



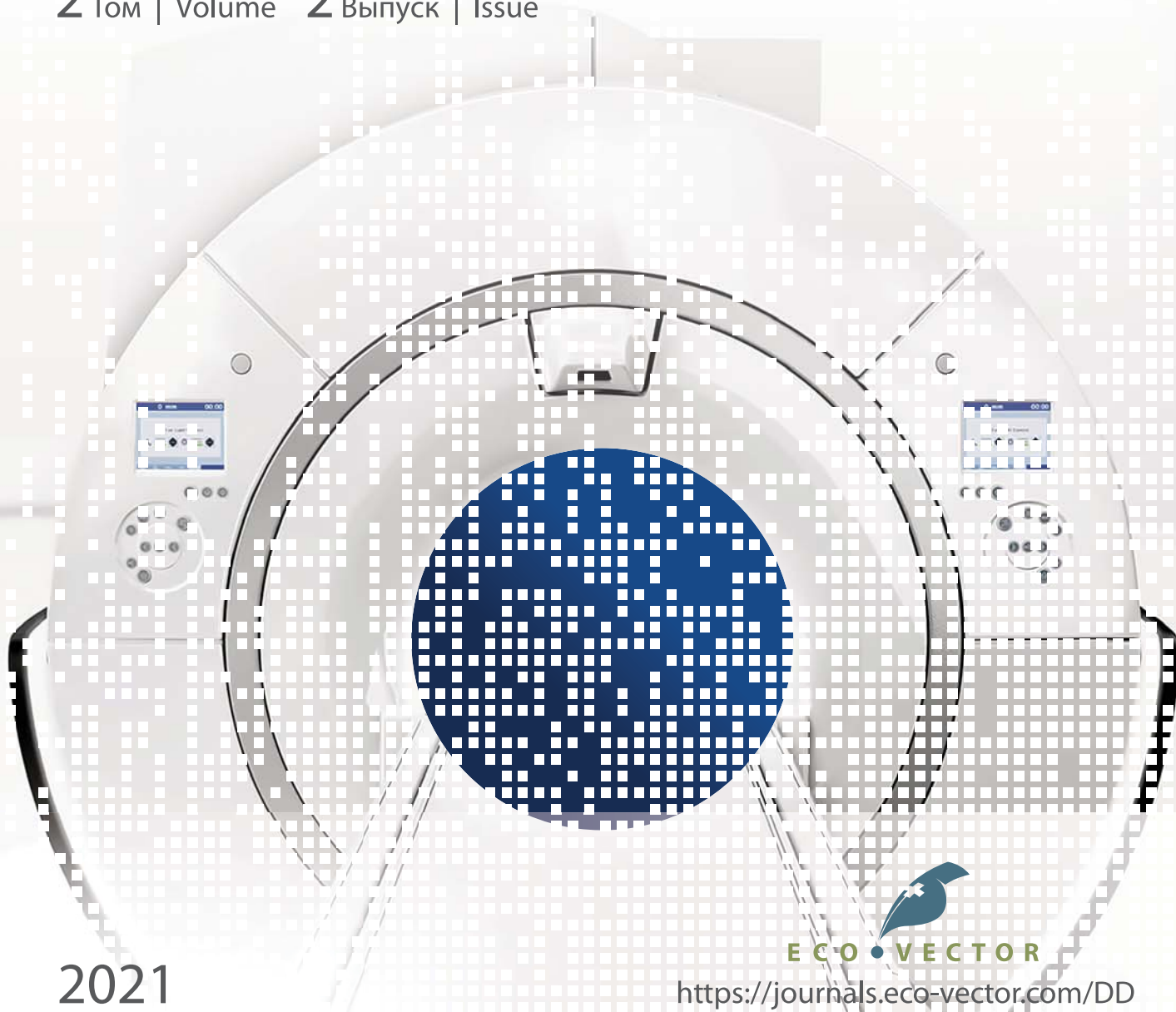
ЦЕНТР ДИАГНОСТИКИ
И ТЕЛЕМЕДИЦИНЫ

ISSN 2712-8490 (Print)
ISSN 2712-8962 (Online)

DIGITAL DIAGNOSTICS

Рецензируемый научный медицинский журнал

2 Том | Volume 2 Выпуск | Issue



2021



<https://journals.eco-vector.com/DD>

УЧРЕДИТЕЛИ

- ГБУЗ «Научно-практический клинический центр диагностики и телемедицинских технологий ДЗМ»
- ООО «Эко-Вектор»

ИЗДАТЕЛЬ

ООО «Эко-Вектор»

Адрес: 191186, г. Санкт-Петербург, Аптекарский переулочек, д. 3, литера А, помещение 1Н

E-mail: info@eco-vector.com

WEB: <https://eco-vector.com>

РЕКЛАМА

Отдел рекламы

Тел.: +7 495 308 83 89

РЕДАКЦИЯ

Зав. редакцией

Елена Андреевна Филиппова

Email: ddjournal@eco-vector.com

Тел: +7 (965) 012 70 72

ПОДПИСКА

Подписка на печатную версию через интернет:

www.journals.eco-vector.com/

www.akc.ru

www.pressa-rf.ru

OPEN ACCESS

В электронном виде журнал распространяется бесплатно — в режиме немедленного открытого доступа

ИНДЕКСАЦИЯ

- РИНЦ
- Google Scholar
- Ulrich's International Periodicals Directory
- WorldCat

Оригинал-макет

подготовлен в издательстве «Эко-Вектор».

Литературный редактор: *М.Н. Шошина*

Корректор: *М.Н. Шошина*

Верстка: *Ф.А. Игнащенко*

Обложка: *И.С. Феофанова*

Сдано в набор 20.07.2021.

Подписано в печать 30.07.2021.

Формат 60 × 88%. Печать офсетная.

Печ. л. 18,5. Усл. печ. л. 17,205.

Уч.-изд. л. 10,09. Тираж 500 экз. Заказ № 1-5121-1v

Отпечатано в ООО «Типография Михаила Фурсова».

196105, Санкт-Петербург, ул. Благодатная, 69.

Тел.: (812) 646-33-77

© ООО «Эко-Вектор», 2021

ISSN 2712-8490 (Print)

ISSN 2712-8962 (Online)

Digital Diagnostics

Том 2 | Выпуск 2 | 2021

ЕЖЕКВАРТАЛЬНЫЙ РЕЦЕНЗИРУЕМЫЙ НАУЧНЫЙ
МЕДИЦИНСКИЙ ЖУРНАЛ

Главный редактор

Синицын Валентин Евгеньевич, д.м.н., профессор (Москва, Россия)

ORCID: 0000-0002-5649-2193

Заместитель главного редактора

Морозов Сергей Павлович, д.м.н., профессор (Москва, Россия)

ORCID: 0000-0001-6545-6170

Редакционная коллегия

Андрейченко А.Е., к.ф.-м.н. (Москва, Россия)

ORCID: 0000-0001-6359-0763

Berlin L., профессор (Иллинойс, США)

ORCID: 0000-0002-0717-0307

Беляев М.Г., к.ф.-м.н. (Москва, Россия)

ORCID: 0000-0001-9906-6453

Bisdas S., MBBS, MD, PhD (Лондон, Великобритания)

ORCID: 0000-0001-9930-5549

Гомболевский В.А., к.м.н. (Москва, Россия)

ORCID: 0000-0003-1816-1315

Frija G., профессор (Париж, Франция)

ORCID: 0000-0003-0415-0586

Guglielmi G., MD, профессор (Фоджа, Италия)

ORCID: 0000-0002-4325-8330

Holodny A., д.м.н. (Нью-Йорк, США)

ORCID: 0000-0002-1159-2705

Li H., MD, профессор, (Пекин, КНР)

Кульберг Н.С., к.ф.-м.н. (Москва, Россия)

ORCID: 0000-0001-7046-7157

Mannelli L., MD (Нью-Йорк, США)

ORCID: 0000-0002-9102-4176

Мокиенко О.А., к.м.н. (Москва, Россия)

ORCID: 0000-0002-7826-5135

Neri E., д.м.н. (Пиза, Италия)

ORCID: 0000-0001-7950-4559

van Ooijen P., к.м.н. (Гронинген, Нидерланды)

ORCID: 0000-0002-8995-1210

Oudkerk M., профессор (Гронинген, Нидерланды)

ORCID: 0000-0003-2800-4110

Ros P.R., MD, MPH, PhD, профессор (Нью-Йорк, США)

ORCID: 0000-0003-3974-0797

Rovira A., профессор (Барселона, Испания)

ORCID: 0000-0002-2132-6750

Решетников Р.В., к.ф.-м.н. (Москва, Россия)

ORCID: 0000-0002-9661-0254

Румянцев П.О., д.м.н. (Москва, Россия)

ORCID: 0000-0002-7721-634X

Редакционный совет

Аншелес А.А., д.м.н. (Москва, Россия)

ORCID: 0000-0002-2675-3276

Арутюнов Г.П., д.м.н. (Москва, Россия)

ORCID: 0000-0002-6645-2515

Белевский А.С., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0001-6050-724X

Васильева Е.Ю., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0003-4111-0874

Гехт А.Б., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0002-1170-6127

Кобякова О.С., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0003-0098-1403

Кремнева Е.И., к.м.н. (Москва, Россия)

ORCID: 0000-0001-9396-6063

Петриков С.С., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0003-3292-8789

Проценко Д.Н., к.м.н. (Москва, Россия)

ORCID: 0000-0002-5166-3280

Хатьков И.Е., д.м.н., профессор (Москва, Россия)

ORCID: 0000-0002-4088-8118

Редакция не несет ответственности за содержание рекламных материалов. Точка зрения авторов может не совпадать с мнением редакции. К публикации принимаются только статьи, подготовленные в соответствии с правилами для авторов. Направляя статью в редакцию, авторы принимают условия договора публичной оферты. С правилами для авторов и договором публичной оферты можно ознакомиться на сайте: <https://journals.eco-vector.com/DD/>. Полное или частичное воспроизведение материалов, опубликованных в журнале, допускается только с письменного разрешения издателя — издательства «Эко-Вектор».



FOUNDERS

- Moscow Center for Diagnostics and Telemedicine
- Eco-Vector

PUBLISHER

Eco-Vector

Address: 3 liter A, 1H, Aptekarsky pereulok, 191186, Saint Petersburg Russian Federation
E-mail: info@eco-vector.com
WEB: <https://eco-vector.com>

ADVERTISE

Adv. department

Phone: +7 (495) 308 83 89

EDITORIAL

Executive editor

Elena A. Philippova

Email: ddjournal@eco-vector.com

Phone: +7(965)0127072

SUBSCRIPTION

For print version:

www.journals.eco-vector.com/

PUBLICATION ETHICS

Journal's ethic policies are based on:

- ICMJE
- COPE
- ORE
- CSE
- EASE

OPEN ACCESS

Immediate Open Access is mandatory for all published articles

INDEXATION

- Russian Science Citation Index
- Google Scholar
- Ulrich's International Periodicals Directory
- WorldCat

TYPESET

complete in Eco-Vector

Copyeditor: *M.N. Shoshina*

Proofreader: *M.N. Shoshina*

Layout editor: *Ph. Ignashchenko*

Cover: *I. Feofanova*

ISSN 2712-8490 (Print)

ISSN 2712-8962 (Online)

Digital Diagnostics

Volume 2 | Issue 2 | 2021

QUARTERLY PEER-REVIEW MEDICAL JOURNAL

EDITOR-IN-CHIEF

Valentin E. Sinitsyn, MD, Dr.Sci. (Med), Professor (Moscow, Russia)

ORCID: 0000-0002-5649-2193

DEPUTY EDITOR-IN-CHIEF

Sergey P. Morozov, MD, Dr.Sci. (Med), Professor (Moscow, Russia)

ORCID: 0000-0001-6545-6170

Редакционная коллегия

A.E. Andreychenko, PhD (Moscow, Russia)

ORCID: 0000-0001-6359-0763

L. Berlin, Professor (Illinois, United States)

ORCID: 0000-0002-0717-0307

M.G. Belyaev, Cand.Sci. (Phys-Math), Assistant

Professor (Moscow, Russia)

ORCID: 0000-0001-9906-6453

S. Bisdas, MBBS, MD, PhD (London, United Kingdom)

ORCID: 0000-0001-9930-5549

V.A. Gomboleviskiy, MD, Dr.Sci. (Med) (Moscow, Russia)

ORCID: 0000-0003-1816-1315

G. Frija, Professor (Paris, France)

ORCID: 0000-0003-0415-0586

G. Guglielmi, MD, Professor (Foggia, Italy)

ORCID: 0000-0002-4325-8330

A. Holodny, MD (New-York, United States)

ORCID: 0000-0002-1159-2705

H. Li, MD, Professor (Beijing, China)

N.S. Kul'berg, Cand.Sci. (Phys-Math) (Moscow, Russia)

ORCID: 0000-0001-7046-7157

L. Mannelli, MD (New-York, United States)

ORCID: 0000-0002-9102-4176

O.A. Mokienko, MD, Cand.Sci. (Med) (Moscow, Russia)

ORCID: 0000-0002-7826-5135

E. Neri, MD, Associate Professor (Pisa, Italy)

ORCID: 0000-0001-7950-4559

P. van Ooijen, PhD, Assoc. Professor (Groningen, Netherlands)

ORCID: 0000-0002-8995-1210

M. Oudkerk, Professor (Groningen, Netherlands)

ORCID: 0000-0003-2800-4110

P.R. Ros, MD, MPH, PhD, Professor (New-York, United States)

ORCID: 0000-0003-3974-0797

A. Rovira, Professor (Barcelona, Spain)

ORCID: 0000-0002-2132-6750

R.V. Reshetnikov, Cand.Sci. (Phys-Math) (Moscow, Russia)

ORCID: 0000-0002-9661-0254

P.O. Rummyantsev, MD, Dr.Sci. (Med) (Moscow, Russia)

ORCID: 0000-0002-7721-634X

Редакционный совет

A.A. Ansheles, MD, Dr.Sci. (Med)

(Moscow, Russia)

ORCID: 0000-0002-2675-3276

G.P. Arutyunov, MD, Dr.Sci. (Med)

(Moscow, Russia)

ORCID: 0000-0002-6645-2515

A.S. Belevskiy, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0001-6050-724X

E.Y. Vasilieva, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0003-4111-0874

A.B. Gekht, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0002-1170-6127

O.S. Kobyakova, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0003-0098-1403

E.I. Kremneva, MD, Cand.Sci. (Med)

(Moscow, Russia)

ORCID: 0000-0001-9396-6063

S.S. Petrikov, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0003-3292-8789

D.N. Protsenko, MD, Cand.Sci. (Med)

(Moscow, Russia)

ORCID: 0000-0002-5166-3280

I.E. Khatkov, MD, Dr.Sci. (Med), Professor

(Moscow, Russia)

ORCID: 0000-0002-4088-8118

The editors are not responsible for the content of advertising materials. The point of view of the authors may not coincide with the opinion of the editors. Only articles prepared in accordance with the guidelines are accepted for publication. By sending the article to the editor, the authors accept the terms of the public offer agreement. The guidelines for authors and the public offer agreement can be found on the website: <https://journals.eco-vector.com/DD/>. Full or partial reproduction of materials published in the journal is allowed only with the written permission of the publisher — the Eco-Vector publishing house.

СОДЕРЖАНИЕ

ОРИГИНАЛЬНЫЕ ИССЛЕДОВАНИЯ

Д.А. Филатова, В.Е. Сеницын, Е.А. Мершина

Возможности снижения лучевой нагрузки

при проведении компьютерной томографии для оценки изменений в лёгких, характерных для COVID-19:

использование адаптивной статистической итеративной реконструкции 94

Н.С. Кульберг, Р.В. Решетников, В.П. Новик, А.Б. Елизаров, М.А. Гусев,

В.А. Гомбелевский, А.В. Владимирский, С.П. Морозов

Вариабельность заключений при интерпретации КТ-снимков: один за всех и все за одного 105

ОБЗОРЫ

J.P. Vandenbroucke, E. von Elm, D.G. Altman, P.C. Gøtzsche, C.D. Mulrow, S.J. Pocock,

C. Poole, J.J. Schlesselman, M. Egger for the STROBE Initiative

Повышение качества отчётов о наблюдательных исследованиях в эпидемиологии (STROBE):

разъяснения и уточнения 119

А.Н. Хоружая, Е.С. Ахмад, Д.С. Семенов

Роль системы контроля качества лучевой диагностики онкологических заболеваний в радиомике 170

В.Г. Говорухина, С.С. Семенов, П.Б. Гележе, В.В. Диденко, С.П. Морозов, А.Е. Андрейченко

Роль маммографии в радиомике рака молочной железы 185

Е.Н. Гиря, В.Е. Сеницын, А.С. Токарев

Кавернозные мальформации головного мозга и современные взгляды на их лечение 200

S. Taylor-Phillips, C. Stinton

Проблема утомления в лучевой диагностике: многообещающая область для научных исследований 211

КЛИНИЧЕСКИЕ СЛУЧАИ

К.В. Прусакова, П.В. Гаврилов

Длительный анамнез бронхоцеле, вызванный типичным карциноидом 223

CONTENTS

ORIGINAL STUDIES

Daria A. Filatova, Valentin E. Sinitsyn, Elena A. Mershina

Opportunities to reduce the radiation exposure during computed tomography to assess the changes in the lungs in patients with COVID-19: use of adaptive statistical iterative reconstruction 94

Nikolas S. Kulberg, Roman V. Reshetnikov, Vladimir P. Novik, Alexey B. Elizarov, Maxim A. Gusev, Victor A. Gomboleviskiy, Anton V. Vladzimirskyy, Sergey P. Morozov

Inter-observer variability between readers of CT images: all for one and one for all. 105

REVIEWS

Jan P. Vandenbroucke, Erik von Elm, Douglas G. Altman, Peter C. Gotzsche, Cynthia D. Mulrow, Stuart J. Pocock, Charles Poole, James J. Schlesselman, Matthias Egger for the STROBE Initiative

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. Translation to Russian 119

Anna N. Khoruzhaya, Ekaterina S. Akhmad, Dmitry S. Semenov

The role of the quality control system for diagnostics of oncological diseases in radiomics. 170

Veronika G. Govorukhina, Serafim S. Semenov, Pavel B. Gelezhe, Vera V. Didenko, Sergey P. Morozov, Anna E. Andreychenko

The role of mammography in breast cancer radiomics 185

Elena N. Giry, Valentin E. Sinitsyn, Aleksei S. Tokarev

Cavernous malformations of the brain and modern views on their treatment 200

Sian Taylor-Phillips, Chris Stinton

Fatigue in radiology: a fertile area for future research 211

CASE REPORTS

Kseniya V. Prusakova, Pavel V. Gavrilov

Long-term broncocele anamnesis, triggered by typical carcinoid. 223

DOI: <https://doi.org/10.17816/DD62477>

Возможности снижения лучевой нагрузки при проведении компьютерной томографии для оценки изменений в лёгких, характерных для COVID-19: использование адаптивной статистической итеративной реконструкции

Д.А. Филатова, В.Е. Сеницын, Е.А. Мершина

Московский государственный университет имени М.В. Ломоносова, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Большинство пациентов с COVID-19 во время госпитализации проходит многократные визуализационные обследования, кумулятивный эффект которых может значительно увеличивать общую дозу полученного облучения. Эффективная доза облучения может быть снижена за счёт уменьшения тока и напряжения рентгеновской трубки, что, однако, снижает качество изображения. Возможным решением этой проблемы может стать внедрение технологии адаптивной статистической итерационной реконструкции «сырых данных» компьютерной томографии (КТ) — Adaptive Statistical Iterative Reconstruction (ASIR). В последнее время в литературе появились сведения об эффективности низкодозной КТ (НДКТ) в диагностике COVID-19.

Цель — анализ качества и диагностической ценности НДКТ-изображений лёгких после применения итеративного алгоритма обработки; оценка возможности снижения лучевой нагрузки на пациента при диагностике COVID-19.

Материал и методы. В проспективном исследовании приняли участие пациенты, проходившие стационарное лечение в инфекционном отделении МНОЦ МГУ им. М.В. Ломоносова. Исследования КТ выполнялись при поступлении и выписке; в период госпитализации их повторяли по мере клинической необходимости. При первом исследовании использовался стандартный протокол КТ с напряжением тока на трубке 120 кВ и автоматическим модулированием силы тока в диапазоне 200–400 мА, при повторных КТ применяли протокол НДКТ с уменьшенными параметрами напряжения тока на трубке (100 или 110 кВ) и автоматической модуляцией тока в диапазоне 40–120 мА. Для оценки диагностической ценности НДКТ по сравнению со стандартной КТ было проведено анкетирование среди врачей отделения лучевой диагностики МНОЦ МГУ. Анкета включала в себя сравнительную характеристику двух методик при выявлении таких патологических процессов, как уплотнение лёгочной ткани по типу матового стекла, уплотнение по типу матового стекла с ретикулярными изменениями, участки консолидации лёгочной ткани, лимфаденопатия.

Результаты. В исследовании принял участие 151 пациент; средний возраст $58 \pm 14,2$ года; 53,6% мужчин. При НДКТ в сравнении со стандартной КТ лучевая нагрузка снижалась в среднем в 2,96 раза, компьютерно-томографический индекс дозы (CTDI) — в 2,6 раза, средняя поглощённая доза (DLP) — в 3,1 раза, сила тока на трубке — в 1,83 раза, напряжение на трубке — в 1,2 раза. Полученные анкетные данные свидетельствуют о том, что при проведении НДКТ эффективность выявления основных признаков вирусной пневмонии и оценки динамики состояния пациента существенно не меняется по сравнению с КТ, проведённой по стандартному протоколу.

Заключение. Результаты сравнения стандартной и НДКТ демонстрируют отсутствие значимых потерь диагностической информации и качества при снижении лучевой нагрузки. Таким образом, НДКТ грудной клетки может использоваться в рутинной практике для успешной диагностики COVID-19.

Ключевые слова: COVID-19; НДКТ; лёгкие; лучевая нагрузка; SARS-CoV-2.

Как цитировать

Филатова Д.А., Сеницын В.Е., Мершина Е.А. Возможности снижения лучевой нагрузки при проведении компьютерной томографии для оценки изменений в лёгких, характерных для COVID-19: использование адаптивной статистической итеративной реконструкции // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 94–104. DOI: <https://doi.org/10.17816/DD62477>

DOI: <https://doi.org/10.17816/DD62477>

Opportunities to reduce the radiation exposure during computed tomography to assess the changes in the lungs in patients with COVID-19: use of adaptive statistical iterative reconstruction

Daria A. Filatova, Valentin E. Sinitsin, Elena A. Merzhina

Lomonosov Moscow State University, Moscow, Russian Federation

ABSTRACT

BACKGROUND: Several COVID-19 patients are subjected to multiple imaging examinations during hospitalization, the cumulative effect of which can significantly increase the total dose of radiation received. The effective radiation dose can be reduced by lowering the current and voltage of the X-ray tube, but this reduces image quality. One possible solution is to use adaptive statistical iterative reconstruction technology on the «raw» CT data. Recently, data on the efficacy of low-dose CT (LDCT) in the diagnosis of COVID-19 have appeared in the literature.

AIM: To analyze the quality and diagnostic value of LDCT images of the lungs after applying an iterative processing algorithm and to assess the possibility of reducing the radiation load on the patient when diagnosing COVID-19.

MATERIALS AND METHODS: Patients from the Infectious Diseases Department of the Moscow State University Hospital participated in the prospective study. CT examinations were performed at the time of patient admission and discharge and were repeated as needed during hospitalization. In the first study, a standard CT protocol with a tube voltage of 120 kV and automatic current modulation in the range of 200–400 mA was used; in repeated CT scans, the LDCT protocol was used with reduced tube voltage parameters (100 or 110 kV) and automatic current modulation in the range of 40–120 mA. To assess the diagnostic value of LDCT in comparison with standard CT, a survey was conducted among doctors from the Department of Radiation Diagnostics at Moscow State University Hospital. The questionnaire included a comparison of the two methods for identifying the following pathological processes: «ground-glass» opacities, compaction of the lung tissue with reticular changes, areas of lung tissue consolidation, and lymphadenopathy.

RESULTS: The study included 151 patients. The average age was 58 ± 14.2 years, with men accounting for 53.6% of the population. During LDCT the radiation load was reduced by 2.96 times on average, CTDI by 2.6 times, DLP by 3.1 times, the current on the tube by 1.83 times, and the voltage on the tube by 1.2 times. The results indicate that the effectiveness of detecting the main signs of viral pneumonia and assessing the dynamics of the patient's condition does not differ significantly from CT performed according to the standard protocol.

CONCLUSIONS: The results of a comparison of standard and low-dose CT show that there is no significant loss of diagnostic information and image quality as the radiation load is reduced. Thus, chest LDCT can be used to successfully diagnose COVID-19 in routine practice.

Keywords: COVID-19; X-ray computed tomography; lung; radiation protection; SARS-CoV-2.

To cite this article

Filatova DA, Sinitsin VE, Merzhina EA. Opportunities to reduce the radiation exposure during computed tomography to assess the changes in the lungs in patients with COVID-19: use of adaptive statistical iterative reconstruction. *Digital Diagnostics*. 2021;2(2):94–104. DOI: <https://doi.org/10.17816/DD62477>

Received: 02.03.2021

Accepted: 20.05.2021

Published: 28.06.2021

DOI: <https://doi.org/10.17816/DD62477>

在计算机断层扫描期间减少辐射负荷以评估COVID-19肺特性变化的可能性：使用自适应统计迭代重建

Daria A. Filatova, Valentin E. Sinitsin, Elena A. Mershina

Lomonosov Moscow State University, Moscow, Russian Federation

简评

论证大多数COVID-19患者在住院期间接受多次成像检查，其累积效应可以显著增加接受的辐射总剂量。有效辐射剂量可以通过降低x射线管的电流和电压来降低，然而，这会降低图像质量。这个问题的一个可能的解决方案是引入自适应统计迭代重建（Adaptive Statistical Iterative Reconstruction (ASIR)）技术，用于计算机断层扫描（CT）的»原始数据»的自适应统计迭代重建。最近，有关低剂量CT（LDCT）有效性的信息已经出现在COVID-19诊断中的文献中。

目的是在应用迭代处理算法后分析肺部LDCT图像的质量和诊断价值，以评估在COVID-19诊断期间减少患者辐射负荷的可能性。

材料与方法。这项前瞻性研究涉及在罗蒙诺索夫莫斯科国立大学医学中心传染病部门接受住院治疗的患者。CT研究在入院和出院时进行；在住院期间，根据临床需要重复进行。在第一项研究中，使用120kV管电压和200–400mA范围内的自动电流调制的标准CT协议，通过重复CT扫描的时候，LDCT协议使用管电压（100或110kV）和40–120mA范围内的自动电流调。为了评估LDCT与标准CT相比的诊断价值，在莫斯科国立大学医学中心辐射诊断系的医生中进行了问卷调查。调查问卷包括两种方法的比较描述，用于检测这种病理过程，如通过磨砂玻璃类型压实肺组织，通过磨砂玻璃类型压实具有网状变化，肺组织固结区域，淋巴结病。

结果该研究涉及151名患者；平均年龄为 58 ± 14.2 岁；男性为53.6%。使用LDCT，与标准CT相比，辐射负荷平均下降2.96倍，计算机断层扫描剂量指数（CTDI）—2.6倍，平均吸收剂量（DLP）—3.1倍，管上的电流—1.83倍，管上的电压—1.2倍。获得的问卷数据表明，在LDCT期间，与根据标准协议进行的CT相比，检测病毒性肺炎的主要体征和评估患者病情动态的有效性没有显著变化。

结论比较标准的CT和LDCT的结果表明，在辐射负荷降低的情况下，诊断信息和质量没有显著损失。因此，胸部的LDCT扫描可以在常规实践中用于成功诊断COVID-19。

关键词：COVID-19; LDCT; 肺; 辐射负荷; SARS-CoV-2。

引用本文：

Filatova DA, Sinitsin VE, Mershina EA. 在计算机断层扫描期间减少辐射负荷以评估COVID-19肺特性变化的可能性：使用自适应统计迭代重建. *Digital Diagnostics*. 2021;2(2):94–104. DOI: <https://doi.org/10.17816/DD62477>

收到: 02.03.2021

接受: 20.05.2021

发布日期: 28.06.2021

ОБОСНОВАНИЕ

Во время пандемии COVID-19 для диагностики коронавируса пневмонии как в амбулаторных, так и стационарных условиях часто применяются компьютерно-томографические исследования (КТ), которые рекомендовано проводить у пациентов с подозрением или верифицированным COVID-19 в день госпитализации для начального обследования, затем повторно через 2–3 дня при недостижении требуемого терапевтического эффекта и через 5–7 дней при отсутствии или улучшении динамики симптомов [1–5].

Ряд пациентов с COVID-19 в ходе госпитализации проходят через многократные визуализационные исследования, кумулятивный эффект которых может значительно увеличивать суммарную дозу полученного облучения. Принцип «as low as reasonably achievable» (ALARA) гласит, что всякий раз, когда необходимо излучение, воздействие должно быть настолько низким, насколько это разумно достижимо. Имея в виду этот важный принцип, крайне важно помнить, что любая КТ должна сопровождаться обоснованием обследования и оптимизацией дозы облучения [6]. Хотя КТ является большим подспорьем в установлении диагноза COVID-19, потенциал для увеличения радиационного облучения большого числа пациентов по всей стране нельзя игнорировать. Важно балансировать между потребностью в эффективной визуализации для быстрой диагностики и попытками минимизировать радиационное облучение.

Эффективная доза облучения при выполнении КТ-исследований может быть уменьшена за счёт снижения тока и напряжения рентгеновской трубки, но это приводит к искажению качества изображений вследствие увеличения количества шума и артефактов. Возможным решением данной проблемы может служить внедрение технологии адаптивной статистической итеративной реконструкции «сырых данных» КТ, например с помощью технологии ASIR (Adaptive Statistical Iterative Reconstruction) и ряда подобных ей методик [7–9].

