0000-0002-3834-7566. E-mail: vladimir komok@mail.ru.

Олег Александрович Гриненко — д-р мед. наук, проректор по лечебной работе Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. E-mail: klinika@spb-gmu.ru.

Александр Сергеевич Немков — д-р мед. наук, профессор, кардиохирург, руководитель отделения кардиохирургии № 2 Первого Санкт-Петербургского медицинского университета им. И.П. Павлова, Санкт-Петербург. https://orcid.org/0000-0002-5152-0001. E-mail: nemk_as@mail.ru.

Контактное лицо / Corresponding author

Николай Сергеевич Буненков / Nikolay S. Bunenkov E-mail: bunenkov2006@gmail.com

Vladimir V. Komok — PhD, Cardiac Surgeon, Department of Cardiac Surgery No. 2 of Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. https:// orcid.org/0000-0002-3834-7566. E-mail: vladimir_komok@ mail.ru.

Oleg A. Grinenko — Doctor of Medical Science, Vice-Rector of Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. E-mail: klinika@spbgmu.ru.

Alexander S. Nemkov — Doctor of Medical Science, Professor, Cardiac Surgeon, Chief of Department of Cardiac Surgery No. 2 of Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia. https://orcid. org/0000-0002-5152-0001. E-mail: nemk_as@mail.ru.