В последнее время в литературе появляются данные об эффективности в диагностике COVID-19 низкодозной КТ (НДКТ) в сравнении со стандартной. Стоит отметить, что низкодозной считается КТ с лучевой нагрузкой 0,2 мЗв и менее. В ретроспективном исследовании НДКТ с итеративной реконструкцией в диагностике COVID-19 продемонстрировала чувствительность, специфичность и прогностическую ценность около 90%. Если у пациентов симптоматика наблюдалась более 48 ч, значения этих параметров увеличивались до 96%. У пациентов с положительной КТ вероятность заболевания повышалась с 43,2% (до теста) до 91,1% или 91,4% (после теста); у пациентов с отрицательной КТ вероятность заболевания снизилась с 43,2% (до теста) до 9,6% или 3,7%

(после теста). Кроме того, НДКТ продемонстрировала аддитивное диагностическое преимущество у пациентов с сопутствующей бактериальной пневмонией или альтернативным диагнозом, отличным от COVID-19 [10]. Исследования в этом перспективном направлении активно продолжаются.

Цель исследования — проанализировать качество и диагностическую ценность НДКТ-изображений лёгких после применения ASIR-алгоритма обработки, оценить возможности снижения лучевой нагрузки на пациента при диагностике COVID-19.

МЕТОДЫ

Дизайн исследования

В проспективном одноцентровом неконтролируемом исследовании приняли участие пациенты, проходившие стационарное лечение в инфекционном отделении МНОЦ МГУ имени М.В. Ломоносова. Исследования КТ выполняли при поступлении и выписке пациента; в период госпитализации их повторяли по мере клинической необходимости, но не реже чем 1 раз в 5 дней. Первое исследование всем пациентам проводили в режиме стандартной КТ, последующие — в режиме НДКТ.

Первичная конечная точка исследования — отсутствие значимых потерь диагностической информации при проведении НДКТ в сравнении со стандартной КТ.

Критерии соответствия

К критериям соответствия относились верифицированное молекулярно-генетическими исследованиями (метод полимеразной цепной реакции, ПЦР) инфицирование новой коронавирусной инфекцией; прохождение стационарного лечения.

Условия проведения

Исследование выполнено в инфекционном отделении МНОЦ МГУ с участием госпитализированных пациентов с новой коронавирусной инфекцией.

Продолжительность исследования

Исследование проводилось с 21 апреля по 11 мая 2020 г.

Описание медицинского вмешательства

КТ лёгких и органов грудной клетки выполняли на 32-рядном компьютерном томографе Somatom Score производства компании Siemens (ФРГ). Исследования выполняли с толщиной среза 1 мм. При первом исследовании использовали стандартный протокол КТ с напряжением тока на трубке 120 кВ и автоматическим модулированием силы тока на трубке в диапазоне 200–400 мА; при повторных КТ применяли протокол НДКТ с уменьшенными параметрами

напряжения тока на трубке (100 или 110 кВ) и автоматической модуляцией тока на трубке в диапазоне 40–120 мА; для снижения лучевой нагрузки применяли алгоритм ASIR (Adaptive Statistical Iterative Reconstruction). Все полученные изображения в формате DICOM хранили в радиологической информационной сети (PACS/RIS) МНОЦ МГУ. Для обработки и анализа КТ использовали рабочие станции Syngo.via (Siemens, ФРГ).

Для оценки диагностической ценности НДКТ в сравнении со стандартной КТ было проведено анкетирование врачей отделения лучевой диагностики МНОЦ МГУ имени М.В. Ломоносова. Анкета включала в себя сравнительную характеристику двух методик при выявлении следующих патологических процессов: уплотнение лёгочной ткани по типу матового стекла, уплотнение по типу матового стекла с ретикулярными изменениями (утолщённые междольковые перегородки; картина «лоскутного одеяла», «булыжной мостовой»), участки консолидации лёгочной ткани, лимфаденопатия. Врачи оценивали каждую из двух методик по пятибалльной шкале, где наихудшая выявляемость того или иного патологического процесса соответствовала 1 баллу, наилучшая — 5, затем по каждому пункту вычисляли среднее арифметическое. В заключение предлагалось оценить в целом, эффективна ли диагностика COVID-19 с помощью НДКТ. Каждое исследование оценивалось двумя врачами, решение в каждом случае принималось независимо.

Основной исход исследования

Основной исход исследования: сопоставимая диагностическая ценность КТ, выполненной по стандартному протоколу, и НДКТ.

Этическая экспертиза

Тема данной статьи была утверждена на заседании ЛЭК Медицинского научно-образовательного центра (МНОЦ) МГУ имени М.В. Ломоносова от 25 мая 2020 г. (в рамках научной темы по диагностике и лечению COVID-19 в МНОЦ).

Статистический анализ

Статистический анализ производили на базе программного обеспечения MS Office Excel.

РЕЗУЛЬТАТЫ

Участники исследования

В исследовании принял участие 151 пациент, проходивший стационарное лечение в инфекционном отделении МНОЦ МГУ имени М.В. Ломоносова. Средний возраст пациентов $58 \pm 14,2$ года; женщин 70 (46,4%), мужчин 81 (53,6%). Диагноз новой коронавирусной инфекции был подтверждён результатами ПЦР.

Основные результаты исследования

Характеристики первого по счёту исследования (стандартная КТ): средняя лучевая нагрузка $3,76 \pm 1,28$ мЗв; средний компьютерно-томографический индекс дозы (Computed Tomography Dose Index, CTDI) $6,69 \pm 2,18$ мГр; средняя поглощённая доза (Dose Length Product, DLP) $222,28 \pm 76,33$ мГр/см; средняя сила тока на трубке $2165,97 \pm 682,83$ мА/с; среднее напряжение на трубке $129,43 \pm 3,21$ мВ. Характеристика последующих исследований (НДКТ): лучевая нагрузка $1,27 \pm 0,47$ мЗв; CTDI $1,57 \pm 1,40$ мГр; DLP $73,01 \pm 19,94$ мГр/см; сила тока на трубке $1182,55 \pm 366,55$ мА/с; напряжение на трубке $111,79 \pm 5,73$ мВ. Если пациенту проводили несколько исследований НДКТ, то при вычислении статистических показателей учитывали среднее арифметическое между ними.

При сравнении стандартной и низкодозной КТ были получены следующие результаты: при проведении НДКТ лучевая нагрузка снижалась в среднем в 2,96 раза, CTDI — в 2,6 раза, DLP — в 3,1 раза, сила тока на трубке — в 1,83 раза, напряжение на трубке — в 1,2 раза. Данные значения суммированы в табл. 1.

Результаты анкетирования врачей отделения лучевой диагностики МНОЦ МГУ по оценке диагностической точности НДКТ в сравнении со стандартной КТ отражены в табл. 2.

Таблица 1. Сравнительная характеристика стандартной и низкодозной компьютерной томографии

Показатель	Стандартная КТ	Низкодозная КТ	Разница, раз
Средняя лучевая нагрузка, мЗв	$3,76 \pm 1,28$	$1,27 \pm 0,47$	2,96
CTDI, мГр	$6,69 \pm 2,18$	$1,57 \pm 1,40$	2,6
DLP, мГр/см	$222,28 \pm 76,33$	$73,01 \pm 19,94$	3,1
Средняя сила тока на трубке, мА/с	$2165,97 \pm 682,83$	$1182,55 \pm 366,55$	1,83
Напряжение на трубке, мВ	$129,43 \pm 3,21$	$111,79 \pm 5,73$	1,2

Примечание. КТ — компьютерная томография; CTDI (Computed Tomography Dose Index) — средний компьютерно-томографический индекс дозы; DLP (Dose Length Product) — средняя поглощённая доза.

Таблица 2. Результаты анкетирования врачей отделения лучевой диагностики МНОЦ МГУ

Характеристика	Стандартная КТ	НДКТ
Выявление уплотнений лёгочной ткани по типу матовых стёкол	5	5
Выявление участков уплотнения по типу матового стекла с ретикулярными изменениями (утолщённые междольковые перегородки) — картина «лоскутного одеяла», «булыжной мостовой»	5	4,43
Выявление участков консолидации лёгочной ткани	5	5
Выявление лимфаденопатии	5	4

Примечание. Указаны средние значения выставленных по каждому пункту баллов: минимальный балл — 1, максимальный — 5. КТ — компьютерная томография; НДКТ — низкодозная компьютерная томография.

Как видно из табл. 2, при НДКТ эффективность выявления основных признаков вирусной пневмонии и оценки динамики состояния пациента существенно не меняется в сравнении со стандартной КТ. Следует также отметить, что, согласно результатам анкетирования, 7 врачей (100% опрошенных) считают, что НДКТ в целом эффективна для диагностики COVID-19.

Приведём наглядные примеры клинических случаев (рис. 1–6), демонстрирующих схожесть диагностической ценности двух вышеописанных методик исследования. На столбцах слева (а) расположены изображения стандартной КТ, выполненной при поступлении пациента в стационар, на столбцах справа (b) — НДКТ в динамике. Верхняя строчка изображений показывает режим лёгочного окна, нижняя — режим медиастинального окна. Для сравнения приведены значения лучевой нагрузки в каждом случае. Временные интервалы между стандартной КТ и НДКТ

составляли 2–7 дней. Таким образом, первичная конечная точка достигнута у всех пациентов, вовлечённых в исследование.

Нежелательные явления

В ходе исследования нежелательных явлений в результате КТ по стандартному протоколу и НДКТ не зафиксировано.

ОБСУЖДЕНИЕ

Резюме основного результата исследования

Результаты проведённого исследования подтверждают отсутствие значимых потерь диагностической информации при НДКТ органов грудной клетки у пациентов с COVID-19. Таким образом, НДКТ грудной клетки может использоваться в рутинной практике для успешной диагностики данного заболевания.

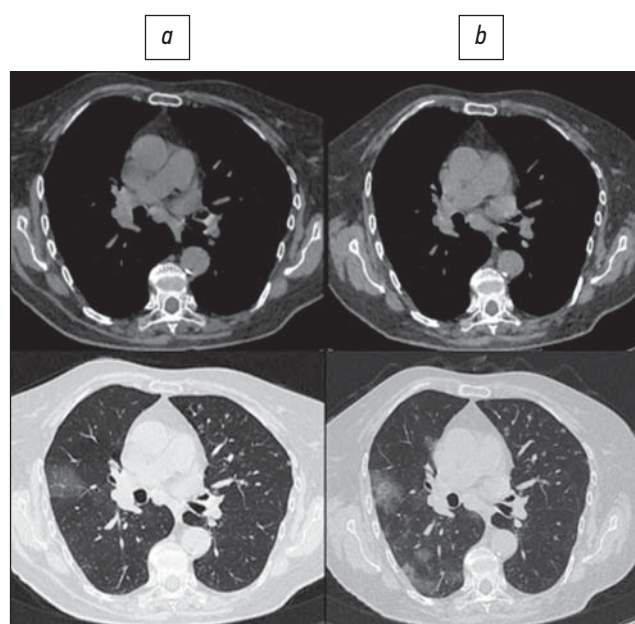


Рис. 1. Пациентка, 78 лет: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 2,5 мЗв (а), низкодозная компьютерная томография — 1,0 мЗв (b).

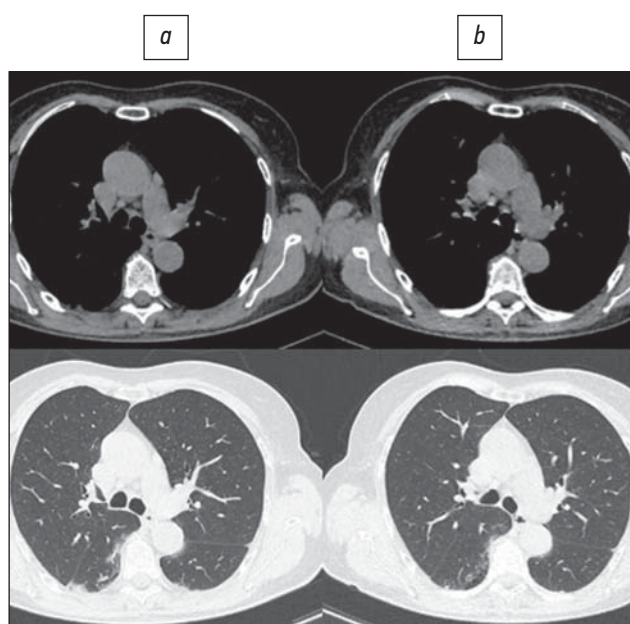


Рис. 2. Пациент, 72 года: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 2,1 мЗв (а), низкодозная компьютерная томография — 0,87 мЗв (b).

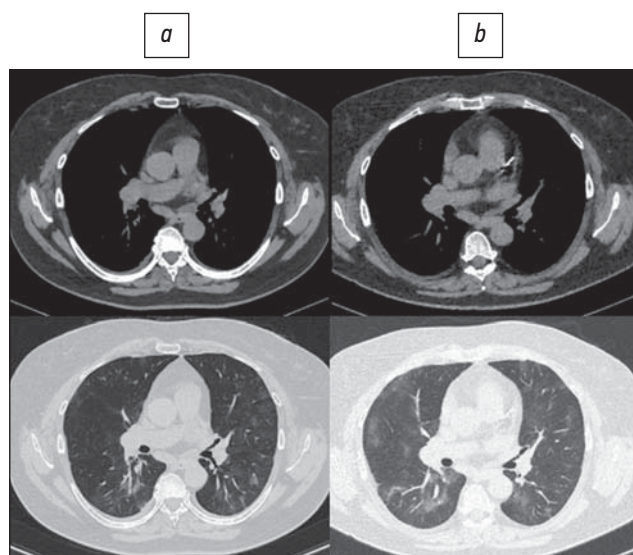


Рис. 3. Пациент, 60 лет: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 3,3 мЗв (a), низкодозная компьютерная томография — 1,1 мЗв (b).

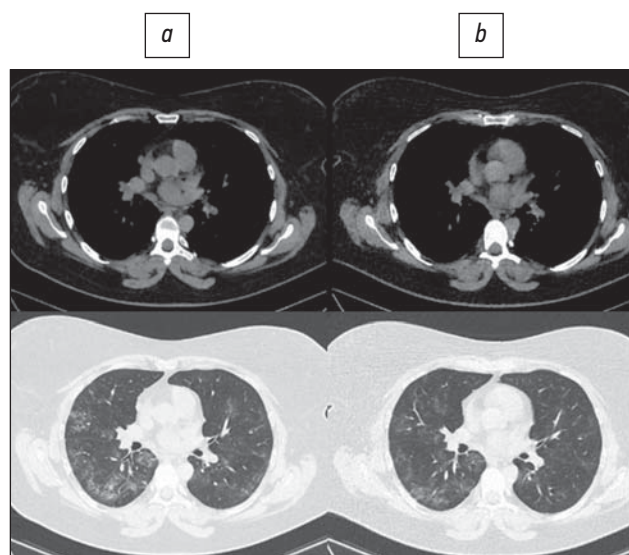


Рис. 4. Пациент, 46 лет: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 5,6 мЗв (a), низкодозная компьютерная томография — 1,7 мЗв (b).

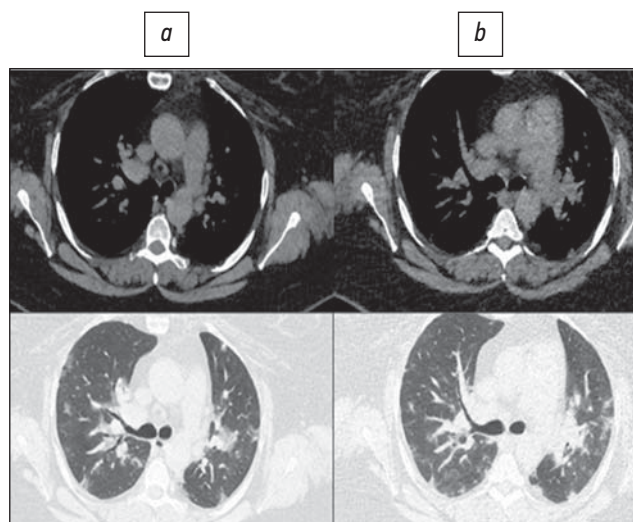


Рис. 5. Пациент, 40 лет: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 6,8 мЗв (a), низкодозная компьютерная томография — 2,0 мЗв (b).

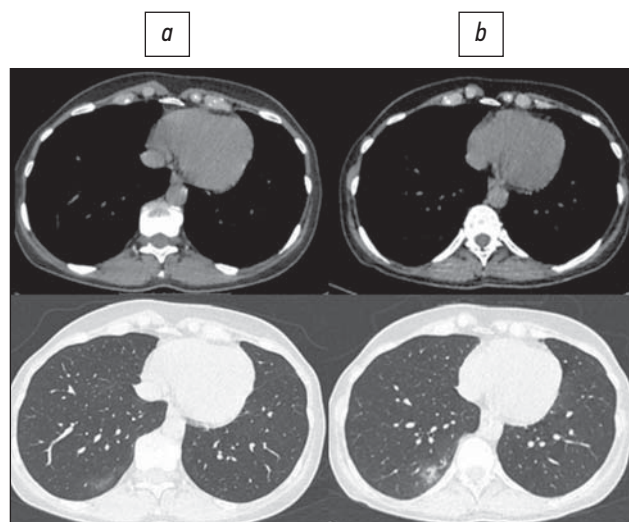


Рис. 6. Пациент, 56 лет: стандартная компьютерная томография при поступлении выполнена с лучевой нагрузкой 1,6 мЗв (a), низкодозная компьютерная томография — 0,87 мЗв (b).

Обсуждение основного результата исследования

В условиях отсутствия этиотропного лечения коронавирусной инфекции особенно важно диагностировать заболевание на ранней стадии и немедленно изолировать инфицированного человека. Согласно клиническим рекомендациям, диагноз коронавирусной инфекции устанавливается на основании клинического обследования, данных эпидемиологического анамнеза и результатов лабораторных исследований [11]. Задачей этиологической лабораторной диагностики является поиск РНК SARS-CoV-2 с помощью методов амплификации нуклеиновых кислот (ПЦР с обратной транскрипцией, ОТ-ПЦР). Обнаружение возбудителя в мазке из носоглотки

возможно уже за неделю до начала клинических проявлений инфекции [12]. Тем не менее получены данные о том, что ОТ-ПЦР может давать и ложноотрицательные результаты. Так, Ch. Long и соавт. [13] сообщили, что среди 36 пациентов с диагностированной коронавирусной инфекцией у 35 пациентов на КТ были признаки характерной пневмонии, в то время как положительный результат ОТ-ПЦР в первый раз был получен только у 30 пациентов. Оставшиеся 6 случаев были подвергнуты повторному тестированию: у 3 из них положительный результат был получен при втором тестировании (через 2 дня), ещё у 3 — при третьем (через 6 дней). Таким образом, чувствительность КТ составила 97,2%, ОТ-ПЦР при первом исследовании — 84,6% [13]. В исследовании Y. Fang и соавт. [14] получили схожие результаты:

чувствительность КТ оказалась равна 98%, ОТ-ПЦР — 71% (при первом исследовании положительный результат был получен у 36 пациентов из 51 с симптомами пневмонии на КТ и подходящим эпидемиологическим анамнезом, при втором исследовании диагноз был подтверждён ещё у 12 пациентов, при третьем — у 2, при четвёртом — у 1). Среди причин того, что чувствительность ОТ-ПЦР в диагностике коронавирусной инфекции оказалась ниже, чем КТ, можно предположить следующие: несовершенство технологий амплификации нуклеиновых кислот; вариабельность порога чувствительности тестов разных производителей; низкая вирусная нагрузка; неправильная техника взятия материала для анализа. Кроме того, количество вирусных частиц варьирует в зависимости от того, откуда взят материал для анализа: имеются данные, что предпочтительнее всего первично исследовать мокроту, следом по чувствительности стоит мазок из носоглотки [15]. Таким образом, даже несмотря на отрицательный результат ОТ-ПЦР, при наличии у пациента характерных симптомов и эпидемиологического анамнеза рекомендуется КТ для визуализации изменений в лёгких. При обнаружении на КТ признаков пневмонии необходимо принять меры к экстренной изоляции пациента, после чего следует провести повторный лабораторный анализ.

В условиях возросшего количества КТ-исследований встаёт вопрос о значительном повышении лучевой нагрузки и ассоциированным с ней риском: так, имеются данные, что порядка 2% онкологических заболеваний в США связаны с радиационными дозами, полученными в результате КТ [16]. Несмотря на отсутствие больших эпидемиологических исследований по этой тематике, имеется большое количество данных относительно радиационно-индуцированного рака у выживших после сброшенных на Японию в 1945 г. атомных бомб. В подгруппе людей, получивших дозы облучения в диапазоне от 5 до 150 мЗв, наблюдалось значительное увеличение общего риска развития рака, средняя доза в этой подгруппе составляла 40 мЗв [17], а средняя эффективная доза облучения при стандартной КТ органов грудной клетки составляет 5 мЗв [18]. Относительно же НДКТ грудной клетки с лучевой нагрузкой порядка 0,4 мЗв в настоящее время не получено достаточных доказательств эффективности в контексте скрининга и диагностики коронавирусной инфекции [3].

Доза облучения, получаемая пациентом по время КТ-исследования, зависит от силы тока на трубке, напряжения, времени сканирования, толщины среза, объёма сканирования и шага. Хотя при использовании современных моделей спиральных томографов время сканирования сокращается, радиационное облучение порой даже возрастает из-за увеличения силы тока и объёма сканирования. В данных условиях разумно прибегнуть к методикам снижения дозы облучения. Известно, что доза прямо пропорциональна току на трубке;

в последнее время было проведено несколько исследований, показавших, что выполнение НДКТ грудной клетки при 10–140 мАс не даёт серьёзного снижения качества изображений, и на них всё ещё обнаруживаются узелковые структуры [19–21]. В исследовании X. Zhu и соавт. [22] показано, что имеется линейная корреляция между током на трубке и поглощённой дозой при постоянном напряжении и времени сканирования, а также оценена целесообразность оптимизации дозы облучения за счёт уменьшения тока на трубке. Путём сравнения изображений, получаемых при разных значениях CTDI, было определено пороговое значение данного параметра, позволяющее получать изображения без существенной потери информативности (25 мАс), причём с увеличением толщины срезов потеря качества изображения происходила медленнее. Статистический анализ не выявил существенной разницы между изображениями, полученными при 115; 40 и 25 мАс. Таким образом, 25 мАс и более — это приемлемый параметр экспозиции для обеспечения удовлетворительного качества изображения для КТ грудной клетки; при этом CTDI при 25 мАс снижается на 70% по сравнению с CTDI при 115 мАс. Несмотря на точность данного параметра, клиницистам стоит учитывать, что его значение может варьировать при использовании различных систем КТ; кроме того, необходимо корректировать его с учётом биологических особенностей пациентов (например, доза облучения должна быть увеличена для пациентов с ожирением и при исследовании верхней доли лёгкого из-за ложной тени, вызываемой лопаткой). Полученные в данном исследовании пороговые значения параметра CTDI согласуются с результатами исследования T. Kubo и соавт. [23], в котором стандартная и низкодозная КТ сравнивались при определении основных характеристик образований в лёгких, позволяющих подтвердить или исключить злокачественную природу. Параметров 20–50 мАс оказалось достаточно для определения природы образования без проведения дополнительной стандартной КТ. При определении таких параметров образований, как структура, характеристика краёв, кальцификация, дольчатость, а также реакция плевры, стандартная и низкодозная КТ показали одинаковую эффективность.

В течение многих лет дискуссионным оставался вопрос, безопасно ли применять НДКТ для скрининга онкологических заболеваний, например, рака лёгких [24]. В исследовании C. Rampinelli и соавт. [25] были проанализированы возможные риски развития радиационных рака лёгких и лейкемии у здоровых людей, которые в течение 10 лет регулярно проходили скрининг с использованием НДКТ. Оказалось, что суммарная кумулятивная доза радиации составила примерно 9 мЗв для мужчин и 13 мЗв для женщин, что эквивалентно одному стандартному исследованию КТ. Кроме того, учитывая, что средняя доза от фоновых источников в США

составляет примерно 30 мЗв за 10 лет, можно заключить, что на долю скрининга с помощью НДКТ приходится лишь 1/3 от воздействия естественного радиационного фона за аналогичный период времени. Результаты исследования показали, что после 10 лет скрининга с помощью НДКТ у 5203 бессимптомных пациентов в возрасте старше 50 лет, имеющих стаж курения более 20 пачко-лет, примерно 1,5 случая рака лёгких и 2,4 случая других видов рака были вызваны воздействием радиации. По сравнению с количеством выявленных случаев рака лёгких можно подсчитать, что на каждый радиационно-индуцируемый рак приходится порядка 100 выявленных посредством скрининга раков. Кроме того, результаты исследования скрининга НДКТ в популяции курильщиков в возрасте 55–74 лет показали снижение смертности от рака лёгких на 20% [26]. Все эти данные свидетельствуют о том, что, несмотря на возможные риски, связанные с воздействием радиационного излучения, метод НДКТ в целом безопасен и эффективен для многократного повторения в рамках скрининга или отслеживания динамики состояния пациента в стационаре. Не вызывает сомнений важность применения НДКТ для снижения лучевой нагрузки и обеспечения большей безопасности исследования для пациента.

ЗАКЛЮЧЕНИЕ

При сравнительном анализе эффективности и диагностической ценности НДКТ и КТ, выполняемой по стандартному протоколу, выявлено, что НДКТ является не только полноценной альтернативой, но также и более предпочтительным вариантом, т.к. её проведение позволяет существенно снизить лучевую нагрузку на пациента. С учётом того, что при стационарном лечении COVID-19 пациент проходит через несколько визуализирующих исследований, вопрос о радиационной безопасности встаёт остро. По мнению практикующих врачей,

объём информации, предоставляемый НДКТ, не уступает по качеству и точности стандартной методике КТ, поэтому для исследований в динамике целесообразно предпочесть НДКТ — метод, позволяющий снизить лучевую нагрузку.

ДОПОЛНИТЕЛЬНО

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

Вклад авторов. Д.А. Филатова — поиск публикаций по теме статьи, написание текста рукописи; В.Е. Синицын — концепция исследования, экспертная оценка информации, редактирование текста рукописи, утверждение финальной версии; Е.А. Мершина — формирование набора данных, экспертная оценка информации, редактирование текста рукописи, утверждение финальной версии. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Competing interests. The authors declare that they have no competing interests.

Funding source. This study was not supported by any external sources of funding.

Authors' contribution. D.A. Filatova — search for publications on the article topic, writing the text of the manuscript; V.E. Sinitsin — the concept of research, expert evaluation of information, editing the text of the manuscript, final version approval; E.A. Merzhina — formation of a data set, expert evaluation of information, editing of the text of the manuscript, final version approval. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

СПИСОК ЛИТЕРАТУРЫ

1. Министерство здравоохранения Российской Федерации. Временные методические рекомендации: профилактика, диагностика и лечение новой коронавирусной инфекции. Версия 8 (03.09.2020). Москва, 2020. Режим доступа: <https://base.garant.ru/74596434/>. Дата обращения: 14.03.2021.
2. Романов Б.К. Коронавирусная инфекция COVID-2019 // Безопасность и риск фармакотерапии. 2020. Т. 8, № 1. С. 3–8. doi: 10.30895/2312-7821-2020-8-1-3-8
3. Морозов С.П., Проценко Д.Н., Сметанина С.В., и др. Лучевая диагностика коронавирусной болезни (COVID-19): организация, методология, интерпретация результатов: препринт № ЦДТ – Версия 2 от 17.04.2020. Москва; 2020. 78 с.
4. Udagama B., Kadhiresan P., Kozłowski H.N., et al. Diagnosing COVID-19: The disease and tools for detection // ACS Nano. 2020. Vol. 14, N 4. P. 3822–3835. doi: 10.1021/acsnano.0c02624
5. Zhao W., Zhong Z., Xie X., et al. Relation between chest ct findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study // AJR Am J Roentgenol. 2020. Vol. 214, N 5. P. 1072–1077. doi: 10.2214/AJR.20.22976
6. Beregi J.P., Greffier J. Low and ultra-low dose radiation in CT: Opportunities and limitations // Diagn Interv Imaging. 2019. Vol. 100, N 2. P. 63–64. doi: 10.1016/j.diii.2019.01.007
7. Cheng L., Fang T., Tyan J. Fast Iterative adaptive reconstruction in low-dose CT imaging // 2006 International Conference on Image Processing. Atlanta: GA: IEEE; 2006. P. 889–892. Режим доступа: <https://ieeexplore.ieee.org/document/4106673/>. Дата обращения: 14.03.2021.
8. Hara A.K., Paden R.G., Silva A.C., et al. Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study // AJR Am J Roentgenol. 2009. Vol. 193, N 3. P. 764–771. doi: 10.2214/AJR.09.2397

9. Prakash P., Kalra M., Kambadakone A., et al. Reducing abdominal CT radiation dose with adaptive statistical iterative reconstruction technique // *Invest Radiol*. 2010. Vol. 45, N 4. P. 202–210. doi: 10.1097/RLI.0b013e3181dzfeec
10. Chen L.G., Wu P.A., Sheu M.H., et al. Automatic current selection with iterative reconstruction reduces effective dose to less than 1 mSv in low-dose chest computed tomography in persons with normal BMI // *Medicine (Baltimore)*. 2019. Vol. 98, N 28. P. e16350. doi: 10.1097/MD.00000000000016350
11. Dangis A., Gieraerts C., De Brueker Y., et al. Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19 // *Radiology Cardiothoracic Imaging*. 2020. Vol. 2, N 2. P. e200196. doi: 10.1148/ryct.2020200196
12. Sethuraman N., Jeremiah S.S., Ryo A. Interpreting diagnostic tests for SARS-CoV-2 // *JAMA*. 2020. Vol. 323, N 22. P. 2249–2251. doi: 10.1001/jama.2020.8259
13. Long C., Xu H., Shen Q., et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? // *Eur J Radiology*. 2020. Vol. 126. P. 108961. doi: 10.1016/j.ejrad.2020.108961
14. Fang Y., Zhang H., Xie J., et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR // *Radiology*. 2020. Vol. 296, N 2. P. E115–E117. doi: 10.1148/radiol.2020200432
15. Yang Y., Yang M., Shen C., et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections // *medRxiv*. 2020. doi: 10.1101/2020.02.11.20021493
16. Brenner D.J., Hall E.J. Computed tomography — an increasing source of radiation exposure // *N Engl J Med*. 2007. Vol. 357, N 22. P. 2277–2284. doi: 10.1056/NEJMra072149
17. Pierce D.A., Preston D.L. Radiation-related cancer risks at low doses among atomic bomb survivors // *Radiat Res*. 2000. Vol. 154, N 2. P. 178–186. doi: 10.1667/0033-7587(2000)154[0178:rrcral]2.0.co;2
18. Маткевич Е.И., Синицын В.Е., Мершина Е.А. Сравнительный анализ доз облучения пациентов при компьютерной томографии в федеральном лечебном учреждении // *Вестник рентгенологии и радиологии*. 2016. Т. 97, № 1. С. 33–39. doi: 10.20862/0042-4676-2016-97-1-33-40
19. Naidich D.P., Marshall C., Gribbin C., et al. Low-dose CT of the lungs: preliminary observations // *Radiology*. 1990. Vol. 175, N 3. P. 729–731. doi: 10.1148/radiology.175.3.2343122
20. Prasad S.R., Wittram C., Sherard J.A., et al. Standard-dose and 50%-reduced-dose chest CT: comparing the effect on image quality // *AJR Am J Roentgenol*. 2002. Vol. 179, N 2. P. 461–465. doi: 10.2214/ajr.179.2.1790461
21. Zwirewich C.V., Mayo J.R., Müller N.L. Low-dose high-resolution CT of lung parenchyma // *Radiology*. 1991. Vol. 180, N 2. P. 413–417. doi: 10.1148/radiology.180.2.2068303
22. Zhu X., Yu J., Huang Z. Low-dose chest CT: optimizing radiation protection for patients // *AJR Am J Roentgenol*. 2004. Vol. 183, N 3. P. 809–816. doi: 10.2214/ajr.183.3.1830809
23. Kubo T., Ohno Y., Takenaka D., et al. Standard-dose vs. low-dose CT protocols in the evaluation of localized lung lesions: Capability for lesion characterization — iLEAD study // *Eur J Radiol Open*. 2016. Vol. 3. P. 67–73. doi: 10.1016/j.ejro.2016.03.002
24. Гомболевский В.А., Чернина В.Ю., Блохин И.А. Основные достижения низкодозной компьютерной томографии в скрининге рака легкого // *Туберкулез и болезни легких*. 2021. Т. 99, № 1. С. 61–70. doi: 10.21292/2075-1230-2021-99-1-61-70
25. Rampinelli C., De Marco P., Origgi D., et al. Exposure to low dose computed tomography for lung cancer screening and risk of cancer: secondary analysis of trial data and risk-benefit analysis // *BMJ*. 2017. Vol. 356. P. j347. doi: 10.1136/bmj.j347
26. Chiles C. Lung cancer screening with low dose CT // *Radiol Clin North Am*. 2014. Vol. 52, N 1. P. 27–46. doi: 10.1016/j.rcl.2013.08.006

REFERENCES

1. Ministry of Health of the Russian Federation. Temporary guidelines: prevention, diagnosis and treatment of new coronavirus infection. Version 8 (03.09.2020). Moscow; 2020. (In Russ). Available from: <https://base.garant.ru/74596434/>
2. Romanov B.K. Coronavirus infection COVID-19. *Safety and Risk of Pharmacotherapy*. 2020;8(1):3–8. (In Russ.)
3. Morozov S.P., Protsenko D.N., Smetanina S.V., et al. Radiation diagnostics of coronavirus disease (COVID-19): organization, methodology, interpretation of results: preprint No. CDT-Version 2 of 17.04.2020. Moscow; 2020. 78 p. (In Russ.)
4. Udugama B., Kadhiresan P., Kozlowski H.N., et al. Diagnosing COVID-19: the disease and tools for detection. *ACS Nano*. 2020;14(4):3822–3835. doi: 10.1021/acsnano.0c02624
5. Zhao W., Zhong Z., Xie X., et al. Relation between chest ct findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. *AJR Am J Roentgenol*. 2020;214(5):1072–1077. doi: 10.2214/AJR.20.22976
6. Beregi J.P., Greffier J. Low and ultra-low dose radiation in CT: Opportunities and limitations. *Diagn Interv Imaging*. 2019;100(2):63–64. doi: 10.1016/j.diii.2019.01.007
7. Cheng L., Chen Y., Fang T., et al. Fast iterative adaptive reconstruction in low-dose CT imaging. In: 2006 International Conference on Image Processing. Atlanta, GA: IEEE; 2006. P. 889–892. Available from: <https://ieeexplore.ieee.org/document/4106673/>
8. Hara A.K., Paden R.G., Silva A.C., et al. Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study. *AJR Am J Roentgenol*. 2009;193(3):764–771. doi: 10.2214/AJR.09.2397
9. Prakash P., Kalra M., Kambadakone A., et al. Reducing abdominal CT radiation dose with adaptive statistical iterative reconstruction technique. *Invest Radiol*. 2010;45(4):202–210. doi: 10.1097/RLI.0b013e3181dzfeec
10. Chen L.G., Wu P.A., Sheu M.H., et al. Automatic current selection with iterative reconstruction reduces effective dose to less than 1 mSv in low-dose chest computed tomography in persons with normal BMI. *Medicine (Baltimore)*. 2019;98(28):e16350. doi: 10.1097/MD.00000000000016350
11. Dangis A., Gieraerts C., De Brueker Y., et al. Accuracy and reproducibility of low-dose submillisievert chest CT for the diagnosis of COVID-19. *Radiology Cardiothoracic Imaging*. 2020;2(2):e200196. doi: 10.1148/ryct.2020200196
12. Sethuraman N., Jeremiah S.S., Ryo A. Interpreting diagnostic tests for SARS-CoV-2. *JAMA*. 2020;323(22):2249–2251. doi: 10.1001/jama.2020.8259

13. Long C, Xu H, Shen Q, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol*. 2020;126:108961. doi: 10.1016/j.ejrad.2020.108961
14. Fang Y, Zhang H, Xie J, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*. 2020;296(2):E115–E117. doi: 10.1148/radiol.2020200432
15. Yang Y, Yang M, Shen C, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv*. 2020. doi: 10.1101/2020.02.11.20021493
16. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277–2284. doi: 10.1056/NEJMr072149
17. Pierce DA, Preston DL. Radiation-related cancer risks at low doses among atomic bomb survivors. *Radiat Res*. 2000;154(2):178–186. doi: 10.1667/0033-7587(2000)154[0178:rrcral]2.0.co;2
18. Matkevich E, Sinitsyn V, Mershina E. Comparative analysis of radiation doses of patients with computed tomography in a Federal medical institution. *Journal of Radiology and Nuclear Medicine*. 2016;97(1):33–39. (In Russ). doi: 10.20862/0042-4676-2016-97-1-33-40
19. Naidich DP, Marshall CH, Gribbin C, et al. Low-dose CT of the lungs: preliminary observations. *Radiology*. 1990;175(3):729–731. doi: 10.1148/radiology.175.3.2343122
20. Prasad SR, Wittram C, Shepard JA, et al. Standard-dose and 50%-reduced-dose chest CT: comparing the effect on image quality. *AJR Am J Roentgenol*. 2002;179(2):461–465. doi: 10.2214/ajr.179.2.1790461
21. Zwirewich CV, Mayo JR, Müller NL. Low-dose high-resolution CT of lung parenchyma. *Radiology*. 1991;180(2):413–417. doi: 10.1148/radiology.180.2.2068303
22. Zhu X, Yu J, Huang Z. Low-dose chest CT: optimizing radiation protection for patients. *AJR Am J Roentgenol*. 2004;183(3):809–816. doi: 10.2214/ajr.183.3.1830809
23. Kubo T, Ohno Y, Takenaka D, et al. Standard-dose vs. low-dose CT protocols in the evaluation of localized lung lesions: Capability for lesion characterization – iLEAD study. *Eur J Radiol Open*. 2016;3:67–73. doi: 10.1016/j.ejro.2016.03.002
24. Gombolevsky VA, Chernina VY, Blokhin IA. Main achievements of low-dose computed tomography in lung cancer screening. *Tuberculosis and Lung Diseases*. 2021;99(1):61–70. (In Russ). doi: 10.21292/2075-1230-2021-99-1-61-7025
25. Rampinelli C, De Marco P, Origgi D, et al. Exposure to low dose computed tomography for lung cancer screening and risk of cancer: secondary analysis of trial data and risk-benefit analysis. *BMJ*. 2017;356:j347. doi: 10.1136/bmj.j347
26. Chiles C. Lung cancer screening with low dose CT. *Radiol Clin North Am*. 2014;52(1):27–46. doi: 10.1016/j.rcl.2013.08.006

ОБ АВТОРАХ

* Филатова Дарья Андреевна;

адрес: Россия, 143430, Московская обл., п.г.т. Нахабино, ул. Школьная, д. 1а; ORCID: <https://orcid.org/0000-0002-0894-1994>; eLibrary SPIN: 2665-5973; e-mail: dariafilatova.msu@mail.ru

Синицын Валентин Евгеньевич, д.м.н., профессор;

ORCID: <https://orcid.org/0000-0002-5649-2193>; eLibrary SPIN: 8449-6590; e-mail: vsini@mail.ru

Мершина Елена Александровна, к.м.н., доцент;

ORCID: <https://orcid.org/0000-0002-1266-4926>; eLibrary SPIN: 6897-9641; e-mail: elena_mershina@mail.ru

AUTHORS' INFO

* Daria A. Filatova;

address: 1a Shkolnaya street, 143430 Nakhabino, Moscow region, Russia; ORCID: <https://orcid.org/0000-0002-0894-1994>; eLibrary SPIN: 2665-5973; e-mail: dariafilatova.msu@mail.ru

Valentin E. Sinitsyn, MD, Dr. Sci. (Med.), Professor;

ORCID: <https://orcid.org/0000-0002-5649-2193>; eLibrary SPIN: 8449-6590; e-mail: vsini@mail.ru

Elena A. Mershina, MD, Cand. Sci. (Med.), Associate Professor;

ORCID: <https://orcid.org/0000-0002-1266-4926>; eLibrary SPIN: 6897-9641; e-mail: elena_mershina@mail.ru

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD60622>

Вариабельность заключений при интерпретации КТ-снимков: один за всех и все за одного

Н.С. Кульберг^{1, 2}, Р.В. Решетников^{1, 3}, В.П. Новик¹, А.Б. Елизаров¹, М.А. Гусев^{1, 4},
В.А. Гомболевский¹, А.В. Владзимирский¹, С.П. Морозов¹

¹ Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения г. Москвы, Москва, Российская Федерация

² Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Российская Федерация

³ Первый Московский государственный медицинский университет имени И.М. Сеченова (Сеченовский Университет), Москва, Российская Федерация

⁴ Московский политехнический университет, Москва, Российская Федерация

АННОТАЦИЯ

Обоснование. Разметка наборов медицинских изображений во многом полагается на субъективную интерпретацию наблюдаемых подозрительных структур. На настоящий момент не существует рекомендованного протокола по определению эталонных данных (ground truth), основанных на врачебных описаниях.

Цель — анализ правильности и согласованности оценок рентгенологов, принимавших участие в подготовке общедоступного набора данных CT LungCa-500; определение взаимосвязи этих показателей с количеством специалистов, проводящих независимую интерпретацию изображений, полученных при компьютерно-томографическом (КТ) исследовании.

Материал и методы. Набор данных, в разметке которого принимали участие 34 рентгенолога, включает 536 КТ-исследований пациентов из группы риска развития рака лёгкого. Каждое КТ-исследование было независимо интерпретировано шестью специалистами, после чего обнаруженные ими подозрительные структуры проходили арбитраж другим экспертом. Для каждого эксперта подсчитывали количество истинно положительных, ложноположительных, истинно отрицательных и ложноотрицательных находок, на основании которых проводили оценку диагностической точности рентгенологов. Для анализа согласованности между заключениями рентгенологов использовали метрику процентного показателя.

Результаты. Увеличение количества специалистов, проводящих независимую интерпретацию КТ-исследований, ведёт к росту правильности их оценок при снижении согласованности. Среди факторов, влияющих на согласованность заключений между парами исследователей, выделяется расхождение мнений по поводу наличия лёгочного очага в конкретном участке КТ-снимка.

Заключение. Увеличение числа независимых первичных интерпретаций способно повысить их комбинированную правильность при условии проведения арбитража, причём квалификация рентгенологов не имеет определяющего значения для качества анализа. Проведение первичной разметки силами четырёх рентгенологов является оптимальным с точки зрения сочетания правильности интерпретации и её стоимости.

Ключевые слова: компьютерная томография; набор данных; эталонные данные; согласованность между заключениями.

Как цитировать

Кульберг Н.С., Решетников Р.В., Новик В.П., Елизаров А.Б., Гусев М.А., Гомболевский В.А., Владзимирский А.В., Морозов С.П. Вариабельность заключений при интерпретации КТ-снимков: один за всех и все за одного // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 105–118. DOI: <https://doi.org/10.17816/DD60622>

DOI: <https://doi.org/10.17816/DD60622>

Inter-observer variability between readers of CT images: all for one and one for all

Nikolas S. Kulberg^{1, 2}, Roman V. Reshetnikov^{1, 3}, Vladimir P. Novik¹, Alexey B. Elizarov¹, Maxim A. Gusev^{1, 4}, Victor A. Gombolevskiy¹, Anton V. Vladzimirskyy¹, Sergey P. Morozov¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russian Federation

³ Institute of Molecular Medicine, The First Sechenov Moscow State Medical University, Moscow, Russian Federation

⁴ Moscow Polytechnic University, Moscow, Russian Federation

ABSTRACT

BACKGROUND: The markup of medical image datasets is based on the subjective interpretation of the observed entities by radiologists. There is currently no widely accepted protocol for determining ground truth based on radiologists' reports.

AIM: To assess the accuracy of radiologist interpretations and their agreement for the publicly available dataset "CTLungCa-500", as well as the relationship between these parameters and the number of independent readers of CT scans.

MATERIALS AND METHODS: Thirty-four radiologists took part in the dataset markup. The dataset included 536 patients who were at high risk of developing lung cancer. For each scan, six radiologists worked independently to create a report. After that, an arbitrator reviewed the lesions discovered by them. The number of true-positive, false-positive, true-negative, and false-negative findings was calculated for each reader to assess diagnostic accuracy. Further, the inter-observer variability was analyzed using the percentage agreement metric.

RESULTS: An increase in the number of independent readers providing CT scan interpretations leads to accuracy increase associated with a decrease in agreement. The majority of disagreements were associated with the presence of a lung nodule in a specific site of the CT scan.

CONCLUSION: If arbitration is provided, an increase in the number of independent initial readers can improve their combined accuracy. The experience and diagnostic accuracy of individual readers have no bearing on the quality of a crowd-tagging annotation. At four independent readings per CT scan, the optimal balance of markup accuracy and cost was achieved.

Keywords: X-ray computed tomography; datasets as topic; ground truth; observer variation.

To cite this article

Kulberg NS, Reshetnikov RV, Novik VP, Elizarov AB, Gusev MA, Gombolevskiy VA, Vladzimirskyy AV, Morozov SP. Inter-observer variability between readers of CT images: all for one and one for all. *Digital Diagnostics*. 2021;2(2):105–118. DOI: <https://doi.org/10.17816/DD60622>

Received: 11.02.2021

Accepted: 07.07.2021

Published: 13.07.2021

DOI: <https://doi.org/10.17816/DD60622>

CT图像解释中结论的可变性： 一个为所有和所有为一

Nikolas S. Kulberg^{1,2}, Roman V. Reshetnikov^{1,3}, Vladimir P. Novik¹, Alexey B. Elizarov¹, Maxim A. Gusev^{1,4}, Victor A. Gomboleviskiy¹, Anton V. Vladzemyrskyy¹, Sergey P. Morozov¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russian Federation

³ Institute of Molecular Medicine, The First Sechenov Moscow State Medical University, Moscow, Russian Federation

⁴ Moscow Polytechnic University, Moscow, Russian Federation

结构简评

理由： 医学图像集的标记在很大程度上依赖于观察到的可疑结构的主观解释。目前，没有推荐的协议用于根据医学描述确定参考数据（ground truth）。

目标： 评估参与编制公开数据集»CTLungCa-500«的放射科医生评估的正确性和一致性，以及确定这些指标与对CT研究进行独立解释的专家数量的关系。

方法： 该数据集包括有患肺癌风险的患者的536项CT研究，其中34名放射科医生参加了该研究。每项CT研究都由六位专家独立解释，之后他们发现的可疑结构由另一位专家进行仲裁。对于每位专家计算真阳性，假阳性，真阴性和假阴性结果的数量，在此基础上评估放射科医生的诊断准确性。为了分析放射科医生的结论之间的一致性，使用了百分比度量。

结果： 对CT研究进行独立解释的专家数量的增加在一致性降低的情况下导致其评估的正确性增加。在影响成对研究人员之间结论一致性的因素中，关于CT图像的特定部分中存在肺焦点的观点不一致。

结论： 独立的初级解释数量的增加使它们的组合正确性会升高，但需要仲裁，放射科医生的资格对分析的质量没有决定性的价值。从结合解释的正确性及其成本的角度来看，由四名放射科医生进行主要标记是最佳的。

关键词： 计算机断层扫描，数据集，参考数据，结论之间的一致性。

引用本文：

Kulberg NS, Reshetnikov RV, Novik VP, Elizarov AB, Gusev MA, Gomboleviskiy VA, Vladzemyrskyy AV, Morozov SP. CT图像解释中结论的可变性：一个为所有和所有为一. *Digital Diagnostics*. 2021;2(2):105–118. DOI: <https://doi.org/10.17816/DD60622>

收到：11.02.2021

接受：07.07.2021

发布日期：13.07.2021

ОБОСНОВАНИЕ

В 2017 г. С.П. Морозов и соавт. подготовили общедоступный набор данных «Тэгированные результаты компьютерных томографий лёгких», впоследствии получивший название «CTLung500-Ca» [1, 2]. Этот набор содержит 536 рентгеновских изображений органов грудной клетки пациентов из группы риска развития рака лёгкого, полученных методом компьютерно-томографического (КТ) исследования. Интерпретацию каждого исследования независимо проводили шестеро рентгенологов, находки впоследствии проходили проверку дополнительным экспертом. При разметке использовали подход со слабой аннотацией находок, т.е. указание ограниченного числа очагов на КТ-снимке, локализацию которых осуществляли заданием координат охватывающих сфер максимального диаметра с последующей их кластеризацией [2, 3]. С.П. Морозов и соавт. разработали такой протокол разметки и аннотации, поскольку интерпретации рентгенологов имеют склонность к субъективности и не защищены от ошибок. В условиях когда стоимость ложноположительных и ложноотрицательных находок одинаково высока, арбитраж первичных интерпретаций может увеличить правильность заключений [4]. Отметим, что такой арбитраж эффективен только в случае, если рентгенологи совершают разные ошибки. Согласно P.G. Herman и S.J. Hessel, вероятность того, что одну и ту же ложноположительную находку сделают два и более рентгенологов, невелика. Однако существенную долю ложноотрицательных ошибок, как правило, совершают два и более специалистов [5]. Таким образом, количество рентгенологов, проводящих независимую интерпретацию КТ-снимков, может существенно повлиять на точность разметки и аннотации.

Цель исследования. Первичной целью настоящего исследования являлось изучение взаимосвязи между числом независимых интерпретаций расположенных в базе данных CTLungCa-500 КТ-снимков и количеством и типом совершаемых ошибок, а также поиск протокола интерпретации КТ-исследований, способствующего оптимальной точности разметки. Вторичной целью исследования было изучение согласованности заключений рентгенологов, принимавших участие в подготовке набора данных.

МЕТОДЫ

Дизайн исследования

В настоящей работе анализировали данные ретроспективного многоцентрового обсервационного исследования, посвящённого изучению перспектив применения технологий компьютерного зрения в системе здравоохранения г. Москвы.

Критерии соответствия

Критерии включения: пациенты поликлиник г. Москвы в возрасте от 50 до 75 лет, которые проходили диагностическое КТ-исследование по направлению лечащего врача в связи с подозрением на рак лёгкого.

Условия проведения

Согласно критериям включения, из Единого радиологического информационного сервиса было выгружено 3897 КТ-исследований. Из этого количества случайным образом отобрали 550 КТ-исследований для создания набора данных «Тэгированные результаты компьютерных томографий лёгких». Из выборки исключили 14 КТ-снимков по причине несоответствия критериям включения либо протоколу медицинского вмешательства.

Продолжительность исследования

Набор данных включает результаты КТ-исследований, проводимых с 01 января 2015 г. по 31 декабря 2017 г.

Описание медицинского вмешательства

Рекомендуемые параметры сканирования для взрослых пациентов (рост 170 см, масса тела 70 кг): автоматическая модуляция тока на трубке при напряжении 120 кВ, FOV 350 мм, толщина среза $\leq 1,5$ мм, расстояние между соседними срезами $\leq$ толщины среза. Сканирование проводили в положении пациента лежа на спине, направление сканирования — от диафрагмы до верхушек лёгких за одну задержку дыхания на вдохе. Ядра реконструкции были специфичны для конкретного производителя томографа: для аппаратов Toshiba — FC50, FC51, FC52, FC53, FC07 для лёгких и FC07, FC08, FC09, FC17, FC18 для мягких тканей; для аппаратов Siemens — B70, B75 и B80; для аппаратов Philips — Y-Sharp и LUNG для лёгких и SOFT для мягких тканей; для аппаратов GE (General Electric) — LUNG для лёгких и SOFT для мягких тканей.

Основной исход исследования

В разметке и аннотировании исследований принимали участие две группы рентгенологов-волонтёров. Представители первой группы (первичные эксперты), состоящей из 15 специалистов с опытом от 2 до 10+ лет, выполняли первичную интерпретацию КТ-снимков. Согласно сформированной методологии, врачи искали на изображениях КТ лёгочные очаги размерами от 4 до 30 мм, сохраняя такую информацию о находках, как локализацию лёгочного очага (положение центра находки по двум измерениям на изображении и номеру среза); диаметр находки; тип лёгочного очага (солидный, полусолидный или очаг по типу матового стекла). Врачам рекомендовали не отмечать кальцинированные и перифиссуральные очаги в лёгких, а также не отмечать более пяти крупнейших лёгочных очагов на одном КТ-снимке. Для уменьшения вероятности пропусков потенциальных лёгочных очагов каждое исследование

независимо просматривали шестеро рентгенологов. Затем один из участников второй группы (арбитры), составленной из трёх рентгенологов с опытом 10+ лет, просматривал разметку, сделанную рентгенологами первой группы, оценивая достоверность каждой отметки. Арбитры также проводили оценку злокачественности обнаруженных очагов, относя их к категории «злокачественный» или «доброкачественный», руководствуясь рекомендациями Флейшнеровского общества [6].

Этическая экспертиза

Исследование, данные которого использовали для анализа в настоящей работе, получило одобрение независимого этического комитета Московского регионального отделения Российского общества рентгенологов и радиологов (протокол № 2 1-II-2020 от 20 февраля 2020 г.). Все процедуры, выполняемые пациентам в ходе исследования, соответствовали стандартам регионального и национального исследовательского комитета, а также Хельсинкской декларации и Декларации Тайбэя Всемирной медицинской ассоциации.

Статистический анализ

Для определения специфичности и чувствительности индивидуальных специалистов подсчитывали количество истинно положительных, ложноположительных, истинно отрицательных и ложноотрицательных находок для каждого рентгенолога, проводившего первичную интерпретацию. Истинно положительными (ИП) признавали случаи, в которых совпадали мнения рентгенолога и арбитра по поводу наличия и типа лёгочного очага (солидный, полусолидный или уплотнение по типу матового стекла) в конкретной области. К ложноположительным (ЛП) относили случаи, в которых арбитр признал оценку первичного эксперта ошибочной в отношении наличия или типа лёгочного очага в данном участке. Истинно отрицательными (ИО) признавали ситуации, в которых рентгенолог не отметил уплотнения, некорректно, по мнению арбитра, принятого за очаговое образование одним или более из пяти других первичных экспертов. Наконец, к ложноотрицательным (ЛО) относили случаи, в которых рентгенолог не распознавал лёгочный очаг, корректно, по мнению арбитра, идентифицированный одним или более из пяти других участников. При анализе данных исходили из предположения, что суждение арбитра всегда верное.

Чувствительность (sensitivity, Se) рассчитывали по формуле

$$Se = \frac{ИП}{(ИП + ЛО)} . \quad (1)$$

Специфичность (specificity, Sp) вычисляли как

$$Sp = \frac{ИО}{ИО + ЛП} . \quad (2)$$

Для каждого участника определяли индекс Юдена (J):

$$J = Se + Sp - 1 . \quad (3)$$

Для расчёта показателя правильности (accuracy, Acc) различных выборок первичных экспертов ИП результатом признавали случаи, в которых хотя бы один специалист из выборки корректно, по мнению арбитра, идентифицировал лёгочный очаг в конкретной области КТ-снимка. К ИО результатам относили случаи, в которых хотя бы один специалист из выборки не отмечал уплотнения, некорректно, по мнению арбитра, принятого за лёгочный очаг любым другим участником исследования. Правильность вычисляли как

$$Acc = \frac{(ИП + ИО)}{(П + О)} \times 100 , \quad (4)$$

где П — количество корректных находок, О — количество некорректных находок.

Существует целый ряд метрик для оценки согласованности у одного или нескольких исследователей. В частности, O. Gerke и соавт. в своих рекомендациях по систематизации исследований согласованности предлагают использовать анализ Бланда–Альмана [7]. Другими распространёнными метриками являются каппа Коэна [8] и Флейса [9]. Однако при всех преимуществах этих методов они сложны в интерпретации, поэтому авторы настоящей работы остановились на простейшем варианте — процентном показателе согласованности между исследователями (Inter-observer agreement, IOA), который не учитывает фактора случайных совпадений заключений рентгенологов, но при этом интуитивно понятен и достоверно отображает основные закономерности при условии проведения повторных экспериментов. Процентный показатель рассчитывали как долю очагов, для которых мнения экспертов (наличие, тип) совпадали в общем числе совместно размеченных очагов:

$$IOA = \frac{Совпадения}{Совпадения + Несовпадения} \times 100 . \quad (5)$$

Статистический анализ проводили, используя пакеты dplyr [10], irr [11] и ggplot2 [12] для R 3.6.3 [13]. При подготовке данных использовали самостоятельно разработанные скрипты на языке Python 3.8.2 [14].

РЕЗУЛЬТАТЫ

Объекты исследования

Всего в первичной интерпретации КТ-снимков принял участие 31 рентгенолог. Каждый рентгенолог из исходной когорты, состоящей из 15 специалистов, в процессе исследования был заменён другим специалистом по причине отказа или невозможности продолжать исследование; один участник был заменён дважды.

Нагрузка на рентгенологов была распределена неравномерно. Каждый специалист из исходной когорты принял участие в разметке и аннотации в среднем 1050 ± 140 подозрительных структур. Заменявшие их рентгенологи разметили в среднем по 110 ± 42 очаговых образования.

По результатам разметки набор данных включал 72 КТ-исследования, на которых рентгенологи не обнаружили лёгочных очагов от 4 до 30 мм, и 464 КТ-снимка с лёгочными очагами, содержащие в сумме 3151 подтверждённую арбитром находку. Из этого количества 1761 очаг эксперты отнесли к вероятно злокачественным образованиям, 445 — к доброкачественным, 945 уплотнений имели иную природу (содержали кальцинаты, жировую, фиброзную ткань либо жидкость).

Основные результаты исследования

Чувствительность и специфичность рентгенологов, принимавших участие в разметке

В процессе работы над набором данных каждому рентгенологу был присвоен трёхзначный идентификационный номер (ID). В случае замены специалиста новый участник наследовал его ID с дополнительным символом «+». Среднее значение чувствительности составило 34,9% (95% доверительный интервал [ДИ] 30,4–39,4), специфичности — 78,4% (95% ДИ 74,9–81,9), что заметно уступает минимальным показателям, продемонстрированным в схожем сценарии исследования рентгенологами в работе D. Ardila и соавт.: 62,5% (95% ДИ 54,4–70,7) и 95,3% (95% ДИ 94,0–96,6) соответственно [15].

Возможной причиной наблюдаемого различия являются условия разметки, руководствуясь которыми первичные эксперты размечали максимум пять очагов на снимке. Эта рекомендация основана на результатах исследования NELSON, согласно которым риск первичного рака повышается с ростом количества очагов до четырёх, но снижается для пациентов с пятью и более очагами [16]. В случаях множественных очагов (>5) такой подход способен искусственно занижить диагностическую точность первичных экспертов, поскольку привносит дополнительную степень свободы, связанную с конкретным набором очагов, который разметил каждый рентгенолог. Корректировать эту неопределённость можно введением альтернативной классификации находок, признавая за ИП случаи, в которых первичный эксперт разметил хотя бы один подтверждённый очаг на КТ-снимке. При подобной схеме оценки средняя чувствительность первичных экспертов составила 66,2% (95% ДИ 62,1–69,9), специфичность — 78,5% (95% ДИ 72,3–84,8). Однако целью разметки было создание набора данных, предназначенного для обучения алгоритмов искусственного интеллекта, вследствие чего интерес представляет каждая подозрительная структура на КТ-снимке. По этой причине в настоящей работе для оценки диагностической точности использовали критерии,

изложенные в разделе «Методы». В соответствии с этими критериями по показателю индекса Юдена наибольшую эффективность продемонстрировал рентгенолог с ID 012+ ($J=0,472$), наименьшую ($J=-0,188$) — специалист с ID 008+ (табл. 1).

Влияние числа исследователей на правильность интерпретации

Интерпретация двумя первичными экспертами.

В данном анализе рассматривали выборку из 97 КТ-исследований, в интерпретации которых принимал участие рентгенолог с ID 012+, продемонстрировавший наивысший индекс Юдена среди всех участников (см. табл. 1). При таком размере выборки все полученные оценки могут отличаться от средних для полного набора данных не более чем на 10% [17]. Размеченная этим специалистом выборка содержала 53 солидных лёгочных очага, 6 полусолидных и 5 уплотнений по типу матового стекла. Правильность оценок рентгенолога 012+ составила 65,98%: он корректно идентифицировал 28 солидных очагов и избежал 32 из 33 ложноположительных ошибок, совершённых другими специалистами на этих же исследованиях, неверно распознав при этом 2 солидных очага и 1 полусолидный и совершив 34 ложноотрицательных ошибки. Помимо него, в разметке всех 97 КТ-исследований в выборке участвовал также рентгенолог с ID 012, имеющий один из самых низких индексов Юдена (0,058, 24-е место, см. табл. 1). Этот специалист корректно распознал 32 солидных очага, 1 полусолидный, 1 уплотнение по типу матового стекла, избежал 18 ложноположительных ошибок. При согласованности между исследователями равной 59,8% комбинированная правильность их оценок составила 81,44%. Источниками несогласованности послужили расхождение мнений в паре по поводу наличия подозрительной структуры в конкретной области (92,3% случаев) и типа лёгочного очага (7,7% случаев).

Распределение КТ-исследований между специалистами проводили случайным образом. По этой причине получилось так, что интерпретацию всех 97 КТ-исследований в изучаемой выборке проводили только первичные эксперты 012 и 012+. Помимо них в разметке выборки приняли участие 17 рентгенологов (в скобках для каждого указано число размеченных очагов): 000(11), 002(54), 003(30), 004(27), 005(18), 006(40), 007(10), 008(16), 009(17), 010(32), 011(24), 013(30), 014(52), 004+(7), 005+(10), 011+(1) и 014+(9), что предоставило возможность сравнить ситуацию, в которой второе мнение по всем исследованиям в выборке выражает один специалист, с моделью многопользовательской разметки (crowd-tagging), в которой это мнение обеспечивает участник, отбираемый случайным образом из некоторой группы экспертов с переменными показателями специфичности и чувствительности.

Таблица 1. Диагностическая точность участников исследования

ID эксперта	Показатели по отдельным очагам			
	Se, %	Sp, %	Индекс Юдена	Число размеченных очагов*
000	39,52	73,17	0,127	1079
001	32,63	79,04	0,117	1068
002	28,25	80,19	0,084	1045
003	44,05	67,75	0,118	1094
004	31,37	68,75	0,001	844
005	33,08	72,76	0,058	1222
006	36,91	71,32	0,082	1085
007	37,31	73,43	0,107	884
008	42,01	68,00	0,100	1227
009	36,79	79,50	0,163	1265
010	38,62	71,16	0,098	1166
011	26,05	79,51	0,056	853
012	33,97	71,88	0,058	1045
013	38,52	77,40	0,159	1028
014	37,16	82,32	0,195	850
000+	31,63	79,17	0,108	194
001+	52,94	82,46	0,354	108
002+	62,50	57,14	0,196	46
003+	60,71	86,21	0,469	86
004+	27,78	86,49	0,143	110
005+	41,49	75,86	0,173	152
006+	31,34	74,14	0,055	125
007+	29,73	85,71	0,154	86
008+	18,99	62,16	-0,188	176
009+	25,76	85,11	0,109	113
010+	25,00	75,36	0,004	145
011+	31,58	93,33	0,249	68
012+	53,85	93,33	0,472	97
013+	34,29	85,71	0,170	77
014+	17,95	100,0	0,179	63
000++	0,00	94,87	-0,051	48

Примечание. * Учитываются все найденные очаги в КТ-исследованиях, в разметке которых принимал участие эксперт, вне зависимости от того, распознал он их или нет.

В первую группу попали 6 исследователей (табл. 2). Средний индекс Юдена в этой группе составил $0,078 \pm 0,045$ (максимальное значение 0,127, минимальное — 0,001), что превышает показатель участника с ID 012 (0,058). Тем не менее согласованность оценок с экспертом 012+ составила всего лишь 40,2%, а комбинированная правильность оценок — 74,23%. Источником большинства несогласий в паре (97,4%) стало

расхождение мнений по поводу наличия лёгочного очага.

В повторном аналогичном эксперименте анализировали группу с другим составом участников (табл. 3). Число и состав участников различались между группой 1 (см. табл. 2) и группой 2 (см. табл. 3); более того, распределение числа очагов, размеченных каждым из них, было неравномерным.

Таблица 2. Распределение размеченных подозрительных структур в группе 1

ID исследователя	000	002	003	004	005	006
Количество размеченных очагов	11	54	9	3	11	9

Таблица 3. Распределение размеченных подозрительных структур в группе 2

ID исследователя	005+	010	003	004	005	006	008	009
Количество размеченных очагов	10	10	21	9	7	31	8	1

Средний индекс Юдена в группе 2 равнялся $0,099 \pm 0,055$ (максимальное значение 0,173, минимальное — 0,01) и был выше, чем у участника 012 и в группе 1. Согласованность и комбинированная правильность оценок участников группы 2 и рентгенолога 012+ также были наивысшими из трёх рассмотренных вариантов интерпретации КТ-исследований двумя экспертами, составляя 71,1 и 83,50% соответственно. Несогласие между исследователями в 89,3% случаев было ассоциировано с наличием лёгочного очага в данном участке и в 10,7% — с его типом. Средняя правильность интерпретаций при первичной разметке двумя специалистами в любых сочетаниях составила $79,72 \pm 4,87\%$.

Интерпретация тремя и более исследователями. При анализе интерпретации тремя и более исследователями все группы включали исследователей 012 и 012+. При первичной разметке и аннотации тремя рентгенологами согласованность их оценок колебалась от 32,0 до 42,3%, средняя комбинированная правильность составила $89,18 \pm 5,10\%$. Согласованность оценок четырёх независимых специалистов упала до $16,5 \pm 5,7\%$ при росте средней комбинированной правильности до $93,82 \pm 3,57\%$. Для пяти рентгенологов согласованность оценок продолжила снижаться до $9,8 \pm 8,1\%$, а правильность — повышаться до $97,94 \pm 0,14\%$. Наконец, комбинированная правильность шести экспертов составила 100% в условиях нашего эксперимента при согласованности 3,1% (рис. 1). Таким образом, наблюдается значительная обратная корреляция между правильностью и согласованностью оценок экспертов: $r = -0,78$, $p < 0,05$.

В подтверждение выводов P.G. Herman и S.J. Hessel [5] в выборке из 97 исследований при интерпретации шестью специалистами 85,7% ложноположительных ошибок были совершены только одним экспертом, 11,4% — двумя, 2,9% — тремя одновременно. Все шесть экспертов корректно идентифицировали 8,1% положительных находок в выборке; 25,8% ложноотрицательных ошибок были совершены одним экспертом из шести, 8,1% — двумя, 8,1% — тремя, 19,3% — четырьмя, 30,6% — пятью (рис. 2).

Стоимость разметки

Для оценки оптимальной эффективности разметки с позиции рационального использования ресурсов необходимо учитывать стоимость задействования дополнительных экспертов в интерпретации КТ-снимков. Таким образом, можно будет сопоставить улучшение правильности с увеличением расходов на аннотацию исследований.

Поскольку в разметке набора данных принимали участие рентгенологи-волонтеры, их труд не оплачивался. Вследствие этого расчёт стоимости разметки целесообразно проводить в терминах затраченного экспертами времени. В среднем первичный эксперт затрачивал на интерпретацию одного КТ-снимка 12 мин, арбитр — 4 мин. В настоящем исследовании стоимость устранения ошибки *C* в изучаемой выборке из 97 КТ-снимков рассчитывали как разницу средней стоимости разметки заданным числом первичных экспертов с привлечением арбитра и стоимости разметки одним рентгенологом

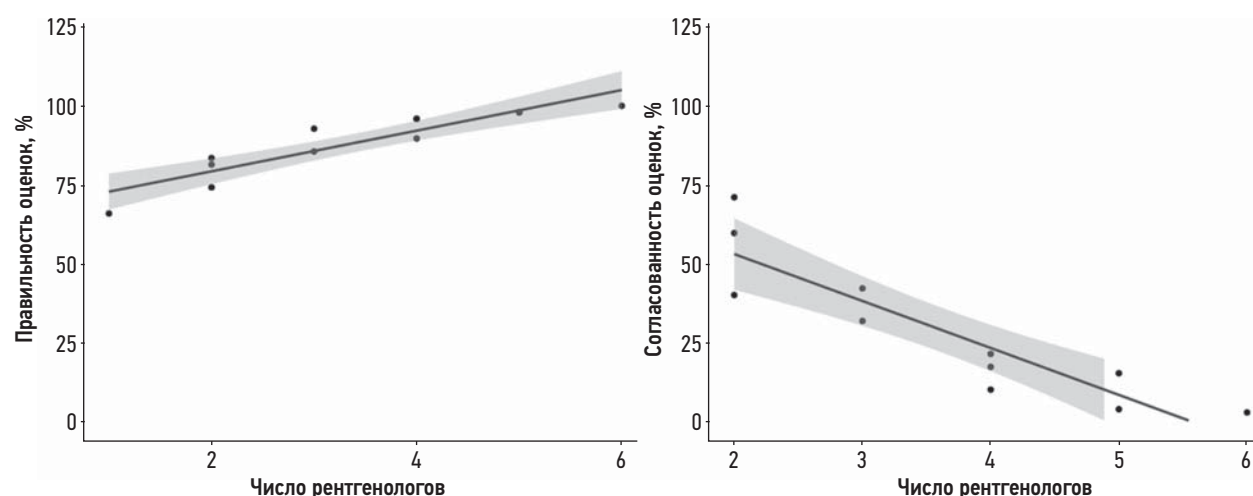


Рис. 1. Правильность и согласованность оценок как функция от числа рентгенологов, принимающих участие в первичной разметке. Серым цветом показан 95% доверительный интервал. Точки соответствуют разным выборкам первичных экспертов. Для экспериментов с двумя, тремя и четырьмя экспертами отбирали по три различающихся выборки из исходных шести рентгенологов; для пяти — по две.

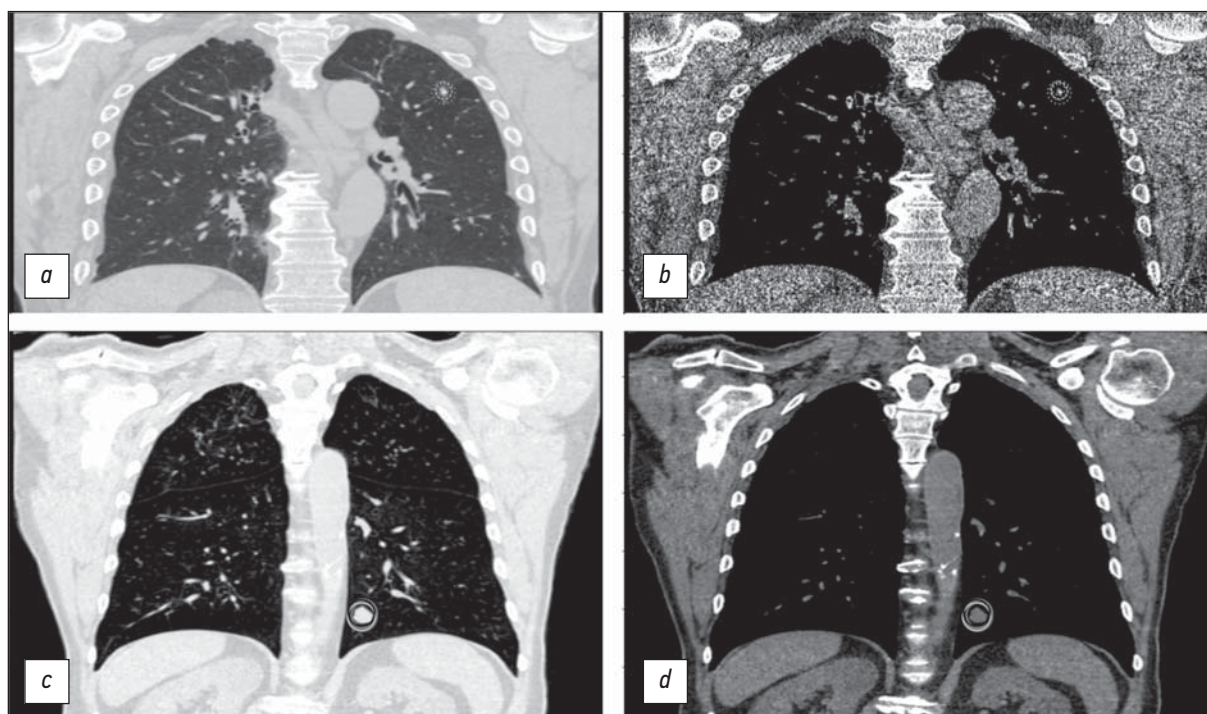


Рис. 2. Примеры КТ-исследований с существенным несогласием (*a, b*, CT LungCa-500 AN RLADD02000018919, ID RLS-DD02000018855) и полным согласием (*c, d*, CT LungCa-500 AN RLAD42D007-25151, ID RLSD42D007-25151) между экспертами. Исследования приведены во фронтальной проекции в лёгочном (*a, c*) и мягкотканном (*b, d*) режимах. Отметки рентгенологов показаны различающимися цветами: *a, b* — очаг разместили пять первичных экспертов из шести, четверо присвоили ему солидный тип и один — полусолидный. Арбитр не согласился с их мнением, признав находку доброкачественным кальцинатом; *c, d* — все шестеро первичных экспертов и арбитр классифицировали очаг как потенциально злокачественный солидный.

без привлечения арбитра, поделённую на количество устранённых ошибок (N_{err}):

$$C = \frac{(n \times 12 \times 97 + n \times 4 \times 97) - 12 \times 97}{N_{err}}, \quad (6)$$

где n — число первичных экспертов.

Эксперт 012+ допустил 33 ложноположительных и ложноотрицательных ошибки. Количество устранённых ошибок, достигаемое за счёт привлечения дополнительных экспертов и проведения арбитража, а также соответствующая стоимость устранения ошибки представлены в табл. 4. Существует закономерность, согласно которой каждый новый первичный эксперт увеличивает стоимость устранения ошибки на $42,5 \pm 10,7$ мин, за исключением одной точки. Разметка набора данных

силами четырёх первичных экспертов с последующим арбитражем сопровождалась резким повышением числа устранённых ошибок и, соответственно, снижением стоимости (см. табл. 4).

Дополнительные результаты исследования

Из-за особенностей дизайна исследования, в котором каждый эксперт интерпретировал индивидуальный КТ-снимок только по одному разу, в рамках настоящей работы не проводили оценку согласованности заключений у отдельных рентгенологов. Среднее значение согласованности оценок между парами специалистов составило $60,5 \pm 5,3\%$, с минимальным значением $53,1\%$ и максимальным $73,0\%$.

Другим способом оценить согласие между первичными экспертами является анализ положительных находок каждого рентгенолога (рис. 3). Для каждого представителя исходной когорты максимальная доля выявленных очагов ($37,6 \pm 5,4\%$) соответствовала уникальным находкам, не распознанным другими экспертами (см. рис. 3, *a*). Затем в порядке убывания следуют находки, с которыми был согласен один ($21,4 \pm 2,8\%$), два ($14,0 \pm 2,0\%$), четыре ($9,5 \pm 2,3\%$), три ($9,2 \pm 1,8\%$) и пять ($8,1 \pm 3,1\%$) первичных экспертов. Только для четырёх рентгенологов из исходной когорты (ID 002, 004, 007 и 010) доля единогласно одобренных находок превышает 10% . Отметим,

Таблица 4. Оценка стоимости устранения ошибки

Число первичных экспертов	Число устранённых ошибок	Стоимость, мин/ошибка
2	15	129,3
3	19	183,8
4	29	173,9
5	31	212,8
6	33	246,9

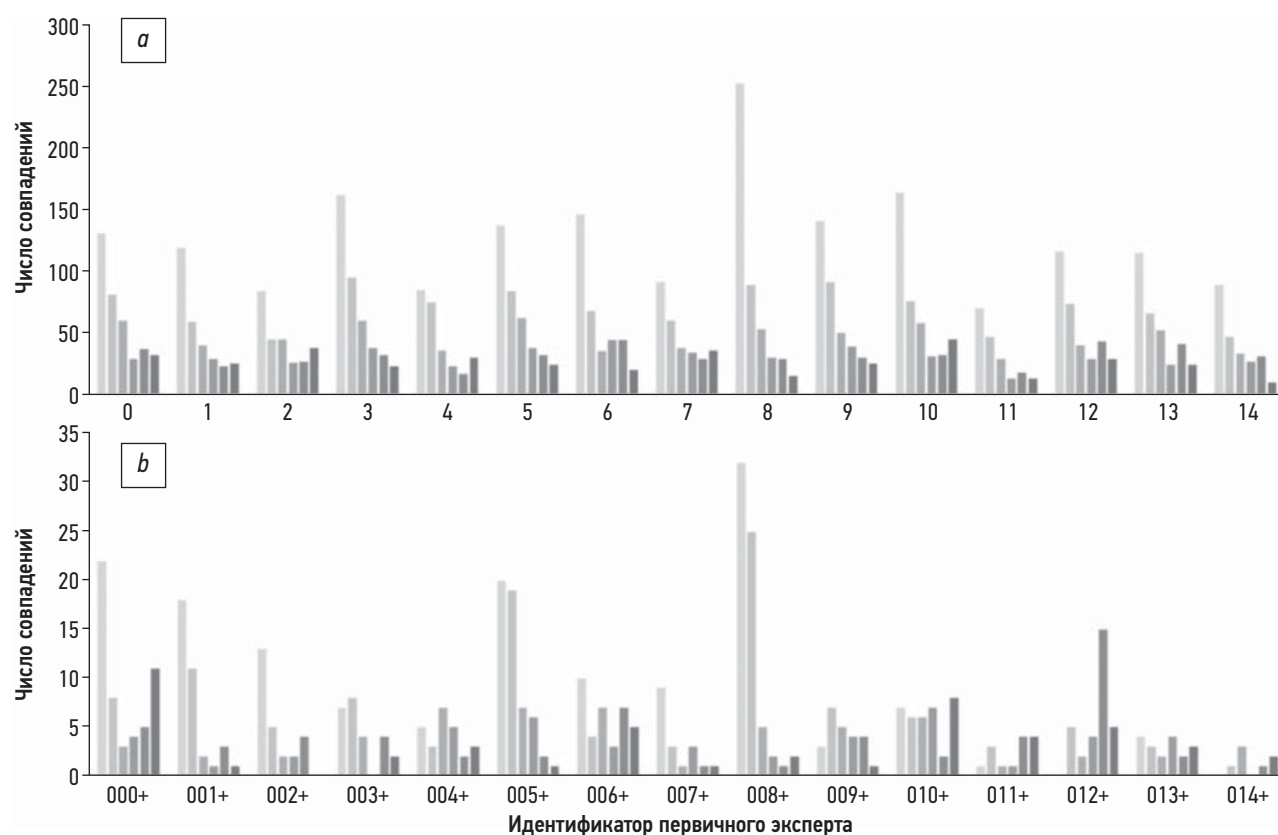


Рис. 3. Согласие между первичными экспертами: *a* — для представителей исходной когорты из 15 рентгенологов; *b* — для рентгенологов, пришедших им на замену. Данные для эксперта с ID 000++ не приведены по причине малого количества отмеченных очагов. Для каждого рентгенолога первый столбец соответствует числу очагов, уникально размеченных этим специалистом (ни один из пяти других экспертов не распознал данную находку). Далее следуют столбцы, соответствующие случаям, когда выявленный рентгенологом очаг отмечали один, два, три, четыре и пять других первичных экспертов. При построении графика не учитывали одобрение арбитра, а также расхождения во мнениях между рентгенологами по поводу типа очага.

что ни один из этих экспертов не входит в лидирующую группу по значению индекса Юдена, рассчитанного по предложенной в настоящей работе методике; более того, эксперт 004 является худшим в когорте по этому показателю (см. табл. 1). В то же время эксперт 014 с максимальным в когорте индексом Юдена (0,195) ничем не выделяется среди своих коллег по согласованности положительных находок (см. рис. 3, *a*).

В когорте рентгенологов, пришедших на замену исходным первичным экспертам, было иное распределение согласованности находок (см. рис. 3, *b*). Максимальная доля выявленных очагов ($28,9 \pm 18,2\%$) была всё так же представлена уникальными находками. Затем следовали находки, выявленные одновременно двумя ($23,3 \pm 11,0\%$), тремя ($13,3 \pm 10,7\%$), пятью ($13,2 \pm 11,9\%$), шестью ($11,5 \pm 9,8\%$) и четырьмя ($9,7 \pm 7,6\%$) экспертами. В этой когорте было уже восемь рентгенологов (ID 000+, 004+, 006+, 010+, 011+, 012+, 013+, 014+), для которых доля единогласно одобренных положительных находок превышала 10%, причём для четырёх из них (ID 000+, 010+, 011+, 014+) она была выше 20%. Тем не менее эти показатели могут быть обусловлены небольшим количеством положительных находок в данной когорте, о чём косвенно свидетельствует высокая вариация их

согласованности, выраженная в отношении средних значений и стандартных отклонений. В качестве примера можно привести эксперта 014+, участвовавшего в интерпретации КТ-исследований, на которых другие эксперты выявили 63 подозрительные структуры (см. табл. 1). Этот эксперт разметил только семь очагов, из которых один был также выявлен одним другим экспертом, три — двумя, один — пятью и два — шестью (см. рис. 3, *b*). При этом эксперт совершил 32 ложноотрицательных ошибки, проигнорировав таким образом ~50% истинно положительных находок. Для этой когорты также не наблюдали корреляции между согласованностью положительных находок и индексом Юдена эксперта.

ОБСУЖДЕНИЕ

Резюме основного результата исследования

Наши результаты демонстрируют, что увеличение количества специалистов, проводящих независимую интерпретацию КТ-исследований, ведёт к росту правильности их оценок, причём уровень квалификации не оказывает существенного воздействия ни на согласованность мнений рентгенологов, ни на их комбинированную правильность. Среди факторов, влияющих

на согласованность заключений между парами исследователей, выделяется расхождение мнений по поводу наличия очага в конкретном участке КТ-снимка.

Обсуждение основного результата исследования

На настоящий момент не существует консенсусного мнения по поводу рекомендуемого количества рентгенологов, участвующих в первичной разметке и аннотации наборов данных медицинских изображений. Это значение, как правило, находится в границах от одного [18, 19] до четырёх [20]. Единственным известным нам исследованием, затрагивающим данный вопрос, является работа P.G. Herman и S.J. Hessel, согласно которой при увеличении числа специалистов, предоставляющих независимые интерпретации исследований, происходит постепенное снижение числа безошибочных описаний [5]. Хотя это, безусловно, интересное наблюдение, оно не представляет собой особой практической ценности, поскольку модель с арбитражем в принципе основана на предположении, что первичные интерпретации будут содержать ошибки. Более того, её эффективность возрастает при условии, что эти ошибки будут разными.

Последнее утверждение не всегда является верным. В частности, результаты настоящей работы показывают, что совершение рентгенологами разных ошибок не приводит автоматически к повышению комбинированной правильности их заключений. В эксперименте с двумя специалистами, проводившими первичную интерпретацию КТ-снимков, наибольший уровень несогласия наблюдали во второй паре (согласованность 40,2%), однако она же продемонстрировала и наименьшую правильность из трёх рассмотренных (74,2 против 81,4 и 83,5%). При этом самое высокое значение правильности с максимальной согласованностью (71,1%) показала третья пара. Тем не менее, согласно полученным в настоящей работе данным, существует значительная отрицательная корреляция между согласованностью оценок экспертов и их правильностью ($r = -0,78$). Так, при первичной интерпретации двумя рентгенологами наблюдали согласованность $57,0 \pm 15,6\%$ при правильности $79,7 \pm 4,9\%$; для пяти рентгенологов эти показатели равнялись $9,8 \pm 8,1\%$ и $97,9 \pm 0,1\%$ соответственно, и эта зависимость сохранялась во всех рассмотренных вариантах разметки набора данных (см. рис. 1).

Согласно результатам настоящего исследования, оптимального сочетания правильности и стоимости разметки позволяет добиться подход с привлечением четырёх первичных экспертов и последующим арбитражем (см. табл. 4). Для него наблюдается резкое увеличение числа устранённых ошибок по сравнению с разметкой силами трёх рентгенологов, что сопровождается снижением времени, затрачиваемого на устранение одной ошибки ($-9,9$ мин). Задействование дополнительных первичных экспертов приводило к дальнейшему росту

правильности интерпретаций, однако это происходило за счёт увеличения стоимости устранения ошибки в среднем на $42,5 \pm 10,7$ мин.

В настоящей работе при отнесении оценок первичных экспертов к категориям ЛО, ИО, ЛП и ИП опирались на предположение, что на каждом КТ-снимке будут размечены все лёгочные очаги. Однако результаты исследования свидетельствуют о том, что участники исследования ограничивались пятью крупнейшими лёгочными очагами на КТ-снимке, выполняя выданные им рекомендации. Таким образом, существенная доля лёгочных очагов была проигнорирована индивидуальными рентгенологами, что сказалось на их показателях диагностической точности, а также значениях согласованности в парах экспертов. Тем не менее расхождения во мнениях первичных экспертов являются желательным исходом при использовании арбитража, поскольку расширяют каталог отмеченных подозрительных структур. Это снижает долю ложноотрицательных находок, даже в условиях искусственных ограничений на число размечаемых очагов. Один из главных выводов настоящей работы — то, что консенсус между несколькими рентгенологами не является необходимым условием для качественной разметки наборов данных. Основная ответственность лежит на арбитрах, которые должны корректно интерпретировать все отмеченные первичными экспертами подозрительные структуры (см. рис. 2, а, б).

Ограничения исследования

Основным ограничением настоящей работы является модель определения эталонных данных (ground truth) — тех находок, которые следует считать лёгочными очагами. При интерпретации КТ-снимков рентгенологи не имели доступа к клиническим, биологическим и геномным данным пациентов; более того, ни для одного из пациентов набор не содержал двух разнесённых во времени исследований, которые позволили бы оценить динамику развития подозрительных структур. Мы исходили также из предположения, что мнение арбитра всегда правильное, и трактовали несогласие между первичным мнением и мнением арбитра всегда в пользу последнего. Однако набор содержит ряд примеров, которые вызывают сомнение в надёжности такого подхода: в частности, 19 лёгочных очагов были отмечены арбитром одновременно как доброкачественные и злокачественные. Это согласуется с результатами S.J. Hessel и соавт., продемонстрировавших, что арбитры способны корректно разрешить лишь порядка 80% несогласий между первичными экспертами [4].

Другим ограничением работы является невозможность проведения оценки воспроизводимости заключений отдельных рентгенологов. Для достижения основных целей исследования использовали ограниченную выборку; для более достоверной статистики оптимальным подходом было бы использование метода размножения

выборки (bootstrap). Наконец, оценка диагностической точности первичных экспертов в настоящем исследовании опиралась на предположение, что они будут замечать все лёгочные очаги. В случае если количество очагов на КТ-снимке превышало пять, это предположение входило в конфликт с рекомендациями по разметке, что могло влиять на итоговые индивидуальные показатели чувствительности и специфичности. Для компенсации этого методологического ограничения авторы исследования провели попытку оценки согласованности по числу положительных находок каждого первичного эксперта, одобренных двумя, тремя, четырьмя и пятью другими рентгенологами (см. рис. 3). Однако такой анализ не учитывает ложноотрицательные ошибки, в связи с чем его результаты не коррелируют с полученными значениями индекса Юдена для каждого эксперта. В довершение, в настоящей работе изучали результаты интерпретации полнодозных КТ-исследований. Таким образом, её выводы могут не распространяться на данные, полученные в ходе скрининг-исследований, для которых характерно использование низкодозных и ультранизкодозных протоколов КТ.

ЗАКЛЮЧЕНИЕ

Несмотря на ограничения, настоящая работа убедительно демонстрирует, что увеличение числа независимых первичных интерпретаций способно повысить их правильность при условии проведения арбитража. При этом квалификация рентгенологов не имеет определяющего значения для качества анализа, поскольку, согласно полученным результатам, комбинированная правильность их оценок не зависела от индивидуальных индексов Юдена. Оптимальное сочетание правильности и стоимости разметки достигается при первичной независимой интерпретации КТ-исследований силами четырёх экспертов. Это наблюдение создаёт теоретическую базу для выработки требований к алгоритмам искусственного интеллекта, предназначенным для использования в диагностике заболеваний посредством разметки подозрительных структур на КТ-снимках и направления внимания врача-рентгенолога. Помимо этого, полученные в работе результаты позволяют обосновать модель проектов многопользовательской разметки наборов данных (crowd-tagging), при которых рост количества разметчиков повлечёт снижение согласованности и одновременный рост качества конечного продукта, обеспечиваемый за счёт арбитража.

СПИСОК ЛИТЕРАТУРЫ

1. Морозов С.П., Кульберг Н.С., Гомболевский В.А., и др. Датасет радиологии Москвы CT LungCa-500. 2018. Режим доступа: https://mosmed.ai/datasets/ct_lungcancer_500/. Дата обращения: 11.02.2021.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

Конфликт интересов. Авторы данной статьи подтвердили отсутствие конфликта интересов, о котором необходимо сообщить.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: Н.С. Кульберг — дизайн набора данных, концептуализация исследования, подготовка и редактирование текста статьи; Р.В. Решетников — статистический анализ, написание текста статьи; В.П. Новик — подготовка набора данных, написание скриптов для сбора данных, статистический анализ; А.Б. Елизаров — подготовка набора данных, написание скриптов для сбора данных; М.А. Гусев — подготовка набора данных, написание скриптов для сбора данных; В.А. Гомболевский — концептуализация исследования, дизайн набора данных; А.В. Владзимирский — концептуализация исследования, редактирование текста статьи; С.П. Морозов — дизайн набора данных, концептуализация и финансирование исследования.

Благодарности. Авторы выражают благодарность Черниной Валерии Юрьевне за методические консультации, а также всем врачам-рентгенологам, принимавшим участие в разметке набора данных.

Funding source. This study was not supported by any external sources of funding.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. The largest contributions are as follows: N.S. Kulberg — dataset design, conceptualization of the study, preparation and editing of the text of the article; R.V. Reshetnikov — statistical analysis, writing of the manuscript; V.P. Novik — dataset preparation, software development for data processing, statistical analysis; A.B. Elizarov — dataset preparation, software development for data processing; M.A. Gusev — dataset preparation, software development for data processing; V.A. Gombolevskiy — conceptualization of the study, dataset design; A.V. Vladzymirskyy — conceptualization of the study, editing of the text of the article; S.P. Morozov — dataset design, conceptualization and funding of the study.

Acknowledgments. The authors express their gratitude to Valeria Chernina for methodological consultations and all radiologists who participated in the dataset markup and annotation.

2. Morozov S.P., Gombolevskiy V.A., Elizarov A.B., et al. A simplified cluster model and a tool adapted for collaborative labeling of lung cancer CT Scans // *Comput Methods Programs Biomed*. 2021. Vol. 206. P. 106111. doi: 10.1016/j.cmpb.2021.106111

3. Kulberg N.S., Gusev M.A., Reshetnikov R.V., et al. Methodology and tools for creating training samples for artificial intelligence systems for recognizing lung cancer on CT images // *Heal Care Russ Fed.* 2020. Vol. 64, N 6. P. 343–350. doi: 10.46563/0044-197X-2020-64-6-343-350
4. Hessel S.J., Herman P.G., Swensson R.G. Improving performance by multiple interpretations of chest radiographs: effectiveness and cost // *Radiology.* 1978. Vol. 127, N 3. P. 589–594. doi: 10.1148/127.3.589
5. Herman P.G., Hessel S.J. Accuracy and its relationship to experience in the interpretation of chest radiographs // *Invest Radiol.* 1975. Vol. 10, N 1. P. 62–67. doi: 10.1097/00004424-197501000-00008
6. MacMahon H., Naidich D.P., Goo J.M., et al. Guidelines for management of incidental pulmonary nodules detected on ct images: from the fleischner society 2017 // *Radiology.* 2017. Vol. 284, N 1. P. 228–243. doi: 10.1148/radiol.2017161659
7. Gerke O., Vilstrup M.H., Segtnan E.A., et al. How to assess intra- and inter-observer agreement with quantitative PET using variance component analysis: a proposal for standardisation // *BMC Med Imaging.* 2016. Vol. 16, N 1. P. 54. doi: 10.1186/s12880-016-0159-3
8. Rasheed K., Rabinowitz Y.S., Remba D., Remba M.J. Interobserver and intraobserver reliability of a classification scheme for corneal topographic patterns // *Br J Ophthalmol.* 1998. Vol. 82, N 12. P. 1401–1406. doi: 10.1136/bjo.82.12.1401
9. Van Riel S.J., Sánchez C.I., Bankier A.A., et al. Observer variability for classification of pulmonary nodules on low-dose ct images and its effect on nodule management // *Radiology.* 2015. Vol. 277, N 3. P. 863–871. doi: 10.1148/radiol.2015142700
10. Wickham H., François R., Henry L., Müller K. *dplyr: A Grammar of Data Manipulation.* R package version 1.0.4. 2021.
11. Gamer M, Lemon J, Fellows I, Singh P. *irr: Various Coefficients of Interrater Reliability and Agreement.* 2019.
12. Wickham H. *ggplot2: elegant Graphics for Data Analysis.* Springer-Verlag New York; 2016. 260 p.
13. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. 2020. Режим доступа: <http://www.r-project.org/index.html>. Дата обращения: 11.02.2021.
14. Van Rossum G., Drake F.L. *Python 3 Reference Manual.* CreateSpace, Scotts Valley, CA; 2009.
15. Ardila D., Kiraly A.P., Bharadwaj S., et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography // *Nat Med.* 2019. Vol. 25, N 6. P. 954–961. doi: 10.1038/s41591-019-0447-x
16. Peters R., Heuvelmans M., Brinkhof S., et al. Prevalence of pulmonary multi-nodularity in CT lung cancer screening. 2015.
17. Creative Research Systems. *The survey systems: Sample size calculator.* 2012.
18. Hugo G.D., Weiss E., Sleeman W.C., et al. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer // *Med Phys.* 2017. Vol. 44, N 2. P. 762–771. doi: 10.1002/mp.12059
19. Bakr S., Gevaert O., Echegaray S., et al. A radiogenomic dataset of non-small cell lung cancer // *Sci Data.* 2018. Vol. 5. P. 180202. doi: 10.1038/sdata.2018.202
20. Armato S.G., McLennan G., Bidaut L., et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on ct scans // *Med Phys.* 2011. Vol. 38, N 2. P. 915–931. doi: 10.1118/1.3528204

REFERENCES

1. Morozov SP, Kulberg NS, Gombolevsky VA, et al. Moscow Radiology Dataset CT LungCa-500. 2018. (In Russ). Available from: https://mosmed.ai/datasets/ct_lungcancer_500/
2. Morozov SP, Gombolevskiy VA, Elizarov AB, et al. A simplified cluster model and a tool adapted for collaborative labeling of lung cancer CT Scans. *Comput Methods Programs Biomed.* 2021;206:106111. doi: 10.1016/j.cmpb.2021.106111
3. Kulberg NS, Gusev MA, Reshetnikov RV, et al. Methodology and tools for creating training samples for artificial intelligence systems for recognizing lung cancer on CT images. *Heal Care Russ Fed.* 2020;64(6):343–350. doi: 10.46563/0044-197X-2020-64-6-343-350
4. Hessel SJ, Herman PG, Swensson RG. Improving performance by multiple interpretations of chest radiographs: effectiveness and cost. *Radiology.* 1978;127(3):589–594. doi: 10.1148/127.3.589
5. Herman PG, Hessel SJ. Accuracy and its relationship to experience in the interpretation of chest radiographs. *Invest Radiol.* 1975;10(1):62–67. doi: 10.1097/00004424-197501000-00008
6. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on ct images: from the fleischner society 2017. *Radiology.* 2017;284:228–243. doi: 10.1148/radiol.2017161659
7. Gerke O, Vilstrup MH, Segtnan EA, et al. How to assess intra- and inter-observer agreement with quantitative PET using variance component analysis: a proposal for standardisation. *BMC Med Imaging.* 2016;16(1):54. doi: 10.1186/s12880-016-0159-3
8. Rasheed K, Rabinowitz YS, Remba D, Remba MJ. Interobserver and intraobserver reliability of a classification scheme for corneal topographic patterns. *Br J Ophthalmol.* 1998;82(12):1401–1406. doi: 10.1136/bjo.82.12.1401
9. Van Riel SJ, Sánchez CI, Bankier AA, et al. Observer variability for classification of pulmonary nodules on low-dose ct images and its effect on nodule management. *Radiology.* 2015;277(3):863–871. doi: 10.1148/radiol.2015142700
10. Wickham H, François R, Henry L, Müller K. *dplyr: A Grammar of Data Manipulation.* R package version 1.0.4. 2021.
11. Gamer M, Lemon J, Fellows I, Singh P. *irr: Various Coefficients of Interrater Reliability and Agreement.* 2019.
12. Wickham H. *ggplot2: elegant Graphics for Data Analysis.* Springer-Verlag New York; 2016. 260 p.
13. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria; 2020. Available from: <http://www.r-project.org/index.html>
14. Van Rossum G, Drake FL. *Python 3 Reference Manual.* CreateSpace, Scotts Valley, CA; 2009.
15. Ardila D, Kiraly AP, Bharadwaj S, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med.* 2019;25(6):954–961. doi: 10.1038/s41591-019-0447-x
16. Peters R, Heuvelmans M, Brinkhof S, et al. Prevalence of pulmonary multi-nodularity in CT lung cancer screening. 2015.

17. Creative Research Systems. The survey systems: Sample size calculator. 2012.

18. Hugo GD, Weiss E, Sleeman WC, et al. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer. *Med Phys.* 2017;44(2):762–771. doi: 10.1002/mp.12059

19. Bakr S, Gevaert O, Echegaray S, et al. A radiogenomic dataset of non-small cell lung cancer. *Sci Data.* 2018;5:180202. doi: 10.1038/sdata.2018.202

20. Armato SG, McLennan G, Bidaut L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on ct scans. *Med Phys.* 2011;38(2):915–931. doi: 10.1118/1.3528204

ОБ АВТОРАХ

*** Кульберг Николай Сергеевич**, к.ф.-м.н.;
адрес: Россия, 127051, Москва, ул. Петровка, д. 24;
ORCID: <https://orcid.org/0000-0001-7046-7157>;
eLibrary SPIN: 2135-9543; e-mail: kulberg@npcmr.ru

Решетников Роман Владимирович, к.ф.-м.н.;
ORCID: <https://orcid.org/0000-0002-9661-0254>;
eLibrary SPIN: 8592-0558; e-mail: reshetnikov@fbb.msu.ru

Новик Владимир Петрович;
ORCID: <https://orcid.org/0000-0002-6752-1375>;
eLibrary SPIN: 2251-1016; e-mail: v.novik@npcmr.ru

Елизаров Алексей Борисович, к.ф.-м.н.;
ORCID: <https://orcid.org/0000-0003-3786-4171>;
eLibrary SPIN: 7025-1257; e-mail: a.elizarov@npcmr.ru

Гусев Максим Александрович;
ORCID: <https://orcid.org/0000-0001-8864-8722>;
eLibrary SPIN: 1526-1140; e-mail: m.gusev@npcmr.ru

Гомболевский Виктор Александрович, к.м.н.;
ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: g_victor@mail.ru

Владимирский Антон Вячеславович, д.м.н., профессор;
ORCID: <https://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: a.vladimirsky@npcmr.ru

Морозов Сергей Павлович, д.м.н., профессор;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720; e-mail: morozov@npcmr.ru

AUTHORS' INFO

*** Nikolas S. Kulberg**, Cand. Sci. (Phys.-Math.);
address: 24 Petrovka str., 109029, Moscow, Russia;
ORCID: <https://orcid.org/0000-0001-7046-7157>;
eLibrary SPIN: 2135-9543; e-mail: kulberg@npcmr.ru

Roman V. Reshetnikov, Cand. Sci. (Phys.-Math.);
ORCID: <https://orcid.org/0000-0002-9661-0254>;
eLibrary SPIN: 8592-0558; e-mail: reshetnikov@fbb.msu.ru

Vladimir P. Novik;
ORCID: <https://orcid.org/0000-0002-6752-1375>;
eLibrary SPIN: 2251-1016; e-mail: v.novik@npcmr.ru

Alexey B. Elizarov, Cand. Sci. (Phys.-Math.);
ORCID: <https://orcid.org/0000-0003-3786-4171>;
eLibrary SPIN: 7025-1257; e-mail: a.elizarov@npcmr.ru

Maxim A. Gusev;
ORCID: <https://orcid.org/0000-0001-8864-8722>;
eLibrary SPIN: 1526-1140; e-mail: m.gusev@npcmr.ru

Victor A. Gombolevskiy, MD, Cand. Sci. (Med.);
ORCID: <https://orcid.org/0000-0003-1816-1315>;
eLibrary SPIN: 6810-3279; e-mail: g_victor@mail.ru

Anton V. Vladzimirskyy, MD, Dr. Sci. (Med.), Professor;
ORCID: <https://orcid.org/0000-0002-2990-7736>;
eLibrary SPIN: 3602-7120; e-mail: a.vladimirsky@npcmr.ru

Sergey P. Morozov, MD, Dr. Sci. (Med.), Professor;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720; e-mail: morozov@npcmr.ru

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD70821>

Повышение качества отчётов о наблюдательных исследованиях в эпидемиологии (STROBE): разъяснения и уточнения

J.P. Vandenbroucke¹, E. von Elm^{2, 3}, D.G. Altman⁴, P.C. Gotzsche⁵, C.D. Mulrow⁶, S.J. Pocock⁷, C. Poole⁸, J.J. Schlesselman^{9, 10}, M. Egger^{2, 11} for the STROBE Initiative

¹ Leiden University Medical Center, Лейден, Нидерланды

² University of Bern, Берн, Швейцария

³ University Medical Centre, Фрайбург, Германия

⁴ Cancer Research UK/NHS Centre for Statistics in Medicine, Оксфорд, Великобритания

⁵ Nordic Cochrane Centre, Rigshospitalet, Копенгаген, Дания

⁶ University of Texas Health Science Center, Сан-Антонио, США

⁷ London School of Hygiene and Tropical Medicine, Лондон, Великобритания

⁸ University of North Carolina School of Public Health, Чапел-Хилл, США

⁹ University of Pittsburgh Graduate School of Public Health, Питтсбург, США

¹⁰ University of Pittsburgh Cancer Institute, Питтсбург, США

¹¹ Department of Social Medicine, University of Bristol, Бристоль, Великобритания

АННОТАЦИЯ

Большинство медицинских исследований являются наблюдательными (*observational*). Сообщения о таких исследованиях часто невысокого качества, что затрудняет оценку сильных и слабых сторон работы, а также обобщаемости (*generalizability*) её результатов. Принимая во внимание эмпирические свидетельства и теоретические соображения, группа методологов, исследователей и научных редакторов разработала рекомендации «Повышение качества отчётов о наблюдательных исследованиях в эпидемиологии (STROBE): разъяснения и уточнения». Рекомендации STROBE содержат 22 пункта, связанных с оформлением следующих разделов научных статей: название, аннотация, введение, методы, результаты и их обсуждение, при этом 18 пунктов являются общими для когортных исследований (*cohort studies*), исследований «случай–контроль» (*case-control studies*) и одномоментных исследований (*cross-sectional studies*); 4 пункта специфичны для каждого из указанных дизайнов исследований (*study designs*). STROBE — руководство для авторов, необходимое для повышения качества отчётов о наблюдательных исследованиях, облегчающее критическую оценку исследования и его интерпретацию рецензентами, редакторами журналов и читателями. Цель этой разъясняющей и уточняющей статьи — способствовать более широкому применению, пониманию и распространению стандартов STROBE. В ней даётся разъяснение смысла и обоснование применения каждого пункта руководства (*checklist*). По каждому пункту приводятся один или несколько опубликованных примеров правильного представления исследований и, при возможности, библиографические ссылки на подходящие эмпирические исследования и методологическую литературу. Представлены примеры потоковых диаграмм (*flow diagrams*) для описания последовательности исследования. Рекомендации STROBE, настоящая статья и соответствующий веб-сайт (<http://www.strobe-statement.org/>) должны стать полезным источником для повышения качества отчётов о результатах наблюдательных исследований.

Настоящая статья является русскоязычным переводом оригинальной публикации [Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLoS Med.* 2007;4(10):e297. doi: 10.1371/journal.pmed.0040297] под редакцией д.м.н. Р.Т. Сайгитова.

Ключевые слова: STROBE; наблюдательные исследования; когортные исследования; одномоментные исследования; случай–контроль; дизайн исследования.

Как цитировать

Vandenbroucke J.P., von Elm E., Altman D.G., Gotzsche P.C., Mulrow C.D., Pocock S.J., Poole C., Schlesselman J.J., Egger M. Повышение качества отчётов о наблюдательных исследованиях в эпидемиологии (STROBE): разъяснения и уточнения // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 119–169. DOI: <https://doi.org/10.17816/DD70821>

DOI: <https://doi.org/10.17816/DD70821>

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. Translation to Russian

Jan P. Vandenbroucke¹, Erik von Elm^{2,3}, Douglas G. Altman⁴, Peter C. Gotzsche⁵, Cynthia D. Mulrow⁶, Stuart J. Pocock⁷, Charles Poole⁸, James J. Schlesselman^{9,10}, Matthias Egger^{2,11} for the STROBE Initiative

¹ Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

² Institute of Social & Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

³ Department of Medical Biometry and Medical Informatics, University Medical Centre, Freiburg, Germany

⁴ Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, United Kingdom

⁵ Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

⁶ University of Texas Health Science Center, San Antonio, United States of America

⁷ Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom

⁸ Department of Epidemiology, University of North Carolina School of Public Health, Chapel Hill, United States of America

⁹ University of Pittsburgh Graduate School of Public Health, Pittsburgh, United States of America

¹⁰ University of Pittsburgh Cancer Institute, Pittsburgh, United States of America

¹¹ Department of Social Medicine, University of Bristol, Bristol, United Kingdom

ABSTRACT

Much medical research is observational. The reporting of observational studies is often of insufficient quality. Poor reporting hampers the assessment of the strengths and weaknesses of a study and the generalisability of its results. Taking into account empirical evidence and theoretical considerations, a group of methodologists, researchers, and editors developed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations to improve the quality of reporting of observational studies. The STROBE Statement consists of a checklist of 22 items, which relate to the title, abstract, introduction, methods, results and discussion sections of articles. Eighteen items are common to cohort studies, case-control studies and cross-sectional studies and four are specific to each of the three study designs. The STROBE Statement provides guidance to authors about how to improve the reporting of observational studies and facilitates critical appraisal and interpretation of studies by reviewers, journal editors and readers. This explanatory and elaboration document is intended to enhance the use, understanding, and dissemination of the STROBE Statement. The meaning and rationale for each checklist item are presented. For each item, one or several published examples and, where possible, references to relevant empirical studies and methodological literature are provided. Examples of useful flow diagrams are also included. The STROBE Statement, this document, and the associated Web site (<http://www.strobe-statement.org/>) should be helpful resources to improve reporting of observational research.

This article is the reprint with Russian translation of the original that can be observed here: Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLoS Med.* 2007;4(10):e297. doi: 10.1371/journal.pmed.0040297

Keywords: STROBE; observational studies; cohort studies; case-control studies; cross-sectional studies; study design.

To cite this article

Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. Translation to Russian. *Digital Diagnostics.* 2021;2(2):119–169. DOI: <https://doi.org/10.17816/DD70821>

Received: 19.05.2021

Accepted: 20.05.2021

Published: 25.05.2021

DOI: <https://doi.org/10.17816/DD70821>

提高流行病学观察性研究报告的质量 (STROBE) : 澄清和澄清

Jan P. Vandenbroucke¹, Erik von Elm^{2,3}, Douglas G. Altman⁴, Peter C. Gotzsche⁵,
Cynthia D. Mulrow⁶, Stuart J. Pocock⁷, Charles Poole⁸, James J. Schlesselman^{9,10},
Matthias Egger^{2,11} for the STROBE Initiative

¹ Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

² Institute of Social & Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

³ Department of Medical Biometry and Medical Informatics, University Medical Centre, Freiburg, Germany

⁴ Cancer Research UK/NHS Centre for Statistics in Medicine, Oxford, United Kingdom

⁵ Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

⁶ University of Texas Health Science Center, San Antonio, United States of America

⁷ Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, United Kingdom

⁸ Department of Epidemiology, University of North Carolina School of Public Health, Chapel Hill, United States of America

⁹ University of Pittsburgh Graduate School of Public Health, Pittsburgh, United States of America

¹⁰ University of Pittsburgh Cancer Institute, Pittsburgh, United States of America

¹¹ Department of Social Medicine, University of Bristol, Bristol, United Kingdom

简评

大多数医学研究都是观察性的。(observational). 此类研究的报告通常质量较差, 难以评估工作的优缺点, 以及其结果(generalizability) 的普遍性。 考虑到经验证据和理论考虑, 一组方法学家、研究人员和科学编辑提出了“提高流行病学观察研究报告的质量 (STROBE): 澄清和澄清”的建议。 STROBE建议包含22个段落, 涉及科学文章的以下部分: 标题、注释、导言、方法、结果及其讨论, 共有18项是队列研究(cohort studies)所共有的, 病例对照研究(case-control studies)和横断面研究(cross-sectional studies); 4个项目适用于所述的每个研究设计(study designs)。STROBE - 提高观察研究报告质量所需的作者指南, 便于评论家对研究进行批判性评估和解释, 杂志编辑和读者。这篇澄清和澄清文章的目的是促进 STROBE 标准的更广泛应用、理解和传播。 它解释了指南中每一段的含义和适用理由(checklist)。每个项目都提供一个或多个已发表的正确介绍研究的例子, 并在可能的情况下, 参考相关实证研究和方法论文献的书目参考. 给出了流程图(Flow Diagrams)实例来描述研究序列。STROBE 指南、本文和相关网站 (<http://www.strobe-statement.org/>) 应该是提高观察报告质量的有用资源。

本文是俄文原版翻译[Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Med 2007;4(10):e297. doi: 10.1371/journal.pmed.0040297] 由医学博士编辑。 R. T. Saygitov。

关键词: STROBE; 观察性研究; 队列研究; 横断面研究; 病例控制; 学习规划。

引用本文:

Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M. 提高流行病学观察性研究报告的质量 (STROBE): 澄清和澄清. *Digital Diagnostics*. 2021;2(2):119-169. DOI: <https://doi.org/10.17816/DD70821>

收到: 19.05.2021

接受: 20.05.2021

发布日期: 25.05.2021

Список сокращений

ДИ — доверительный интервал
ИОР — избыточный относительный риск (вследствие взаимодействия)
ОР — относительный риск

STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) — повышение качества отчётов о наблюдательных исследованиях в эпидемиологии

ВВЕДЕНИЕ

Осмысленная практика оказания медицинской помощи требует знаний этиологии и патогенеза, диагностики, прогнозирования и терапии заболеваний. Рандомизированные испытания дают ценные сведения о методах лечения и иных вмешательствах. Однако существенная часть клинических знаний, а также информация в области общественного здравоохранения основываются на наблюдательных исследованиях [1]. Примерно 9 из 10 научных работ, опубликованных в медицинских журналах, описывают наблюдательные исследования [2, 3].

РЕКОМЕНДАЦИИ STROBE

Отчёты о наблюдательных исследованиях часто бывают недостаточно подробными и чёткими, что не позволяет оценить сильные и слабые стороны работы [4, 5]. Для повышения качества таких отчётов мы разработали правила оформления разделов статей — названия, аннотации, введения, методов, результатов и их обсуждения (табл. 1). Эти правила опубликованы в нескольких журналах [6]. Наша цель — обеспечить чёткое представление того, что было запланировано, сделано и обнаружено в результате наблюдательных исследований. Мы обращаем внимание на то, что наши рекомендации не являются обязательным предписанием для организации или проведения исследований, выбора методологии или единого формата представления результатов.

STROBE даёт общие рекомендации к оформлению отчётов об описательных наблюдательных исследованиях, а также тех исследованиях, в ходе которых изучается связь между воздействием (*exposures*) и исходами (*outcomes*) здоровья. Рекомендации STROBE относятся к 3 основным типам наблюдательных исследований: когортным, «случай–контроль» и одномоментным. Авторы используют различную терминологию для описания таких исследований. Например, термины «проспективное исследование» и «продольное исследование» используются как синонимы «когортного исследования», а «исследование распространённости» — как синоним «одномоментного исследования». Мы выбрали данную терминологию, поскольку она широко используется, но часто некорректно [7] или неточно [8]. Во вставке 1 представлены отличительные признаки трёх дизайнов исследования.

СФЕРА НАБЛЮДАТЕЛЬНЫХ ИССЛЕДОВАНИЙ

Наблюдательные исследования служат широкому спектру целей: от сообщения о полученных впервые данных в отношении потенциальной причины заболевания до проверки значимости связей (*associations*), о которых сообщалось ранее. Идеи для исследований могут возникать в результате клинических наблюдений или из описания биологических явлений. Неформальное видение определённых данных может также способствовать возникновению идей, ведущих к дальнейшим исследованиям. Подобно врачу-клиницисту, который осмотрел тысячи пациентов и отмечает одного, чем-то привлёкшего к себе его внимание, исследователь также может заметить нечто особенное в данных. И хотя изучение одних и тех же данных с разных точек зрения может быть невозможным или нежелательным [9], для подтверждения или опровержения первоначальных наблюдений часто возникает необходимость в проведении дальнейших исследований [10]. Существующие данные могут использоваться для изучения новых предположений о потенциальных причинных факторах, и вместе с тем могут быть достаточными для их опровержения или подтверждения. В некоторых случаях исследования проводятся именно с целью преодоления потенциальных проблем, связанных с предыдущими сообщениями. Главная задача таких исследований — сбор новых данных, а не анализ уже существующих. Это приводит к появлению различных точек зрения, например о преимуществах рассмотрения подгрупп или о важности заранее определённого размера выборки (*sample size*). Инициатива STROBE направлена на достижение согласованного описания наблюдательных исследований с разнообразными целями — от научных открытий до опровержения или подтверждения ранее полученных результатов. Для отдельных случаев приводятся конкретные рекомендации.

КАК ИСПОЛЬЗОВАТЬ ЭТУ СТАТЬЮ

Эта статья связана с краткой версией рекомендаций STROBE, представленных в форме перечня и опубликованных ранее [6], и является неотъемлемой частью стандарта STROBE. Наша задача — объяснить, как правильно

Таблица 1. STROBE: перечень пунктов, которые должны быть представлены в отчётах о наблюдательных исследованиях

	№ пункта	Рекомендации
НАЗВАНИЕ И АННОТАЦИЯ	1	а) Укажите в названии статьи или в аннотации наименование дизайна исследования, используя общепринятые термины. б) Включите в аннотацию информативное и сбалансированное обобщение того, что было сделано и обнаружено.
ВВЕДЕНИЕ		
Актуальность / обоснование	2	Объясните актуальность и обоснуйте необходимость проведения представляемого научного исследования.
Цели	3	Укажите цели исследования, включая все предварительные гипотезы.
МЕТОДЫ		
Дизайн исследования	4	Укажите основные элементы дизайна исследования.
Условия проведения	5	Опишите условия, место и соответствующие даты проведения исследования и все связанные с этим данные, включая периоды привлечения участников, воздействия, наблюдения и сбора данных.
Участники	6	а) Когортное исследование. Укажите критерии, источники и методы отбора участников. Опишите методику наблюдения. Исследование «случай–контроль». Укажите критерии отбора, источники и методы установления «случаев» и «контролей». Обоснуйте выбор «случаев» и «контролей». Одномоментное исследование. Укажите критерии, источники и методы отбора участников. б) Когортное исследование. Для исследований с подобранными группами опишите критерии подбора и количество участников, подвергшихся и не подвергавшихся воздействию изучаемого фактора. Исследование «случай–контроль». Для исследований с подобранными группами опишите критерии подбора и количество «контролей» на один «случай».
Переменные	7	Чётко определите все исходы, воздействия, прогностические факторы, потенциальные вмешивающиеся факторы, модификаторы эффекта. Опишите диагностические критерии, если применимо.
Источник данных, измерения	8*	Для каждой переменной, представляющей интерес, приведите источники данных и подробные сведения о методах оценки (измерения). Укажите, были ли сопоставимы методы измерения в группах.
Систематические ошибки	9	Опишите все усилия по устранению потенциальных источников систематических ошибок.
Объём исследования	10	Объясните, как был определён объём (размер) исследования.
Количественные переменные	11	Укажите, какие количественные переменные были использованы в анализе и, если применимо, какие категории (группы) на основании этих переменных были выбраны и почему.
Статистические методы	12	а) Опишите все статистические методы, включая использованные для контроля влияния вмешивающихся факторов. б) Опишите все методы, использованные для изучения подгрупп и взаимодействий. в) Объясните, как устраняли последствия отсутствующих (неполных) данных. г) Когортное исследование. Если применимо, опишите, как решали проблему выбывших из-под наблюдения. Исследование «случай–контроль». Если применимо, объясните, как производили подбор (matching) «случаев» и «контролей». Одномоментное исследование. Если применимо, опишите аналитические методы, учитывающие стратегию формирования выборки. д) Опишите использованные анализы чувствительности.

Таблица 1. Окончание

	№ пункта	Рекомендации
РЕЗУЛЬТАТЫ		
Участники	13*	а) Укажите количество лиц на каждом этапе исследования: число потенциально подходящих для участия в исследовании, обследованных на предмет соответствия критериям отбора, допущенных к участию, включённых в исследование, завершивших последующее наблюдение, а также участников, чьи результаты были проанализированы. б) Укажите причины выбывания на каждом этапе исследования. в) Рассмотрите возможность использования потоковой диаграммы.
Описательные данные	14*	а) Охарактеризуйте участников исследования (например, укажите демографические, клинические, социальные характеристики). Приведите данные о воздействиях и потенциальных вмешивающихся факторах. б) Для каждой из анализируемых переменных укажите количество участников с пропущенными данными. в) Когортное исследование. Опишите характеристики времени наблюдения (например, укажите среднюю и общую продолжительность наблюдения).
Исходы	15*	Когортное исследование. Укажите количество событий исхода или представьте сводку основных показателей исследования за время наблюдения. Исследование «случай–контроль» Укажите количество случаев по каждой категории воздействия или сводные показатели воздействия. Одномоментное исследование. Укажите количество событий исхода или представьте сводку основных показателей исследования.
Основные результаты	16	а) Укажите нескорректированные оценки и, если применимо, оценки с поправкой на вмешивающиеся факторы и их точность (например, 95% доверительный интервал). Объясните, влияние каких вмешивающихся факторов было учтено, и почему именно эти факторы были выбраны. б) При категоризации непрерывных переменных укажите границы категорий. в) Если применимо, рассмотрите возможность преобразования относительного риска в абсолютный для значимого периода времени.
Другие анализы	17	Если проводились другие виды анализа, например анализ в подгруппах, взаимодействий, чувствительности, — укажите это.
ОБСУЖДЕНИЕ		
Основные результаты	18	Обобщите основные результаты, ссылаясь на цели исследования.
Ограничения	19	Обсудите ограничения исследования с анализом источников потенциальных систематических ошибок или неточностей. Обсудите направление действия и значимость потенциальных систематических ошибок.
Интерпретация результатов	20	Дайте осторожную, но достаточно полную интерпретацию результатов с учётом целей, ограничений, многочисленных анализов, результатов схожих исследований и других уместных свидетельств.
Обобщаемость	21	Обсудите обобщаемость (внешнюю достоверность) результатов исследования.
ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ		
Финансирование	22	Укажите источники финансирования и роль спонсоров в настоящем исследовании и, если применимо, в оригинальном исследовании, на котором основывается представляемая статья.

Примечание. * В исследованиях «случай–контроль» рекомендуется давать информацию отдельно для «случаев» и «контролей» и, если применимо, для подвергшихся и не подвергавшихся воздействию групп в когортных и одномоментных исследованиях. В данной статье приводятся все пункты перечня рекомендаций (*checklist*) с методологическим обоснованием и примерами качественного представления исследований в опубликованных работах. Перечень рекомендаций STROBE эффективнее всего использовать в сочетании с этой статьей (имеется в свободном доступе на сайтах журналов *PLoS Medicine* (<http://www.plosmedicine.org/>), *Annals of Internal Medicine* (<http://www.annals.org/>), *Epidemiology* (<http://www.epidem.com/>). Версии перечня отдельно для когортных исследований, исследований «случай–контроль» и одномоментных исследований доступны на сайте STROBE (<http://www.strobe-statement.org/>). doi: 10.1371/journal.pmed.0040297.t001

представлять результаты исследования, а не как его проводить. Мы прилагаем подробные пояснения к каждому пункту перечня. Пояснению предшествует пример того, что мы считаем правильно оформленным сообщением. Это не означает, что исследование, из которого был взят данный пример, было качественно представлено или выполнено. Это также не означает, что его результаты были надёжными (*reliable*) и позже подтверждены другими исследователями. Это всего лишь означает, что данный конкретный фрагмент текста статьи был корректно представлен в этом исследовании. Предложенные нами пояснения и примеры мы сопровождали вставками (№ 1–8) с дополнительной справочной информацией о некоторых теоретических моментах или технических деталях. Для более полного их понимания может потребоваться изучение учебников или методологических источников, которые нами цитируются.

Рекомендации STROBE не ориентированы на специфические исследования, такие как изучение генетических связей, моделирование инфекционных заболеваний, описание случая или серии случаев [11, 12]. Поскольку многие ключевые пункты перечня STROBE применимы к подобным дизайнам, наши рекомендации могут оказаться уместными и полезными для тех авторов, которые представляют результаты таких исследований. Для авторов работ, посвящённых изучению методов

диагностики, онкомаркеров и генетических связей, могут быть полезны рекомендации STARD [13], REMARK [14] и STREGA [15].

О ПУНКТАХ ПЕРЕЧНЯ РЕКОМЕНДАЦИЙ STROBE

Далее мы обсудим и объясним 22 пункта перечня рекомендаций STROBE (см. табл. 1) и приведём опубликованные примеры к каждому из них. В некоторых примерах были удалены цитаты или добавлены сокращения. Ко всем трём типам исследований относятся 18 пунктов, тогда как 4 пункта касаются конкретных дизайнов. Пункты, помеченные звёздочкой (например, пункт 8*), указывают на то, что информацию следует представлять отдельно для «случаев» и «контролей» в исследованиях «случай–контроль» или для групп, подвергшихся и не подвергавшихся воздействию в когортных и одномоментных исследованиях. Мы советуем авторам применять в своей работе все предложенные пункты, но их расположение или очерёдность может быть произвольной. Например, мы предлагаем представлять результаты отдельными пунктами, признавая вместе с тем, что авторы могут объединить информацию по нескольким пунктам в одном разделе текста или в таблице.

Вставка 1. Основные дизайны исследований, охватываемые рекомендациями STROBE

Когортное исследование, исследование «случай–контроль» и одномоментное исследование представляют собой различные подходы к изучению возникновения событий, связанных со здоровьем, в данной популяции и период времени. Эти исследования могут охватывать такие события, как возникновение болезни или наступление ремиссии, инвалидность или осложнения, смерть или выживаемость (*survival*), а также появление факторов риска.

В **когортных исследованиях** исследователи наблюдают за людьми в течение времени. Они получают данные об участниках исследования и воздействии на них изучаемых факторов на момент включения в исследование и через какое-то время оценивают результаты воздействия (исходы). Авторы обычно противопоставляют участников, которые подверглись и не подвергались воздействию, или группы людей с различными категориями воздействия. Они также могут оценивать несколько различных исходов, а также изучать переменные воздействия и исхода в разные моменты времени в ходе наблюдения. **Закрытые когорты** (к примеру, возрастные) включают определённое количество участников, после чего за ними ведётся наблюдение, часто на протяжении установленных промежутков времени до фиксированной даты окончания исследования. В **открытых когортах** (*open cohorts*) выборка исследования (*study population*) динамична: участники исследования (например, жители города) попадают в неё и выбывают из неё в разные моменты времени. Открытая когорта может меняться в результате смертей, рождения или миграции участников исследования, но состав когорты с учётом возраста и пола может оставаться примерно одинаковым, особенно на протяжении короткого периода времени. В закрытых когортах (*closed cohorts*) могут оцениваться кумулятивная инцидентность (*cumulative incidences*) (риски) и относительная инцидентность (*incidence rates*); при сравнении групп, подвергшихся и не подвергавшихся воздействию, на основании упомянутых показателей определяется отношение рисков (*risk ratio*) или отношение частот (*rate ratio*). В открытых когортах оценивается относительная инцидентность и отношение относительных инцидентностей (*rate ratios*).

В исследованиях «случай–контроль» авторы сравнивают эффекты воздействия между участниками с определённым исходом заболевания («случаи») и участников без такого исхода («контроли»). Авторы стремятся отбирать «случаи» и «контроли», которые репрезентативны когорте, из которой их набирают, или происходили

из поперечного «среза» популяции. Такую популяцию можно определить географически, но в более широком смысле — это территория, обслуживаемая медицинскими учреждениями. Выборка «случаев» может составлять всю или большую часть имеющихся «случаев», в то время как выборка «контролей» обычно составляет лишь небольшую часть лиц, не имеющих соответствующего исхода. «Контроли» представляют когорту или популяцию, в которой могут появляться «случаи». Исследователи рассчитывают отношение шансов воздействия предполагаемых причин заболевания между «случаями» и «контролями» (см. вставку 7). В зависимости от стратегии формирования групп «случаев» и «контролей», а также от особенностей изучаемой популяции отношение шансов (*odds ratio*) в исследованиях «случай–контроль» интерпретируется как отношение рисков, отношение частот (*rate ratio*) или отношение шансов для распространённости (*prevalence*) [16, 17]. Большинство опубликованных исследований «случай–контроль» представляют собой открытые когорты и, таким образом, позволяют получать прямые оценки отношения частот.

В **одномоментных исследованиях** исследователи изучают всех участников выборки в один момент времени, часто для того, чтобы изучить распространённость воздействия, факторов риска или болезни. Некоторые одномоментные исследования являются аналитическими и нацелены на количественную оценку потенциальных причинно-следственных связей между воздействием и заболеванием. Такие исследования можно анализировать как когортное исследование, сравнивая показатели распространённости заболевания между группами воздействия, или как исследование «случай–контроль», сравнивая шансы воздействия в группах с заболеванием и без него. Сложность, которая может возникнуть в исследовании любого дизайна, и которая особенно характерна для одномоментных исследований, состоит в том, чтобы установить, что воздействие предшествовало заболеванию, хотя временная последовательность воздействия и исхода иногда может быть вполне очевидной. Например, в исследовании, где переменная воздействия является врождённой или генетической характеристикой, можно быть уверенным, что воздействие предшествовало исходу, даже если мы определяем то и другое одновременно.

РАЗДЕЛЫ

Название и аннотация

1А. Укажите в названии статьи или в аннотации наименование дизайна исследования, используя общепринятые термины.

Пример

«Заболеваемость лейкемией среди работников обувной промышленности: исследование «случай–контроль» [18].

Пояснение

Читатели должны иметь возможность легко идентифицировать дизайн исследования по названию статьи или аннотации. Использование общепринятых терминов способствует правильной индексации статей в электронных базах данных [19, 20].

1В. Включите в аннотацию информативное и сбалансированное обобщение того, что было сделано и обнаружено.

Пример

«**Актуальность:** ожидаемая выживаемость ВИЧ-инфицированных пациентов представляет большой интерес для общественного здравоохранения.

Цель — оценить время выживания и по возрасту смертность в когорте ВИЧ-инфицированных пациентов в сравнении с общей популяцией.

Дизайн: популяционное когортное исследование.

Условия проведения: все ВИЧ-инфицированные лица, получающие медицинскую помощь в Дании в период с 1995 по 2005 год.

Пациенты: каждому участнику Датского общенационального когортного исследования ВИЧ-инфицированных подбирали для сравнения по полу, дате рождения и месту проживания до 99 лиц из общей популяции.

Измерения. Авторы рассчитывали таблицы дожития Каплана–Мейера с возрастом в качестве временной шкалы для оценки выживаемости после 25-летнего возраста. ВИЧ-инфицированные пациенты и соответствующие им лица из общей популяции наблюдались от момента установления диагноза и до смерти, эмиграции или наступления 1 мая 2005 года.

Результаты. В исследование были включены 3990 ВИЧ-инфицированных и 379 872 лица из общей популяции. Суммарное время наблюдения составило 22 744 человеко-года (медиана 5,8 лет) и 2 689 287 человеко-лет в группе контроля (медиана 8,4 лет) соответственно. Исходы 3% участников не были отслежены. После достижения возраста 25 лет медиана выживаемости составила 19,9 лет (95% ДИ; 18,5–21,3) в группе ВИЧ-инфицированных и 51,1 года (ДИ; 50,9–51,5) в выборке из общей популяции. В группе ВИЧ-инфицированных выживаемость возросла до 32,5 лет (ДИ; 29,4–34,7) за период с 2000 до 2005 год. При исключении лиц с гепатитом С (16%) медиана выживаемости за тот же период составила 38,9 года (ДИ; 35,4–40,1). Относительный риск смерти в когорте ВИЧ-инфицированных по сравнению с таковым в выборке из общей популяции снижался с возрастом, тогда как смертность с увеличением возраста увеличивалась.

Ограничения: наблюдаемая смертность признаётся допустимой для анализа, если текущая максимальная длительность наблюдения превышает 10 лет.

Выводы. В эпоху высокоактивной антиретровирусной терапии оценка медианы выживаемости для молодых ВИЧ-инфицированных лиц составляет более 35 лет. Однако для дальнейшего снижения смертности ВИЧ-инфицированных в сравнении с общей популяцией по-прежнему необходимо прикладывать постоянные усилия» [21].

Пояснение

Аннотация содержит ключевую информацию, которая позволяет читателям понять исследование и принять решение о том, читать ли статью. Типичные компоненты аннотации — формулировка исследовательского вопроса, краткое описание методов и результатов, выводы [22]. В аннотации необходимо резюмировать ключевые детали исследования и включать только ту информацию, которая представлена в статье. Рекомендуется представлять ключевые результаты в цифровой форме: количество участников, оценка связей, изменчивости (*variability*) и неопределённости (*uncertainty*) (например, отношения шансов с доверительными интервалами). Мы считаем недостаточной простую констатацию наличия или отсутствия статистически значимой связи между воздействием и исходом.

Серия заголовков, относящихся к актуальности, дизайну, проведению и анализу данных исследования, поможет читателям быстро найти интересующую их информацию [23]. Во многих журналах требуют представлять структурированные аннотации, которые, как правило, имеют более высокое качество и более информативны, чем неструктурированные резюме [24, 25].

Введение

В разделе «Введение» следует описать, почему было проведено исследование, и какие вопросы и гипотезы оно затрагивает. Это позволит другим понять контекст исследования и оценить его потенциальный вклад в современные знания.

2. Актуальность/обоснование. Объясните актуальность и обоснуйте необходимость проведения представляемого научного исследования.

Пример

«Рост распространённости ожирения среди детей и подростков определил интерес к получению качественных данных о связи ожирения у детей с повышенным риском сердечно-сосудистых заболеваний и смерти во взрослом возрасте. Ожирение приводит к серьёзным социальным и психологическим проблемам в детском и подростковом возрасте, однако немного известно о том, какие социальные, социально-экономические и психологические последствия оно вызывает в последующем. Недавний систематический

обзор показал, что длительные когортные исследования последствий детского ожирения, помимо последствий для физического здоровья, не проводились. Было проведено лишь два исследования социально-экономических последствий ожирения у подростков. Gortmaker и соавт. обнаружили, что жительницы США с ожирением в позднем подростковом возрасте в 1981 г., семь лет спустя реже были замужем и имели более низкий доход, чем женщины, не страдавшие ожирением в подростковом возрасте, в то время как мужчины с избыточным весом реже вступали в брак. Sargent и соавт. выяснили, что женщины в Великобритании (но не мужчины), страдавшие ожирением в возрасте 16 лет в 1974 г., в возрасте 23 лет зарабатывали на 7,4% меньше, чем их сверстники, не страдавшие ожирением. <...> Мы использовали данные длительного наблюдения за британской когортой 1970 года рождения для изучения социально-экономических, образовательных, социальных и психологических исходов детского ожирения во взрослом возрасте» [26].

Пояснение

Научные предпосылки исследования дают читателям необходимое представление о контексте, раскрывая суть и направление исследования, обеспечивая общее представление о том, что известно по данному вопросу, и каковы пробелы в современных знаниях. В этом разделе можно ссылаться на недавно проведённые исследования по схожей проблематике и систематические обзоры таких исследований.

3. Цели. Укажите цели исследования, включая все предварительные гипотезы.

Пример

«Нашими основными целями были: 1) определить распространённость домашнего насилия среди пациентов, обратившихся в 4 общественных центра первичной медицинской помощи, которые обслуживают пациентов с разным социально-экономическим статусом, и 2) выявить демографические и клинические различия между пациентами, которые подвергаются и не подвергаются насилию в настоящее время» [27].

Пояснение

Цели — это подробное изложение желаемого результата исследования. Тщательно продуманные цели помогают точно определить популяции, воздействия, исходы, а также другие параметры, которые будут оцениваться. Они могут быть сформулированы в форме конкретных гипотез или как вопросы, для ответа на которые и было спланировано исследование. В некоторых случаях цели могут быть менее чёткими, например, на ранних фазах научного исследования. Тем не менее в рукописи должны быть чётко отражены намерения авторов. Например, если важный анализ в подгруппах или любой другой дополнительный анализ не был изначально целью исследования, но необходимость этого возникла в ходе изучения данных, это должно быть

описано соответствующим образом (см. также пункты 4; 17 и 20).

Методы

В разделе «Методы» рекомендуется подробно описать, что было запланировано и что было сделано, чтобы другие могли понять основные аспекты исследования; судить о том, являются ли использованные методы приемлемыми для получения надёжных (*reliable*) и достоверных (*valid*) данных, а также оценить обоснованность каких-либо отклонений от первоначального плана.

4. Дизайн исследования. Укажите основные элементы дизайна исследования.

Пример

«Мы использовали перекрёстное сравнение случаев (*case-crossover design*), уместное в ситуациях, когда кратковременное воздействие фактора (использование телефона водителем) вызывает преходящее повышение риска редкого исхода (авария). Мы сравнивали использование мобильного телефона водителем в момент аварии и в другие подходящие для сравнения периоды. Поскольку водители выступали и в роли собственных «контролей», в исследовании учитывали их характеристики, которые могли повлиять на риск аварии и вместе с тем были неизменными в течение короткого периода времени. Ввиду того что риски в контрольные периоды и во время аварийных поездок должны быть схожими, мы сравнивали активность использования телефона в интервале риска (время непосредственно перед аварией) и в контрольные интервалы (отрезки времени, когда участники управляли автомобилем, но не попали в аварию) в течение предыдущей недели» [28].

Пояснение

Мы советуем представлять основные элементы дизайна исследования в начале раздела «Методы» или в конце «Введения», чтобы читатели могли понять схему исследования (*study design*). Например, необходимо указать, что авторы проводили когортное исследование, где велось наблюдение за участниками в течение определённого времени, описать группу лиц, составляющих когорту, указать статус оказываемого на них воздействия. Аналогичным образом, если использован дизайн исследования «случай–контроль», должны быть описаны «случаи», «контроли» и исходная популяция (*source population*). Если проводилось одномоментное обследование, следует описать популяцию и временную точку, в которой производили поперечный «срез». Если исследование является промежуточным вариантом трёх основных дизайнов, необходимо внести дополнительную ясность. Например, в примере выше мы кратко описали принципы перекрёстного сравнения случаев, одного из вариантов исследования «случай–контроль» [28].

Мы рекомендуем воздержаться от простого обозначения исследования как «проспективное» или «ретроспективное», поскольку эти термины не имеют чётких

определений [29]. Одни используют термины «проспективный» и «когортный» как синонимы, и оставляют слово «ретроспективный» для исследований «случай–контроль» [30]. Другие различают проспективные и ретроспективные когортные исследования в зависимости от периода сбора данных по отношению к тому времени, когда возникла идея исследования [31]. Третьи различают проспективные и ретроспективные исследования «случай–контроль» в зависимости от того, имелась ли информация об исследуемом воздействии в тот момент, когда отбирали подходящие «случаи» [32]. Некоторые советуют не использовать эти термины [33] или употреблять альтернативные варианты для описания когортных исследований — «параллельные» и «исторические» [34]. В рекомендациях STROBE мы не используем термины «проспективный» и «ретроспективный», впрочем, так же как и их альтернативы — «параллельный» и «исторический». Авторам рекомендуется всякий раз при использовании данной терминологии объяснять, что конкретно имеется в виду, а самое главное — точно описывать, как и когда происходил сбор данных.

В первой части раздела «Методы» можно также указать, является ли рукопись частью серии публикаций по результатам исследования. Если новая рукопись соответствует первоначальным целям этого исследования, на это обычно указывает ссылка на более раннюю публикацию и краткое повторение её основных особенностей. Однако цели исследования могут меняться со временем.

Авторы часто используют данные в целях, для которых они изначально не предназначались, включая в исследование, например, официальную статистику естественного движения населения, которую первоначально собирали в административных целях; анкетные данные, которые изначально были включены лишь для полноты информации, или образцы крови, которые собирали в связи с другими причинами. Например, исследование Physicians' Health Study (рандомизированное контролируемое исследование аспирина и каротина) впоследствии было использовано для демонстрации того, что точечная мутация гена фактора V ассоциирована с высоким риском венозного тромбоза, но не инфаркта миокарда или инсульта [35]. Вторичное использование имеющихся данных — творческая часть наблюдательного исследования, которое не обязательно делает полученные результаты менее убедительными или менее важными. Тем не менее краткое повторение первоначальных целей поможет читателям понять контекст исследования и возможные ограничения в использовании данных.

5. Условия проведения. Опишите условия, место и соответствующие даты проведения исследования, включая периоды привлечения участников, воздействия, наблюдения и сбора данных.

Пример

«В исследование Pasitos Cohort Study включали беременных женщин, наблюдавшихся в клиниках «Мать,

младенец и дитя» в городах Сокорро и Сан-Элизарио округа Эль-Пасо (штат Техас, США), а также в клинике «Мать и дитя» Мексиканского института социального обеспечения в городе Сьюдад-Хуарес (Мексика) с апреля 1998 г. по октябрь 2000 г. Изначально, до рождения детей, женщин, включённых в когорту, расспрашивали о домашней обстановке. В этом продолжающемся когортном исследовании наблюдения проводили с полугодовой периодичностью, начиная с достижения новорождёнными возраста 6 месяцев» [36].

Пояснение

Читателям необходимо знать условия и место проведения исследования, чтобы оценить контекст и обобщаемость результатов исследования. Такие воздействия, как средовые факторы и терапевтические вмешательства, могут со временем меняться, как и методы исследования. Информация о времени проведения исследования, наборе участников и наблюдении за ними, помещает исследование в исторический контекст, что важно для интерпретации результатов.

Информация об условиях проведения включает место набора пациентов или источник информации о них (например, список избирателей, поликлиника, система регистрации онкологических заболеваний или центр третичной медицинской помощи). К информации о месте проведения исследования относятся страна, город, госпиталь или медицинская практика, где проводилось исследование. Мы советуем указывать даты, а не только длительность периодов времени. Могут быть разные даты для воздействия, возникновения заболевания, набора участников, начала и конца периода наблюдения, сбора данных. Следует отметить, что почти 80% из 132 статей, опубликованных в онкологических медицинских журналах, в которых были представлены результаты анализа выживаемости, включали даты начала и окончания набора пациентов, и лишь в 24% статей сообщали дату окончания наблюдения [37].

6. Участники

6А. Когортное исследование. Укажите критерии, источники и методы отбора участников. Опишите методику наблюдения.

Пример

«Участницы исследования Iowa Women's Health Study были случайной выборкой из числа всех женщин в возрасте 55–69 лет, получивших автомобильные права в штате Айова в 1985 г., что составило примерно 94% всех женщин штата Айова этой возрастной группы. <...> Вопросники для определения жизненного статуса и изменения адреса проживания рассылали по почте в октябре 1987 г. и августе 1989 г. <...> Заболеваемость раком (за исключением немеланомного рака кожи) устанавливали по данным Регистра здравоохранения штата Айова <...>. Для участниц Iowa Women's Health Study из регистра подбирали пары, сопоставимые по имени,

настоящей фамилии и девичьей фамилии, почтовому коду и дате рождения» [38].

6А. Исследование «случай–контроль». Укажите критерии отбора, источники и методы установления «случаев» и «контролей». Обоснуйте выбор «случаев» и «контролей».

Пример

«Случаи меланомы кожи, диагностированные в 1999–2000 гг., выявляли по данным Регистра рака штата Айова <...>. Контролем, по данным этого же регистра, были случаи колоректального рака, диагностированные в тот же период. Пациентов с колоректальным раком использовали в качестве «контролей», поскольку это распространённое заболевание с высокой выживаемостью, а воздействие мышьяка не связано с развитием этого заболевания» [39].

6А. Одномоментное исследование. Укажите критерии, источники и методы отбора участников.

Пример

«Ретроспективно выявляли пациентов с основным диагнозом «инфаркт миокарда» (код 410) в соответствии с МКБ-9 по кодам диагнозов при выписке, за исключением кодов с пятой цифрой «2», которые обозначают последующий эпизод лечения <...>. Из числа всех пациентов с диагнозом «инфаркт миокарда», получавших медицинскую помощь по программе Medicare с февраля 1994 г. по июль 1995 г., была сформирована случайная выборка <...> Критериями включения были госпитализация более чем через 30 мин, но менее чем через 12 ч после появления болей в груди, с подъёмом сегмента ST не менее 1 мм в двух последовательных отведениях на электрокардиограмме при поступлении» [40].

Пояснение

Детальное описание участников исследования помогает читателям понять применимость (*applicability*) результатов. Исследователи обычно ограничивают целевую популяцию (*study population*) с помощью ряда клинических, демографических и других характеристик. Типичные критерии отбора (*eligibility criteria*) — возраст, пол, диагноз, сопутствующие заболевания. Несмотря на их важность, критерии отбора часто не описывают должным образом. Так, критерии отбора участников не были указаны в 17 из 49 статей (35%) с результатами наблюдательных исследований инсульта [5].

Критерии отбора в исследовании могут быть описаны как критерии включения (*inclusion criteria*) и невключения (*exclusion criteria*), хотя такое разделение не всегда необходимо и полезно. Мы советуем авторам описать все критерии отбора, а также указать, из какой группы отбирали участников для исследования (например, общая популяция жителей региона или страны), и описать метод отбора (например, направление участников клиницистами или их самостоятельная регистрация по объявлению о наборе в исследование).

Знание деталей процедуры наблюдения (действия по минимизации доли не ответивших или выбывших из-под наблюдения участников, одинаковость процедуры наблюдения для всех участников) позволяет сделать выводы о достоверности (*validity*) результатов. Например, если в исследовании для выявления острой инфекции использовали антитела IgM, необходимо указать интервал между их определением в крови, что позволяет читателям оценить, не были ли некоторые случаи инфекции пропущены из-за слишком длительного интервала между анализами [41]. В исследованиях, где процедуры наблюдения различались в группах, подвергшихся и не подвергавшихся воздействию, читатели могут распознать существенную систематическую ошибку, возникшую из-за различий в установлении событий или разницы в числе неответивших или выбывших из-под наблюдения участников [42]. Именно поэтому мы советуем описывать методы наблюдения за участниками и указывать, были ли они одинаковы для всех участников, и позволяют ли они описать изучаемые переменные достаточно полно (см. пункт 14).

В исследованиях «случай–контроль» выбор «случаев» и «контролей» имеет решающее значение для интерпретации результатов, а метод их отбора имеет большое значение для определения достоверности исследования (*study validity*). В целом «контроли» должны отражать популяцию, из которой происходят «случаи». Для подбора «контролей» используют различные методы, каждый из которых имеет свои преимущества и недостатки. Для «случаев», происходящих из общей популяции, — это выборка из реестра популяции, случайный обзвон, привлечение соседей или друзей. «Контроли», подобранные из соседей или друзей, могут помочь лучше оценить воздействие, поскольку позволяют скорректировать влияние других факторов риска, общих для «случаев» и «контролей» [17]. Подбор «контролей» из лиц, страдающих другими заболеваниями, может иметь преимущества перед отбором из популяции, в частности для «случаев», связанных с госпитализацией, поскольку он проще в исполнении, лучше отражает госпитальную популяцию, отличается сопоставимым откликом «случаев» и «контролей» при включении в исследование. Однако здесь могут возникать проблемы, если изучаемое воздействие повышает риск развития заболевания или связанной с ним госпитализации в группе «контролей» [43, 44]. Для решения этой проблемы часто используют несколько наиболее подходящих для сравнения болезней [45].

6В. Когортное исследование. Для исследований с подобранными группами опишите критерии подбора и количество участников, подвергшихся и не подвергавшихся воздействию изучаемого фактора.

Пример

«Для каждого пациента, изначально получавшего статин, в качестве контроля подбирали пару — участника,

не получавшего статин, в соответствии со следующим протоколом. Сначала для каждого пациента в общей когорте на основании обширного списка факторов рассчитывали показатель вероятности использования статинов или риска возникновения сепсиса. На втором этапе каждому пациенту, получающему статин, подбирали небольшой пул из не получающих статин и сопоставимых по полу, возрасту (плюс-минус 1 год) и указателю даты (плюс-минус 3 месяца). На третьем этапе из этого пула выбирали «контроль», наиболее близкий по показателю вероятности (в пределах 0,2 величины стандартного отклонения) для каждого «случая» в соотношении 1:1. Контроли, оставшиеся без пары, исключали из исследования» [46].

6В. Исследования «случай–контроль». Для исследований с подобранными группами опишите критерии подбора и количество «контролей» на один «случай».

Пример

«Мы планировали отобрать по 5 «контролей» для каждого «случая» из числа индивидуумов исследуемой популяции, не имеющих записи о диагнозе «аутизм» или других расстройств аутистического спектра (РАС) в медицинской карте врача общей практики, живых на момент проведения исследования и зарегистрированных в том же учреждении, что и «случаи» с диагнозом РАС. «Контроли» подбирали к «случаям» индивидуально — по году рождения (не более чем на год младше или старше), полу и общей практике. Для каждого из 300 «случаев» подбирали 5 «контролей», отвечающих всем критериям подбора. Оставшиеся 994 «контроля» были исключены из исследования...» [47].

Пояснение

Подбор пар (*matching*) наиболее характерен для исследований «случай–контроль», но порой его используют и в когортных исследованиях для формирования групп, сопоставимых в начале наблюдения. Использование метода подбора пар в когортных исследованиях позволяет получить группы, напрямую сопоставимые по потенциальным вмешивающимся факторам (*confounders*), и вызывает меньше сложностей, чем в исследованиях «случай–контроль». Например, нет необходимости учитывать переменные, использованные для подбора пар, при оценке относительного риска [48]. Поскольку применение метода подбора пар в когортном исследовании повышает статистическую точность, исследователи могут использовать его при анализе данных и таким образом получить более узкие доверительные интервалы.

В исследованиях «случай–контроль» подбор пар проводят для повышения эффективности исследования (*study's efficiency*), поскольку это помогает добиться схожего распределения переменных в группах «случаи» и «контроли», в частности распределения потенциальных вмешивающихся факторов [48, 49]. Поскольку подбор пар может быть выполнен различными способами,

с одним или более «контролями» для каждого «случая», следует обосновать выбор переменных для сопоставления и описать детали используемого метода. Часто используемыми методами подбора пар являются частотное сопоставление (также называемое групповым сопоставлением; *frequenc/group matching*) и индивидуальное сопоставление (*individual matching*). В случае частотного сопоставления «контроли» выбирают так, чтобы распределение переменных, используемых для подбора пар, стало идентичным или подобным их распределению в группе «случаев». Индивидуальное сопоставление означает поиск одного или нескольких «контролей» для каждого «случая». Хотя метод подбора пар интуитивно кажется привлекательным, а иногда — полезным, сопоставление

в исследованиях «случай–контроль» имеет множество недостатков и не всегда уместен, что необходимо учитывать при анализе (вставка 2).

Даже очень простая процедура подбора пар может быть плохо описана. Например, авторы могут утверждать, что «контроли» подбирали к «случаям» «в пределах 5 лет» или используя «5-летний диапазон». Означает ли это, что для «случая» в возрасте 54 лет возраст «контроля» должен быть от 50 до 54 лет? Или от 49 до 59? А если используется более широкий временной промежуток (например, 10 лет), возникает опасность остаточного влияния (*residual confounding*) возраста как вмешивающегося фактора (см. вставку 4), например, потому что «контроли» будут в среднем моложе, чем «случаи».

Вставка 2. Подбор пар в исследованиях «случай–контроль»

При проведении исследований «случай–контроль» необходимо продумать, сопоставлять ли «контроли» со «случаями», и если да, то по каким переменным, насколько точный метод сопоставления использовать, и какой метод статистического анализа применять. Отказ от подбора пар может привести к тому, что распределение некоторых ключевых вмешивающихся факторов (например, возраст, пол) будет радикально отличаться в группах «случаи» и «контроли». Хотя эти различия можно скорректировать в ходе анализа, статистическая эффективность исследования может значительно снизиться.

Использование процедуры подбора пар в исследованиях «случай–контроль» и интерпретация её результатов сопряжены с трудностями, особенно если подбор пар проводится по нескольким факторам риска, некоторые из которых могут быть связаны с воздействием, представляющим наибольший интерес [50, 51]. Например, при изучении влияния оральных контрацептивов на возникновение инфаркта миокарда, исходя из фармако-эпидемиологической базы данных с информацией о тысячах женщин, которые могут быть потенциальными «контролями», у исследователей может возникнуть соблазн выбрать «контроли» со значениями факторов риска, схожими с таковыми в каждом случае инфаркта миокарда. Одна из задач — учесть факторы, которые могут повлиять на назначение оральных контрацептивов, и таким образом устранить *искажение показателя (confounding by indication*; здесь: искажение оценки влияния лечения на исход болезни при наличии показателя или противопоказания к данному лечению, которое также связано с исходом). Однако в результате будет сформирована *нерепрезентативная* (для исходной популяции женщин, получающих оральные контрацептивы) контрольная группа, поскольку «контроли» будут старше лиц в исходной популяции (*source population*) по причине того, что пациенты с инфарктом миокарда, как правило, старшего возраста. Это может иметь ряд последствий. Простой анализ данных выявит отношения шансов, которые обычно смещены в сторону единицы, если фактор, использованный для подбора пар, связан с воздействием. Выход из данной ситуации — сопоставление «случаев» и «контролей» или проведение стратифицированного анализа (см. пункт 12D). Кроме того, поскольку подобранная контрольная группа перестаёт быть репрезентативной для популяции в целом, распределение фактора воздействия среди «контролей» больше не может использоваться для оценки дополнительного популяционного риска (*population attributable fraction*) (см. вставку 7) [52]. К тому же влияние фактора, используемого для подбора пар, уже не может быть изучено, а поиск хорошо подобранных «контролей» может быть обременительным. В таких случаях дизайн с использованием контрольной группы как есть (*прим. ред.*: сформированной без применения процедур подбора пар) более предпочтителен, поскольку получить такие «контроли» легче, а сама контрольная группа будет больше. *Избыточный подбор (overmatching)* — ещё одна проблема, которая может снизить эффективность исследований «случай–контроль», в которых применяется процедура подбора пар, а в некоторых ситуациях стать причиной систематических ошибок. Информация теряется, а эффективность исследования снижается, если переменная сопоставления тесно связана с фактором воздействия. В результате, многие участники в сопоставляемых группах будут, как правило, иметь одинаковые или схожие уровни воздействия, и, следовательно, при сравнении не будет получена необходимая информация. Подбор пар способен внести некорректируемую систематическую ошибку в том случае, если переменная, применяемая для подбора пар, не будет вмешивающимся фактором, но будет определять причинно-следственную связь между воздействием и заболеванием. Например, экстракорпоральное оплодотворение связано с повышенным риском перинатальной смерти вследствие

увеличения числа многоплодных родов и рождения младенцев с низким весом [53]. Сопоставление по многоплодию или весу при рождении приведёт к смещению результатов в сторону нуля, и такие данные нельзя будет скорректировать в ходе анализа.

Хотя интуитивно подбор пар кажется привлекательным методом формирования сравниваемых групп, методологи не рекомендуют прибегать к подбору пар в исследованиях «случай–контроль» без должного на то обоснования. Они советуют тщательно и продуманно рассмотреть каждый потенциальный фактор сопоставления, с пониманием того, что он может быть измерен и использован для статистической коррекции, а не с целью подбора пар. В результате можно добиться сокращения числа используемых факторов сопоставления, более широкого использования частотного сопоставления, что позволяет избежать некоторых проблем, рассмотренных выше, а также проведение большего числа исследований «случай–контроль» без какого-либо сопоставления вообще [54]. Подбор пар остаётся наиболее предпочтительным или даже необходимым методом, если распределение вмешивающегося фактора (например, возраста) в сравниваемых группах существенно отличается [48, 49].

7. Переменные. Чётко определите все исходы, воздействия, прогностические факторы, потенциальные вмешивающиеся факторы, модификаторы эффекта. Опишите диагностические критерии, если применимо.

Пример

«В анализ включали только случаи серьёзных врождённых пороков развития. Незначительные аномалии, определяемые согласно списку исключений Европейского регистра врождённых аномалий (EUROCAT), исключали из рассмотрения. Если у ребёнка имелось более 1 врождённого порока развития в одной системе органов, эти пороки считали как 1 исход при анализе по системам органов. <...> При статистическом анализе возраст беременной в момент родов и число предыдущих родов рассматривали в качестве потенциально вмешивающихся факторов. Возраст матери в момент отмены противоэpileптических средств и в момент родов рассматривали как потенциальные модификаторы эффекта» [55].

Пояснение

Авторы должны дать определение всем переменным, рассматриваемым и включённым в анализ, в том числе исходам, воздействиям, факторам риска, потенциальным вмешивающимся факторам и модификаторам эффекта (факторы, изменяющие эффект изучаемого воздействия). Исходы заболевания необходимо достаточно подробно описать в соответствии с диагностическими критериями. Это относится к критериям отбора «случаев» в исследовании «случай–контроль»; к случаям заболевания, диагностированным во время наблюдения в когортных исследованиях; к заболеваниям, изучаемым в одномоментных исследованиях. Чёткие определения и шаги, предпринятые для их соблюдения, особенно важны в случае заболевания, представляющего первичный интерес для исследования.

Для некоторых исследований уместными терминами для описания воздействия могут быть «детерминанта» или «предиктор», а для описания исходов — «конечные точки» (*endpoints*). В многомерных моделях авторы иногда используют понятие «зависимые переменные» (*dependent variable*) для описания исходов и «независимые»

(*independent variable*) или «объясняющие переменные» (*explanatory variable*) для воздействия и вмешивающихся факторов. Последнее не является точным, поскольку не позволяет отличить воздействие от вмешивающегося фактора.

Если измеряется большое количество переменных, которые затем включаются в предварительный анализ на ранней фазе исследования, обдумайте возможность создания списка с детальным описанием каждой переменной в приложении, в дополнительной таблице или отдельной публикации. Следует отметить, что в журнале *International Journal of Epidemiology* недавно ввели новый раздел «Профили когорты», который включает подробную информацию о том, какие измерения проводились в различные моменты времени в конкретных исследованиях [56, 57]. Мы советуем авторам перечислять все переменные, рассматриваемые в ходе статистического анализа, а не описывать только те переменные, которые вошли в окончательную модель (см. пункт 16А) [58, 59].

8. Источники данных, измерения. Для каждой переменной, представляющей интерес, приведите источники данных и подробные сведения о методах оценки (измерения). Укажите, были ли сопоставимы методы измерения в группах.

Пример 1

«Общее потребление кофеина рассчитали, используя в основном данные Министерства земледелия США о составе продуктов. В соответствии с этими расчётами предполагалось, что в чашке кофе содержится 137 мг кофеина, в чашке чая — 47 мг, в банке или бутылке газированного напитка на основе ореха колы — 46 мг, в порции шоколадных конфет — 7 мг. Этот метод измерения потребления кофеина был признан достоверным как при изучении когорты NHS I, так и в похожем исследовании когорты мужчин — медицинских работников <...>. Выявление артериальной гипертензии на основании сообщаемой пациентом информации было признано надёжным методом установления заболевания согласно данным когорты NHS I» [60].

Пример 2

«Образцы для анализов, полученные от подобранных в пары «случаев» и «контролей», всегда исследовали вместе одной партией, и персонал лаборатории не имел возможности отличить образцы пациентов группы «случаев» от группы «контролей» [61].

Пояснение

Способ измерения воздействия, вмешивающихся факторов и исходов влияет на надёжность (*reliability*) и достоверность (*validity*) исследования. Ошибки измерения и классификации воздействия или исходов могут затруднить выявление причинно-следственных связей или привести к появлению ложных связей. Ошибки измерения потенциальных вмешивающихся факторов могут увеличить риск остаточного (сохраняющегося после коррекции) искажения (*residual confounding*) [62, 63]. Поэтому если авторы сообщают о результатах любых исследований достоверности или надёжности оценок или измерений, полезно включить в описание характеристики референсных стандартов, если таковые использовались. Вместо того чтобы просто процитировать исследование, установившее достоверность методики (как это сделано в первом примере), мы советуем авторам представить оценку достоверности и надёжности, что может в дальнейшем использоваться для корректировки ошибок измерения или при анализе устойчивости результатов исследования к изменениям исходных условий (*sensitivity analyses*) (см. пункт 12Е и 17).

Кроме того, важно знать, различались ли сравниваемые группы по способу сбора данных. Это может быть существенно при лабораторной оценке (как во втором примере) и в других ситуациях. Например, если при интервьюировании опрашивать сначала все «случаи», а потом все «контроли» (или наоборот), возможно возникновение систематической ошибки в результате обучения интервьюеров (*learning curve*). Решение этой проблемы — рандомизация порядка опроса. Риск систематической ошибки информации возрастает, если в сравниваемых группах проводили неодинаковые диагностические тесты или в одной группе проводили больше тестов того же типа, что и в другой (см. пункт 9).

9. Систематические ошибки. Опишите все усилия по устранению потенциальных источников систематических ошибок.

Пример 1

«В большинстве исследований «случай–контроль», посвящённых изучению суицида, в контрольные группы включали живых участников, мы же решили сформировать контрольную группу из умерших вследствие других причин <...>. В качестве источников информации о факторах риска в контрольной группе выступали лица, недавно потерявшие члена семьи или близкого человека. Эти источники более сопоставимы с источниками информации, используемыми при изучении суицида, чем контрольная группа из живых лиц» [64].

Пример 2

«Систематическая ошибка выявления может влиять на связь между сахарным диабетом 2-го типа (СД2) и первичной открытоугольной глаукомой (ПОГ), если женщины с СД2 находятся под более пристальным контролем офтальмолога, чем женщины без такового диагноза. Мы сравнили среднее число обследований зрения у женщин, страдающих и не страдающих СД2. Также пересчитали относительный риск развития ПОГ с учётом ковариат, ассоциированных с более тщательным наблюдением за состоянием органов зрения (сообщения пациентов о признаках катаракты, макулярной дегенерации, частота проверок зрения и медицинских осмотров)» [65].

Пояснение

Систематические ошибки (*bias*) в исследованиях приводят к результатам, отличающимся от истины (см. вставку 3). Читателям важно знать, какие меры предпринимались во время проведения исследования для уменьшения возможных систематических ошибок. В идеале исследователи должны тщательно оценить возможные источники систематических ошибок ещё на этапе планирования исследования. На стадии написания отчёта мы рекомендуем авторам всегда оценивать вероятность наличия соответствующих систематических ошибок. В частности, следует обсудить и по возможности оценить направление и величину смещения. Например, в исследованиях «случай–контроль» могут возникать систематические ошибки информации (измерения; *information bias*), но их риск можно уменьшить, тщательно отбирая подходящую контрольную группу (см. пример 1) [64]. Различия в медицинском наблюдении за участниками — проблема, описанная во втором примере [65]. Следовательно, авторы приводят более подробные сведения, описывают дополнительные данные, собранные для того, чтобы решить эту проблему. Если авторы использовали программы контроля качества на этапе сбора данных для предотвращения возможных отклонений (*drift*) при измерении переменных в длительных когортных исследованиях (*longitudinal studies*) или для того, чтобы свести к минимуму вариабельность при участии нескольких наблюдателей, это следует описать.

К сожалению, авторы зачастую не учитывают серьёзные систематические ошибки, когда представляют результаты исследований. Среди 43 исследований («случай–контроль» и когортные), опубликованных в период 1990–1994 гг., в которых изучался риск повторного возникновения рака у пациентов с онкологическими заболеваниями в анамнезе, систематические ошибки медицинского наблюдения (*medical surveillance bias*) упоминались лишь в 5 статьях [66]. Исследование отчётов об исследованиях психического здоровья, опубликованных в течение 1998 г. в трёх журналах по психиатрии, показало, что только в 13% статей из 392 сообщали о систематических ошибках отклика

(различия в характеристиках тех, кто вызвался принять участие в исследовании, и тех, кто отказался; *response bias*) [67]. Обзор когортных исследований, посвящённых изучению инсульта, показал, что авторы 14 из 49 статей (28%), опубликованных в период 1999–2003 гг.,

упоминали о потенциальной систематической ошибке отбора (*selection bias*) участников исследования, а в 35 публикациях (71%) допускалась возможность того, что любой тип систематической ошибки мог повлиять на результаты [5].

Вставка 3. Систематическая ошибка

Систематическая ошибка (смещение; *bias*) — это одностороннее отклонение результата исследования от истинного значения. Как правило, такие смещения возникают в результате ошибок планирования или проведения исследования, и в последующем не могут быть исправлены. Систематическая ошибка и конфаундинг (влияние вмешивающихся факторов; *confounding*) не являются синонимами. Систематические ошибки возникают в результате использования неверной информации или ошибок в отборе участников исследования, что приводит к ложным ассоциациям. Конфаундинг порождает связи, которые действительно существуют, но не могут быть интерпретированы как причинно-следственные, поскольку в основе этих связей — некоторые неучтённые факторы, ассоциированные и с фактором воздействия, и с анализируемым исходом (см. вставку 5). Кроме того, следует отличать систематическую ошибку от случайной ошибки (*random error*) — отклонение от истинного значения, вызванное статистическими колебаниями (в любом направлении) в результатах измерений. Описаны многие возможные источники систематических ошибок, при этом используются разные термины [68, 69]. Мы предлагаем к рассмотрению две простые категории ошибок — ошибки информации (*information bias*) и ошибки отбора (*selection bias*).

Ошибка информации возникает в связи с систематическими различиями в полноте или точности данных, что приводит к неодинаковой в сравниваемых группах классификации (*differential misclassification*) индивидов в отношении факторов воздействия или исходов. Например, если женщины, страдающие диабетом, проходят более регулярные и тщательные офтальмологические обследования, выявление глаукомы будет более полным, чем у женщин, не страдающих диабетом (см. пункт 9) [65]. Пациенты, принимающие лекарство, вызывающее неспецифическую диспепсию, могут проходить гастроскопию чаще, и, как результат, в этой группе обнаруживается большее количество язв желудка, чем у пациентов, не принимающих такое лекарство, даже если этот препарат не способствует язвообразованию. Этот тип ошибки информации также называется «ошибкой выявления» (*detection bias*) или «ошибкой медицинского наблюдения» (*medical surveillance bias*). Один из способов оценить её влияние — измерить интенсивность медицинского наблюдения в различных исследуемых группах и внести поправки в статистический анализ. В исследованиях «случай–контроль» ошибка информации возникает, если «случаи» более или менее точно вспоминают прошлые воздействия по сравнению с «контролями», не страдающими таким заболеванием, или если они более или менее охотно сообщают о них (так называемая ошибка памяти; *recall bias*). «Ошибка интервьюера» (*interviewer bias*) может возникнуть, если интервьюеры осведомлены о гипотезе исследования и подсознательно или сознательно собирают данные выборочно [70]. В таких случаях полезно прибегать к процедуре «ослепления» (*blinding*) как участников исследования, так и исследователей.

Ошибка отбора может возникать в исследованиях «случай–контроль», если вероятность отбора «случаев» или «контролей» связана с изучаемым воздействием. Например, врач, набирающий участников для исследования тромбоза глубоких вен, может диагностировать это заболевание у пациентки, которая жалуется на боли в ногах и принимает оральные контрацептивы, а у женщины с аналогичными жалобами, не принимающей такие препараты, тромбоз глубоких вен может быть не диагностирован. Подобных ошибок можно избежать, подбирая «случаи» и «контроли» из числа тех, кто прошёл одинаковое диагностическое обследование [70].

Аналогичным образом, ошибки отбора могут возникать при использовании данных регистров пациентов. В частности, если известна возможная связь между воздействием и развитием заболевания, «случаи» могут с большей вероятностью быть внесены в регистр, если они подверглись воздействию предполагаемого этиологического фактора [72]. «Ошибка отклика» (*response bias*) — ещё один тип ошибки отбора, который возникает, если различия в характеристиках между теми, кто отвечает, и теми, кто отказывается от участия в исследовании, влияют на оценку распространённости (*prevalence*), инцидентности (*incidence*) и иногда связей. В целом ошибка отбора влияет на внутреннюю валидность (*internal validity*) исследования, тогда как обычные проблемы, которые могут возникнуть при наборе участников в исследование, влияют на внешнюю (*external validity*), а не на внутреннюю валидность исследования (см. пункт 21).

10. Объём исследования. Объясните, как был определён объём (размер) исследования.

Пример 1

«Количество случаев в регионе в период исследования определило объём выборки» [73].

Пример 2

«Исследование случаев послеродовой депрессии в регионе выявило, что распространённость этого состояния составляет 19,8%. Предположив, что депрессия возникает у 20% матерей, родивших детей с нормальной массой тела, при отношении шансов равном 3 для матерей маловесных детей, пришли к выводу о необходимости набора 72 участников в исследование «случай–контроль» (1 «контроль» на 1 «случай») для достижения мощности исследования 80% и статистической значимости 5%» [74].

Пояснение

Исследование должно быть достаточно крупным, чтобы получить точечную оценку с достаточно узким доверительным интервалом (*confidence interval*), и убедительно ответить на вопрос, поставленный авторами работы. Большой размер выборки необходим, чтобы отличить слабые связи от отсутствия таковых. Небольшие исследования часто дают ценную информацию, но широкие доверительные интервалы могут указывать на то, что эта информация вносит меньший вклад в современные знания по сравнению с исследованиями, предоставляющими оценки с более узкими доверительными интервалами. Кроме того, небольшие исследования, демонстрирующие «интересные» или «статистически значимые» связи, публикуются чаще, чем такого же размера исследования без каких-либо «значимых» результатов. Хотя такие работы могут сигнализировать о первых шагах к важному открытию, читатели должны быть осведомлены об их потенциальных недостатках.

Важность определения необходимого размера выборки (*sample size*) в наблюдательных исследованиях зависит от контекста. Если предпринимается анализ данных, которые уже доступны и были получены для других целей, основной вопрос — даст ли анализ данных результаты с достаточной статистической точностью (*statistical precision*), чтобы внести существенный вклад в литературу. Соображения относительно размера выборки в таких случаях не имеют принципиального значения. Формально, *предварительный* расчёт размера выборки может быть полезен, когда планируется новое исследование [75, 76]. Такие расчёты связаны с большей неопределённостью, чем предполагает одно число, которое обычно указывается. Например, оценки частоты наступления интересующего события или другие предположения, важные для расчётов, обычно неточны, если не сказать гипотетичны [77]. Точность, достигаемая при окончательном анализе, не может быть спрогнозирована заранее, так как может снижаться в результате включения вмешивающихся факторов в многомерный

анализ [78] вместе с точностью измерения ключевых переменных [79] и при исключении из анализа отдельных участников.

Лишь немногие эпидемиологические исследования содержат данные о размерах выборки [4, 5]. Мы поддерживаем авторов в их стремлении описать надлежащим образом расчёт размера выборки, если таковой проводили. В других ситуациях уместно будет указать соображения, определившие размер выборки (например, доступная выборка фиксированных размеров, как в *Примере 1*, упомянутом выше). Если наблюдательное исследование было прервано до получения статистически значимых результатов, читателям необходимо сообщить об этом. Не стоит запутывать читателей обоснованием размера исследования (*study size*), выполненного после его завершения, или ретроспективными расчётами его мощности [77]. С точки зрения читателя, доверительный интервал хорошо описывает достигнутую в итоге статистическую точность исследования. Необходимо понимать, что доверительные интервалы отражают лишь статистическую неопределённость, а не всю неопределённость, которая может присутствовать в исследовании (см. пункт 20).

11. Количественные переменные. Укажите, какие количественные переменные были использованы в анализе и, если применимо, какие категории (группы) на основании этих переменных были выбраны и почему.

Пример

«Пациенты с оценкой по шкале комы Глазго (ШКГ) менее 8 баллов считаются тяжело травмированными. Оценка 9 баллов и более говорит о менее опасных повреждениях мозга. Мы оценивали связь оценки по ШКГ и этим двум категориям с наступлением смерти пациентов в течение 12 месяцев после повреждения» [80].

Пояснение

Именно исследователи выбирают, как собирать и анализировать количественные данные о воздействии, модификаторах эффекта и вмешивающихся факторах. Например, они могут разбить область значений непрерывной переменной воздействия (*continuous exposure variable*) на несколько интервалов, создав тем самым новую, категориальную переменную (*categorical variable*) (см. вставку 4). Выбор способа группировки данных может оказать существенное влияние на последующий анализ [81, 82]. Советуем авторам объяснять, почему и как группировали количественные данные, в том числе указывать количество категорий, точки разделения (*cut-points*), среднее арифметическое значение или медиану для каждой из групп. Если данные описываются в форме таблицы, для каждой категории необходимо указать число «случаев», «контролей», лиц, подвергшихся воздействию (*persons at risk*), человеко-время для периода воздействия (*person-time at risk*). Таблицы не должны включать лишь оценки эффективности или результаты подбора модели.

Вставка 4. Группировка данных

Непрерывные данные могут быть разбиты на интервалы (*grouping*) по нескольким причинам [86]. При сборе данных, возможно, лучше использовать порядковую переменную (*ordinal variable*), чем искать искусственно созданную непрерывную меру воздействия, основанную на воспоминаниях за несколько лет. Категории также могут быть уместны в презентациях, например для представления всех переменных в одном стиле, или для демонстрации зависимости «доза–ответ».

Данные можно группировать для упрощения анализа, к примеру, чтобы избежать предположения о линейности. Однако при группировании данных теряется информация, следовательно, может снизиться статистическая мощность (*statistical power*) [87], особенно, если авторы прибегают к дихотомии [82, 85, 88]. Группировка значений вмешивающегося фактора может привести к появлению остаточного искажения (*residual confounding*), в результате которого влияние этой переменной останется нескорректированным (см. вставку 5) [62, 89]. Увеличение числа категорий может уменьшить потери статистической мощности и остаточное искажение, что особенно уместно в крупных исследованиях. В небольших исследованиях может использоваться несколько групп из-за ограниченного количества данных.

Исследователи могут определять точки разделения (*cut-points*) для выделения категорий на основе значений, часто используемых в диагностике и прогнозировании в силу их практичности, или обосновывая свой выбор статистически. Можно, например, выбрать равное количество участников в каждой группе с помощью квантилей [90]. С другой стороны, можно получить более полное представление о связи с исходом, выбрав группы с крайними значениями анализируемого показателя, при этом средняя группа(-ы) должна быть больше, чем крайние [91]. В исследованиях «случай–контроль» предпочтительна группировка данных исходя из значений показателя в контрольной группе, предназначенной для описания исходной популяции (*source population*). Если точки разделения на группы выбраны из нескольких альтернативных вариантов, но уже после завершения исследования, читателей следует проинформировать об этом. В частности, если точки разделения были выбраны, чтобы минимизировать *p*-значение, истинная сила связи (*strength of an association*) будет преувеличена [81].

При анализе переменных, полученных в результате группирования данных, важно учитывать их исходно непрерывную природу. Например, можно исследовать возможную тенденцию изменения риска в упорядоченных группах. Общий подход — изучать ранги групп как непрерывную переменную. Такая линейность будет приближаться к фактической линейной зависимости, если категории имеют одинаковый интервал (например, возрастные группы с интервалом 10 лет), но не иначе. D. Il'yasova и соавт. [92] рекомендуют представлять категориальные и непрерывные оценки эффекта со стандартными ошибками, чтобы облегчить метаанализ, равно как и представить ценную информацию о зависимости «доза–ответ». Один анализ может быть источником для другого, и ни один из них не свободен от допущений. Авторы часто игнорируют упорядочение категориальных переменных и рассматривают оценки (и *p*-значения) отдельно для каждой категории в сравнении с референсной категорией. Это может быть уместным при описании, но реальная тенденция к изменению риска по группам может остаться незамеченной. Если такая тенденция обнаружена, доверительный интервал для оценки зависимости одной переменной от другой (*confidence interval for a slope*) будет показателем значимости наблюдения.

11. Количественные переменные. Укажите, какие количественные переменные были использованы в анализе и, если применимо, какие категории (группы) на основании этих переменных были выбраны и почему.

Пример

«Пациенты с оценкой по шкале комы Глазго (ШКГ) менее 8 баллов считаются тяжело травмированными. Оценка 9 баллов и более говорит о менее опасных повреждениях мозга. Мы оценивали связь оценки по ШКГ и этим двум категориям с наступлением смерти пациентов в течение 12 месяцев после повреждения» [80].

Пояснение

Именно исследователи выбирают, как собирать и анализировать количественные данные о воздействии,

модификаторах эффекта и вмешивающихся факторах. Например, они могут разбить область значений непрерывной переменной воздействия (*continuous exposure variable*) на несколько интервалов, создав тем самым новую, категориальную переменную (*categorical variable*) (см. вставку 4). Выбор способа группировки данных может оказать существенное влияние на последующий анализ [81, 82]. Советуем авторам объяснять, почему и как группировали количественные данные, в том числе указывать количество категорий, точки разделения (*cut-points*), среднее арифметическое значение или медиану для каждой из группы. Если данные описываются в форме таблицы, для каждой категории необходимо указать число «случаев», «контролей», лиц, подвергшихся воздействию (*persons at risk*), человеко-время для периода

воздействия (*person-time at risk*). Таблицы не должны включать лишь оценки эффективности или результаты подбора модели.

Исследователи могут анализировать воздействие как непрерывное, чтобы сохранить максимум информации. При этом необходимо понимать характер влияния воздействия на исход. Поскольку это влияние может быть по ошибке автоматически принято как линейное, все отклонения от линейной зависимости должны быть изучены. Необходимо упомянуть альтернативные модели, применимость которых для анализа авторы изучали (логарифмическое преобразование, возведение в квадрат или сглаживание). Существует несколько методов для анализа связей в случае нелинейной зависимости между воздействием и исходом [82–84]. Кроме того, повысить информативность исследования поможет одно-временное представление результатов анализа и непрерывных, и сгруппированных данных количественного показателя воздействия, представляющего основной интерес.

В недавно опубликованном обзоре было показано, что в 2/3 статей по эпидемиологии изучали количественные переменные воздействия (*quantitative exposure variables*) [4]. В 42 статьях из 50 (84%) переменные были сгруппированы в несколько упорядоченных категорий, но часто без какого-либо обоснования сделанного выбора. В 15 статьях сообщалось об использовании линейных ассоциаций для моделирования непрерывного воздействия, но только 2 из них содержали результаты проверки на линейность. В другом обзоре литературных источников по психологии дихотомия была обоснована в 22 статьях из 110 (20%) [85].

12. Статистические методы.

12А. Опишите все статистические методы, включая использованные для контроля влияния вмешивающихся факторов.

Пример

«Скорректированный относительный риск рассчитывали, используя тест Mantel-Haenszel с целью оценки влияния возраста или пола в сравниваемых группах. Рассчитывали также 95% доверительный интервал скорректированного относительного риска, используя дисперсию, как это было описано в работе Greenland и соавт.» [93].

Пояснение

Не существует единственного корректного метода статистического анализа, напротив, есть несколько возможных методов решения одного и того же вопроса, исходя из различных предположений. Именно поэтому авторы должны предварительно определить тип анализа по крайней мере основных показателей исследования (*primary study objectives*) и зафиксировать его в протоколе. Часто необходим дополнительный анализ, вместо или наряду с изначально запланированным. Составляя отчёт об исследовании, авторы должны разъяснить

читателям, необходимость каких вариантов анализа была определена в ходе проверки данных. Несмотря на то, что различие между запланированным и разведочным (*exploratory*) анализом может быть нечётким, авторам следует разъяснить причины проведения последнего.

Если сравниваемые группы различаются по какому-либо признаку, необходимо проводить коррекцию по потенциально вмешивающимся факторам путём стратификации или с помощью многомерной регрессии (см. вставку 5) [94]. Часто выбор типа регрессионного анализа определяется дизайном исследования. Например, метод регрессии пропорциональных рисков по Коксу (*Cox proportional hazard regression*) обычно используется в когортных исследованиях [95], логистическая регрессия (*logistic regression*) часто становится методом выбора в исследованиях «случай–контроль» [96, 97]. Аналитики должны представить полное описание особенностей процедуры отбора переменных, а не только результаты для окончательной модели [98, 99]. Следует описать результаты сравнения моделей, проведённого для сокращения списка потенциальных вмешивающихся факторов, включаемых в окончательную модель. Даже если одна или две переменные вызывают большую часть очевидных искажений (*confounding*), читатели должны об этом знать. Нестандартные или новые статистические подходы должны быть подкреплены ссылками. Необходимо также указать использованное статистическое программное обеспечение. В качестве руководящего принципа рекомендуем следующее: статистические методы должны быть описаны настолько детально, чтобы квалифицированный читатель при доступе к исходным данным мог проверить результаты исследования [100].

В эмпирическом исследовании только 93 статьи из 169 (55%), представляющие оценки с поправкой на вмешивающиеся факторы, содержали чёткую информацию о том, каким образом непрерывные переменные и переменные с множеством категорий были включены в статистическую модель [101]. Другое исследование показало, что среди 67 статей, в которых статистический анализ был выполнен с учётом вмешивающихся факторов, большинство не содержало разъяснений того, как авторы выбирали эти факторы [4].

12В. Опишите все методы, использованные для изучения подгрупп и взаимодействий.

Пример

«Половые различия в восприимчивости к трём факторам риска, связанным с образом жизни, были изучены путём тестирования биологического взаимодействия согласно Rothman: создали новую композитную переменную с 4 категориями (a^-b^- , a^-b^+ , a^+b^- , and a^+b^+) для пола и дихотомической переменной изучаемого воздействия, где a^- и b^- означают отсутствие воздействия. ОР рассчитывали для каждой категории, делая поправку на возраст. Эффект взаимодействия определяли

как отклонение от аддитивности абсолютных эффектов, и рассчитывали избыточный ОР вследствие взаимодействия (ИОР):

$\text{ИОР} = \text{ОР}(a^+b^+) - \text{ОР}(a^-b^+) - \text{ОР}(a^+b^-) - 1$, где $\text{ОР}(a^+b^+)$ — относительный риск среди тех, кто подвержен воздействию обоих факторов; $\text{ОР}(a^-b^-)$ — референсная категория ($\text{ОР}=1,0$). 95% доверительный интервал рассчитывали, как предложено Hosmer и Lemeshow. Значение ИОР равное 0 означало отсутствие взаимодействия» [103].

Пояснение

Как будет подробно обсуждено ниже (см. пункт 17), много вопросов вызывает применимость и ценность анализа в ограниченных подгруппах исследуемой популяции [4, 104]. Тем не менее анализ в подгруппах проводится часто [4]. Читателей необходимо уведомить о том, какие анализы в подгруппах были запланированы заранее, а какие проводились исходя из полученных данных. Важно разъяснить также, какие методы были использованы для выявления различий в величине эффектов или связей между группами (см. пункт 17).

О взаимодействии («модификации эффекта») говорят, когда один фактор изменяет эффект другого. Совместное действие 2 факторов можно охарактеризовать двумя способами: используя аддитивную шкалу в терминах разности риска или мультипликативную шкалу — в терминах относительного риска (см. вставку 8). Многие авторы и читатели могут иметь собственные предпочтения в отношении того, как анализировать взаимодействие. Тем не менее им может быть интересно узнать, в какой степени совместный эффект воздействий отличается от отдельных эффектов. Все согласны с тем, что аддитивная шкала, в которой используются абсолютные риски, больше подходит для общественного здравоохранения и принятия клинических решений [105]. Какой бы ни была точка зрения, она должна быть чётко разъяснена читателям, как это сделано в приведённом примере выше [103]. Наиболее информативной может оказаться схема, представляющая отдельные эффекты обоих воздействий и их совместный эффект относительно отсутствия воздействия. Подобная схема приводится в примере пункта 17, а расчёты по разным шкалам — во вставке 8.

Вставка 5. Влияние вмешивающихся факторов

Влияние вмешивающихся факторов приводит к искажению эффектов. Может показаться, что исследование может либо подтвердить наличие связи между воздействием и риском возникновения заболевания, либо её опровергнуть. В действительности, кажущаяся связь или её отсутствие обусловлены другим фактором, который определяет возникновение заболевания и в то же время связан с воздействием. Этот фактор называется «вмешивающимся фактором» или «конфаундером», который приводит к неверной оценке потенциальной «причинно-следственной» связи с воздействием. Например, если женщинам среднего возраста, у которых повышается артериальное давление, оральные контрацептивы назначаются реже, то простое сравнение частоты сердечно-сосудистых заболеваний между теми, кто принимает контрацептивы, и теми, кто их не принимает, может создать ложное представление о том, что контрацептивы защищают от болезней сердца.

Исследователям следует заранее подумать о потенциальных вмешивающихся факторах. Если предусмотреть их влияние заранее, изучив всю необходимую информацию, дизайн исследования будет более обоснованным, а собранные данные — полными. Для коррекции влияния вмешивающихся факторов можно использовать ограничение (*restriction*) или подбор пар (*matching*). В приведённом выше примере количество участников исследования может быть ограничено женщинами, не подверженными влиянию вмешивающегося фактора — повышенного артериального давления. Подбор пар по значениям артериального давления также возможен, но нежелателен (см. вставку 2). Чтобы уменьшить влияние вмешивающихся факторов при анализе данных, исследователи могут прибегнуть к стратификации или многомерному анализу. Стратификация заключается в разделении данных на страты (категории) для вмешивающегося фактора (например, страты артериального давления), оценивании связей внутри каждой страты и вычисления общей оценки связи как средневзвешенного значения по всем слоям. Многомерный анализ даёт тот же результат, но позволяет учитывать большее количество переменных одновременно. Этот метод — более гибкий, но может повлечь за собой ряд дополнительных предположений о математической форме связи между воздействием и заболеванием.

Учёт вмешивающихся факторов имеет решающее значение в наблюдательных исследованиях, но у читателей не должно складываться впечатление, что анализ, скорректированный в этом отношении, устанавливает причинно-следственные связи (*association*). Искажение результатов может быть обусловлено ещё и остаточным влиянием вмешивающихся факторов (искажение, которое не удалось скорректировать [102]; *residual confounding*), случайной ошибкой выборки (*random sampling error*), систематической ошибкой отбора (*selection bias*) или ошибкой информации (*information bias*) (см. вставку 3).

12С. Объясните, как устраняли последствия отсутствующих (неполных) данных.

Пример

«Наш анализ неполных данных основывался на допущении о пропусках данных вследствие случайных причин (*missing at random*, MAR). Для множества многомерных импутаций (подстановки значений) мы использовали метод MICE (*multivariate imputation by chained equations*), реализованный в пакете STATA. Независимо анализировали 10 копий данных, каждая с подстановкой взамен отсутствующих значений, с помощью многомерной логистической регрессии. Мы усреднили полученные оценки переменных для получения одной средней оценки и скорректировали стандартную ошибку в соответствии с правилами Rubin» [106].

Пояснение

Отсутствующие данные — обычное явление в наблюдательных исследованиях. Разосланные участникам исследования опросники не всегда заполняются целиком, участники пропускают контрольные визиты, источники рутинно собираемых данных и клинические базы данных часто неполны. Несмотря на это, немногие статьи подробно описывают проблему отсутствующих данных [5, 107]. Существует несколько подходов, которые можно использовать для решения данной проблемы. Их преимущества и недостатки описаны во вставке 6. Мы советуем авторам указать количество отсутствующих данных для каждой из исследуемых переменных (воздействие, исходы, вмешивающиеся факторы) на каждом этапе анализа. При описании «потока» участников в ходе исследования авторы должны указать причины отсутствия данных, если это возможно, и указать число участников, исключённых из-за отсутствия данных (см. пункт 13). Для анализа, учитывающего отсутствующие данные, следует описать его характер (например, множественные импутации) и сделанные допущения (например, пропуск данных вследствие случайных причин) (см. вставку 6).

12D. Когортное исследование. Если применимо, опишите, как решали проблему выбывших из-под наблюдения.

Пример

«При проведении лечебных программ с активным наблюдением выявили, что у пациентов, выбывших из-под наблюдения, и тех, кто находился под наблюдением в течение 1 года, исходное количество CD4-клеток было схожим (медиана 115 и 123 клетки/мкл). Среди пациентов,

проходивших программу без активного контроля и потерянных для наблюдения, количество CD4-клеток было значительно ниже, чем среди тех, кто оставался под наблюдением (медиана 64 и 123 клетки/мкл). <...> Данные лечебной программы с пассивным контролем результатов были исключены из последующего анализа» [116].

Пояснение

При анализе данных когортных исследований используют таблицы смертности или иные подходы, основанные на человеко-времени наблюдения и времени до развития изучаемого заболевания. Предполагается, что среди лиц, у которых по окончании периода наблюдения не отмечали признаков заболевания, время наблюдения не связано с вероятностью развития исхода. Такое предположение будет верным, если период наблюдения заканчивается определённой датой или по достижении определённого возраста. Если участники прекращают участие в исследовании до наступления этой даты, речь идёт о выбывании из-под наблюдения (*loss to follow-up*). Достоверность (*validity*) исследования снижается, если выборочно выбывают участники, подвергшиеся воздействию, или пациенты с более высоким риском развития заболевания (цензурирование информации; *informative censoring*). В приведённом выше примере выбывавшие из-под наблюдения пациенты, получавшие лечение по программе, не предусматривавшей активного контроля, имели меньше CD4-хелперов, чем те, за кем продолжали наблюдать, и, следовательно, имели более высокий риск летального исхода [116].

Важно проводить различие между пациентами, завершившими исследование и выбывшими из него. К сожалению, статистические программы обычно не различают эти ситуации и в обоих случаях автоматически обрывают («цензурируют») время отслеживания исходов в момент окончания наблюдения. Поэтому исследователям необходимо решить, желательно ещё на этапе планирования исследования, как поступить с данными о пациентах, выбывших из-под наблюдения. Если потери невелики, исследователи могут исключить из анализа данные о пациентах, не завершивших наблюдение, или обрабатывать эти данные, предполагая, что участники умерли или к моменту выбывания из-под наблюдения, или к окончанию исследования. Советуем авторам указывать, сколько пациентов выбыло из-под наблюдения, и какая стратегия цензурирования данных была использована.

Вставка 6. Отсутствующие (неполные) данные: проблемы и способы их решения

Общепринятый подход — ограничить исследование участниками, в отношении которых имеются полные данные по всем переменным для проведения конкретного анализа. Хотя такой анализ «завершённых случаев» во многих случаях не подвержен систематическим ошибкам, он может быть таковым и всегда будет неэффективным [108]. Систематическая ошибка возникает, если участники с отсутствующими данными нетипичны для всей выборки, а неэффективность — вследствие уменьшения размера выборки.

Использование данных последнего наблюдения для повторных измерений может со временем привести к искажению тенденций, если участники с высокой вероятностью определённого исхода выборочно исключаются из анализа [109]. Добавление показателей отсутствующей категории для вмешивающегося фактора может увеличить остаточное искажение [107]. Метод импутаций, при котором вместо отсутствующего значения указывается предполагаемое или вычисленное значение, может привести к ослаблению или преувеличению исследуемой связи, а без использования сложных методов, описанных ниже, — к слишком малым стандартным ошибкам.

D.B. Rubin разработал типологию проблем, связанных с пропуском данных, основанную на модели вероятности отсутствия наблюдения [108, 110]. Отсутствие данных описывается как полностью случайное (*missing completely at random*, MCAR), если вероятность отсутствия конкретного наблюдения не зависит от значения какой-либо исследуемой переменной или переменных; как случайное (*missing at random*, MAR), если вероятность того, что наблюдения отсутствуют, не зависит от фактических значений отсутствующих данных. К примеру, если предположить, что у детей младшего возраста чаще отмечаются неполные данные спирометрического обследования, вероятность этого не будет связана с истинной ненаблюдаемой функцией лёгких после поправки с учётом возраста обследуемых. В этом случае отсутствующие значения измерения функции лёгких будут считаться MAR в статистических моделях, учитывающих возраст. Отсутствие данных может быть неслучайным (*missing not at random*, MNAR), если вероятность этого зависит от пропущенного значения даже после учёта имеющихся данных. В случае MNAR для большей убедительности выводы исследования необходимо дополнить разъяснениями того, что привело к потере данных.

Методы работы с данными категории MAR делятся на три обширных класса [108, 111] — вероятностные подходы [112], взвешенные оценки [113] и множественные импутации [111, 114]. Третий подход — наиболее гибкий и часто используемый, особенно если несколько переменных имеют пропущенные значения [115]. Результаты, полученные с использованием любого из этих подходов, следует сравнить с результатами анализа «завершённых случаев» и обсудить важные различия. Правдоподобность предположений, сделанных при анализе пропущенных данных, как правило, не может быть подтверждена или опровергнута. В частности, невозможно доказать, что пропущенные данные относятся к категории MAR, а не MNAR, поэтому в таких исследованиях рекомендуется проводить анализ чувствительности (*sensitivity analysis*) (см. пункты 12E и 17).

12D. Исследование «случай–контроль». Если применимо, объясните, как производили подбор (*matching*) «случаев» и «контролей».

Пример

«Для сравнения пациентов с деменцией и подобранных им в пары «контролей» с целью оценки факторов риска сердечно-сосудистых заболеваний, развития спонтанной эмболии сосудов мозга, поражения сонных артерий и артерио-венозных шунтов использовали тест McNemar, парный t-тест и условную логистическую регрессию» [117].

Пояснение

В исследованиях «случай–контроль» с индивидуальным подбором пар определение отношения шансов без учёта факта подбора обычно приводит к смещению оценок к единице (см. вставку 2). Именно поэтому зачастую необходим парный анализ (*matched analysis*). Интуитивно это можно представить как стратифицированный анализ, где каждый случай рассматривается как одна страта для своих «контролей». Анализ основывается на предположении, что «случаи» подвергаются воздействию чаще, чем «контроли», несмотря на то, что они одинаковы по признаку, по которому производится подбор. Исследователи могут проводить такой стратифицированный анализ, используя метод

Mantel-Haenszel для таблиц 2×2. В его простейшем варианте отношение шансов становится отношением пар, дискордантных по оказываемому на них воздействию. Если подбор происходит по таким универсальным характеристикам, как пол или возраст, нет необходимости в индивидуальном («пациент к пациенту») анализе, достаточно простого анализа по категориям пола и возраста [50]. Для других переменных, используемых для подбора пар, таких как соседство, родство или дружба, каждый подобранный набор «случаев» и «контролей» необходимо рассматривать как отдельную страту.

В исследованиях с индивидуальным подбором пар наиболее широко используется метод условной логистической регрессии: каждый «случай» и его «контроли» рассматриваются вместе. Этот метод необходим, когда число «контролей» для разных «случаев» неодинаково и когда кроме переменных, используемых для подбора пар, есть и другие переменные, которые необходимо использовать для коррекции. Для того чтобы читатели могли решить, был ли дизайн исследования с подобранными парами (*matched design*) принят в расчёт при анализе, рекомендуем авторам детально описывать применённые статистические методы анализа данных. Если учёт подбора пар мало влияет на полученные оценки,

авторы могут прибегнуть к анализу, используемому для сравнения независимых выборок.

12D. Одномоментное исследование. Если применимо, опишите аналитические методы, учитывающие стратегию формирования выборки.

Пример

«Стандартную ошибку рассчитывали, используя метод разложения Тейлора для оценки выборочной ошибки при комплексном дизайне выборки. <...> Выявили, что общее влияние дизайна исследования для диастолического артериального давления составило 1,9 у мужчин и 1,8 у женщин, для систолического артериального давления — 1,9 и 2,0 соответственно» [118].

Пояснение

В большинстве одномоментных исследований используется предварительно описанная стратегия формирования выборки (*sampling strategy*) из исходной популяции (*source population*). Стратегия может быть более сложной, чем простой случайный отбор. Она может включать несколько этапов, в том числе кластеризацию участников (например, по районам или деревням). Пропорциональная стратификация может гарантировать правильное представление подгрупп с определённой характеристикой. Непропорциональная стратификация может быть полезной для дополнительной выборки подгруппы, представляющей особый интерес.

Оценка связей в комплексной выборке может быть более или менее точной, чем в простой случайной выборке. Критерии точности, такие как стандартная ошибка или доверительный интервал, должны быть представлены с учётом эффекта дизайна (*design effect*), т.е. показателя отношения, который описывает, насколько точным будет исследование, если используется более сложная стратегия вместо простого случайного отбора [119]. Большинство сложных методик формирования выборки исследования приводят к снижению точности, в результате чего эффект дизайна превышает единицу.

Советуем авторам чётко описать методы, используемые для коррекции данных в комплексных стратегиях формирования выборки, что поможет читателям понять, как выбор метода создания выборки влияет на точность получаемых оценок. Например, при кластерной выборке неявный компромисс между более лёгким методом сбора данных и потерей точности очевиден, если авторы сообщают о влиянии дизайна. В приведённом выше примере расчёты показывают, что влияние дизайна составило 1,9 для мужчин. Это показывает, что необходимый размер выборки должен быть в 1,9 раз больше, чем при простом случайном отборе для получения оценок аналогичной точности.

12E. Опишите использованные анализы чувствительности.

Пример

«Поскольку в исследовании оказалась высокой доля отсутствующих данных по причине смерти участников

(38 из 148; 25,7%), по сравнению с выжившими (15 из 437; 3,4%) <...>, это могло привести к возникновению систематической ошибки. Поэтому мы провели анализ чувствительности. Предположили, что доля женщин, принимающих оральные контрацептивы (19,1% среди умерших, 11,4% среди выживших) в исследуемой группе, может быть применима к популяции в целом. Далее учитывали два крайних сценария: либо все подвергнутые воздействию и выбывшие из-под наблюдения принимали контрацептивы второго поколения, либо все — третьего» [120].

Пояснение

Анализ чувствительности или устойчивости результатов исследования к изменениям исходных условий (*sensitivity analyses*) помогает исследователям понять, согласуются ли основные результаты исследования с теми, что могли бы быть получены при использовании альтернативных подходов к анализу данных или допущений [121]. Таким образом, можно оценить влияние на результат критериев включения в анализ, вариантов определений воздействия и исхода [122], вмешивающихся факторов, отсутствующих данных [120, 123], возможных систематических ошибок отбора, неточного и неодинакового определения воздействия, заболеваний и других переменных, выбора методов анализа, например обработки количественных переменных (см. пункт 11). Для создания моделей одновременного влияния нескольких систематических ошибок или допущений все чаще используются более сложные методы [124–126].

В 1959 г. J. Cornfield и соавт. продемонстрировали, что относительный риск возникновения рака лёгких у курильщиков по сравнению с некурящими равен 9, и совершенно невероятно, чтобы это могло объясняться другими вмешивающимися факторами, поскольку таковые должны были бы встречаться у курильщиков минимум в 9 раз чаще, чем среди некурящих [127]. Этот анализ не исключил возможность наличия такого фактора, но определил, каковой должна быть распространённость такого фактора. Аналогичный подход недавно был использован для выявления правдоподобных вмешивающихся факторов, объясняющих связь между возникновением лейкемии у детей и проживанием вблизи линий электропередач [128]. В общем, анализ чувствительности может использоваться для определения степени влияния вмешивающихся факторов (*degree of confounding*), систематических ошибок отбора (*selection bias*) или ошибок информации (*information bias*). Важным и, возможно, недостаточно признанным является использование анализа чувствительности в случае, когда исследование демонстрирует незначительную связь между воздействием и исходом либо её отсутствие, с целью выявления действия вмешивающихся факторов или систематических ошибок, смещающих результаты в сторону нуля.

Результаты

Раздел «Результаты» должен содержать описание обнаруженных фактов: от набора участников исследования, описания выборки исследования до основных результатов и дополнительных анализов. Здесь не должно быть интерпретаций и рассуждений, отражающих взгляды и мнения авторов.

13. Участники.

13А. Укажите количество лиц на каждом этапе исследования: число потенциально подходящих для участия в исследовании, обследованных на предмет соответствия критериям отбора, допущенных к участию, включённых в исследование, завершивших последующее наблюдение, а также участников, чьи результаты были проанализированы.

Пример

«Из выборки 105 отдельно расположенных баров и таверн 13 заведений больше не работали, 9 располагались при ресторанах, оставшиеся 83 отвечали условиям отбора. Собственникам 22 баров не удалось дозвониться после шести и более попыток. Собственники 36 баров отказались участвовать в исследовании <...>. В 25 барах и тавернах, принявших участие в исследовании, состояло в штате 124 бармена, 67 из них как минимум 1 раз в неделю работали в дневную смену. Из числа последних 54 бармена (81%), были полностью проинтервьюированы в начале исследования и прошли спирометрию, 53 из них (98%) завершили период наблюдения» [129].

Пояснение

Детальная информация о том, как происходил набор участников в исследование, важна по нескольким причинам. Включённые в исследование часто по разным причинам отличаются от целевой популяции (*target population*), к которой применяются результаты исследования. Это может привести к оценкам распространённости (*prevalence*) и инцидентности (*incidence*), которые не отражают положения дел в целевой популяции. Например, лица, соглашающиеся участвовать в почтовом опросе об особенностях сексуального поведения, посещают церковь реже, придерживаются менее консервативных взглядов на секс и получили первый сексуальный опыт в более раннем возрасте, а также с большей вероятностью курят сигареты и употребляют алкоголь, чем лица, отказавшиеся участвовать в таком исследовании [130]. Эти различия позволяют предположить, что в ходе почтового опроса можно переоценить сексуальный либерализм и активность популяции. Такая систематическая ошибка отклика (*response bias*) (см. вставку 3) может исказить связь воздействия и заболевания, если эта связь различается у подходящих для включения в исследование и включённых в него.

Другой пример: связь между ранним материнством и возникновением лейкемии у их потомства, продемонстрированная в нескольких исследованиях «случай–контроль» [131, 132], объяснялась неспособностью

молодых женщин в группах «случаи» и «контроли». Юные матери здоровых детей были менее склонны участвовать в исследованиях, чем те, чьи дети болели [133]. Хотя низкий уровень участия не обязательно ставит под сомнение достоверность (*validity*) исследования, прозрачность информации об участии и причинах неучастия имеет важное значение. Кроме того, поскольку не существует универсального определения таких понятий, как «доля приглашённых участвовать» (*participation rate*), «доля согласившихся участвовать» (*response rate*), «доля завершивших наблюдение» (*follow-up rate*), необходимо пояснять читателям, как авторы рассчитывали эти показатели [134].

В идеале исследователи должны проводить подсчёт числа индивидуумов на каждом этапе набора участников: от определения целевой популяции до включения данных об участнике в анализ. В зависимости от типа исследования может потребоваться подсчёт числа потенциально подходящих участников (*potentially eligible*), проверенных на соответствие критериям отбора (*assessed for eligibility*), признанных подходящими для участия (*to be eligible*), включённых в исследование (*included in the study*), прошедших обследование (*examined*), продолживших участие (*followed up*) и включённых в анализ (*included in the analysis*). Если выборка участников исследования формируется в 2 и более стадии (многоступенчатый отбор), как это произошло в описанном выше примере, может потребоваться информация о составе выборки на каждом из этапов. В исследованиях «случай–контроль» мы советуем авторам описывать «поток» (последовательность и результаты формирования выборки) участников отдельно для групп «случаев» и «контролей» [135]. Иногда «контроли» отбираются из нескольких источников, например госпитализированные пациенты и обитатели домов престарелых. В этом случае рекомендуем отдельно подсчитывать число участников каждого типа. S.H. Olson и соавт. разработали полезные рекомендации по составлению отчётов о подборе «контролей» путём случайного обзвона и применения других методов [136].

Недавний обзор эпидемиологических исследований, опубликованный в 10 журналах по общей эпидемиологии, общественному здравоохранению и медицине, показал, что некоторая информация об участии была предоставлена в 47 исследованиях «случай–контроль» из 107 (59%), 49 когортных исследованиях из 154 (32%) и 51 одномоментном исследовании из 86 (59%) [137]. В двух других обзорах литературы было также показано, что публикуемые сведения об участии или неучастии в эпидемиологических исследованиях были неполными либо отсутствовали вовсе [4, 5]. Наконец, есть свидетельства того, что участие в эпидемиологических исследованиях могло снизиться в последние десятилетия [137, 138], что свидетельствует о необходимости прозрачной отчётности [139].

13В. Укажите причины выбывания на каждом этапе исследования.

Пример

«Основными причинами не включения были тяжёлая болезнь или смерть потенциального участника до проведения интервью (30% в группе «случаев», менее 1% в группе «контролей»), неполучение ответа (2% в группе «случаев», 21% в группе «контролей»), отказ от участия (10% в группе «случаев», 29% в группе «контролей») и другие причины (отказ от сотрудничества с консультантом или врачом общей практики, незнание английского языка, психические нарушения) (7% в группе «случаев», 5% в группе «контролей»)» [140].

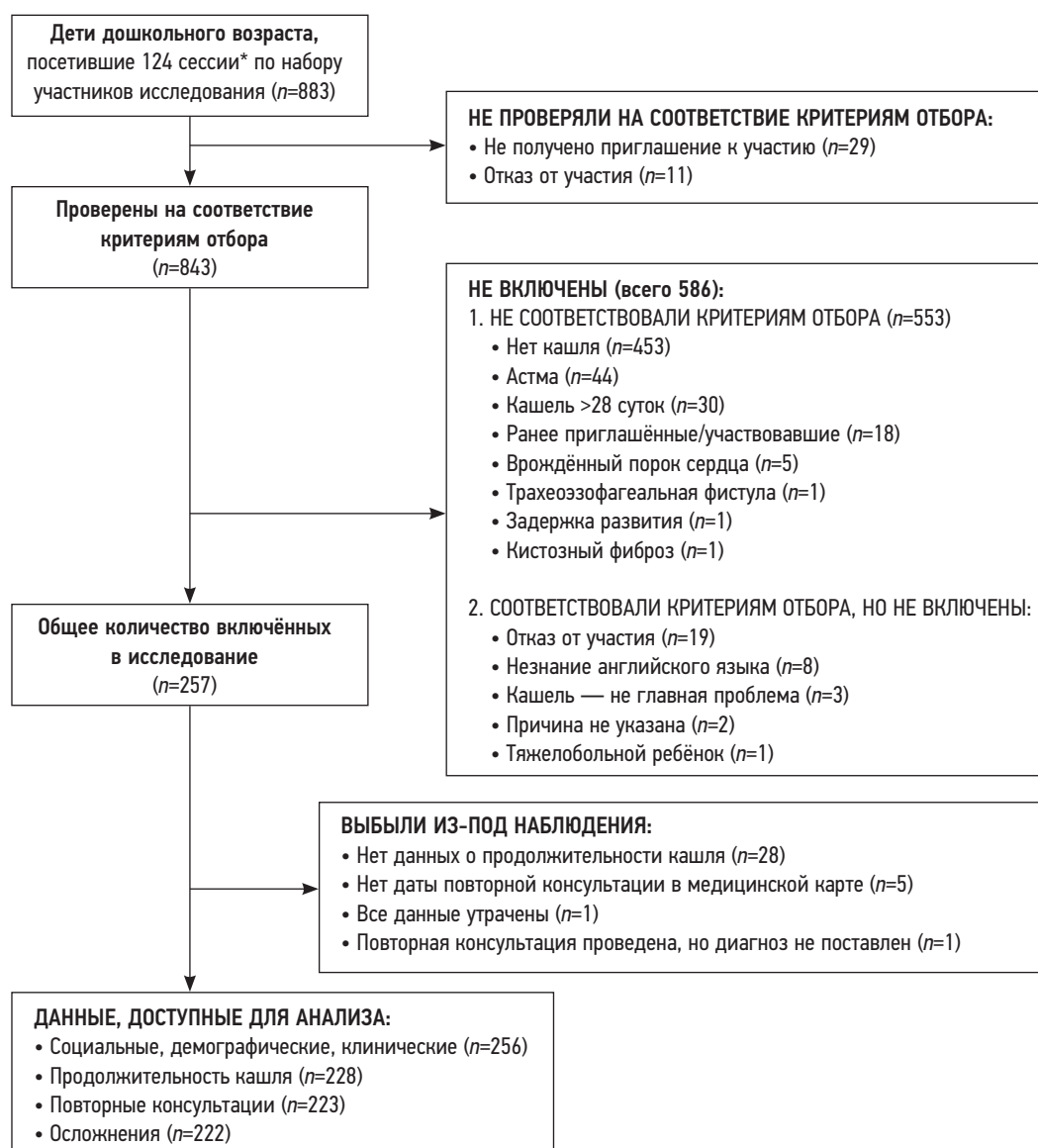
Пояснение

Раскрытие причин не включения, а также прекращения участия в исследовании или исключения информации об участниках из статистического анализа

помогает читателям прийти к заключению, репрезентативна ли выборка исследования (*study population*) целевой популяции (*target population*), и был ли риск возникновения систематической ошибки. Например, в опросе о здоровье неучастие в силу причин, не относящихся к состоянию здоровья (например, неполучение письма-приглашения из-за неверно указанного адреса), может влиять на точность оценок, но почти наверняка не приведёт к систематической ошибке. В противоположность этому, если многие уклоняются от участия в опросе из-за болезни или, наоборот, из-за ощущения себя здоровыми, это может привести к переоценке или недооценке распространённости проблем со здоровьем в популяции.

13С. Рассмотрите возможность использования потоковой диаграммы.

Пример



Примечание. * Отсутствуют данные об одной из сессий по набору участников, на которой присутствовали не менее 3 посетителей с кашлем, отобрано 2 участника.

Потоковая диаграмма адаптирована из работы A.D. Нау и соавт. [141].

Пояснение

Информативная и хорошо структурированная потоковая диаграмма (как в примере выше) помогает сделать информацию более доступной и легко читаемой, чем подробное описание в тексте [142]. Диаграмма может включать основные результаты, такие как число основных исходов. Хотя мы рекомендуем использовать потоковую диаграмму, в частности для сложных наблюдательных исследований, мы не предлагаем для неё конкретного формата.

14. Описательные данные.

14А. Охарактеризуйте участников исследования (например, укажите демографические, клинические, социальные характеристики). Приведите данные о воздействиях и потенциальных вмешивающихся факторах.

Пример (см. табл. 2)

Пояснение

Описание участников и изучаемого воздействия необходимо читателям для того, чтобы судить об обобщаемости (*generalisability*) результатов. Информация о потенциальных вмешивающихся факторах, включающая описание способов их измерения, влияет на суждение о достоверности исследования (*study validity*). Мы советуем авторам описывать непрерывные переменные для каждой группы исследования, указывая среднее арифметическое и стандартное отклонение или, в случае асимметричного распределения, медиану и процентильный разброс (например, 25-й и 75-й процентиля). Переменные, включающие небольшое число упорядоченных категорий (например, стадии заболевания от I до IV), не должны быть представлены в виде непрерывных переменных: лучше указывать число и долю для каждой категории (см. вставку 4). В исследованиях с группами сравнения описательные характеристики и числа должны приводиться для каждой группы, как в примере выше.

Расчётные величины, такие как стандартная ошибка и доверительный интервал, не должны использоваться

для описания вариабельности характеристик, а в описательных таблицах не следует приводить результаты статистических тестов. Кроме того, *p*-значения не подходят для отбора вмешивающихся факторов с целью коррекции анализа; даже небольшое различие групп в распределении значений фактора, который сильно влияет на исход, может быть важным [144, 145].

В когортных исследованиях может оказаться полезным документировать, как воздействие влияло на другие характеристики и вмешивающиеся факторы. Авторы могут представить эту информацию в таблице со столбцами для двух или более категорий воздействия, что позволит судить о разнице вмешивающихся факторов между этими категориями.

В исследованиях «случай–контроль» суждение о потенциальных вмешивающихся факторах не может выноситься на основании сравнения «случаев» и «контролей». Лица, включённые в контрольную группу, представляют исходную популяцию (*source population*) и обычно отличаются от группы «случаев» по многим аспектам. Например, в исследовании, посвящённом оральным контрацептивам и инфаркту миокарда, молодые женщины с инфарктом миокарда отличались более высокой концентрацией холестерина, курением, отягощённым семейным анамнезом, чем участницы группы «контролей» [146]. Это не влияло на оценку воздействия оральных контрацептивов, в том числе потому, что при прописывании этих лекарственных средств не руководствовались наличием факторов риска, поскольку их оценку проводили после развития инфаркта (см. вставку 5). В исследованиях «случай–контроль» эквивалент сравнения подвергшихся и не подвергавшихся воздействию на наличие вмешивающихся факторов (как это делается в когортных исследованиях) может быть достигнут путём изучения исходной для случаев популяции: если контрольная группа достаточно велика и представляет исходную популяцию, подвергшиеся и не подвергавшиеся воздействию «контроли» можно сравнить на наличие потенциальных вмешивающихся факторов [121, 147].

Таблица 2. Характеристики участников при включении в исследование, G. Castellana (Италия), 1985–1986 гг.

	ВГС отрицательные (<i>n</i> =1458)	ВГС положительные (<i>n</i> =511)	Неизвестно (<i>n</i> =513)
Пол (%)			
Мужской	936 (64%)	296 (58%)	197 (39%)
Женский	522 (36%)	215 (42%)	306 (61%)
Средний возраст при регистрации (СО)	45,7 (10,5)	52,0 (9,7)	52,5 (9,8)
Ежедневное потребление алкоголя (%)			
Нет	250 (17%)	129 (25%)	119 (24%)
Умеренное ¹	853 (59%)	272 (53%)	293 (58%)
Чрезмерное ²	355 (24%)	110 (22%)	91 (18%)

Примечание. СО — стандартное отклонение; ВГС — вирус гепатита С. ¹ Мужчины <60 г этанола/сутки, женщины <30 г этанола/сутки;

² Мужчины >60 г этанола/сутки, женщины >30 г этанола/сутки.

Таблица адаптирована из работы A.R. Osella и соавт. [143].

Таблица 3. Симптомы заболевания как конечные точки в анализе выживаемости

	Кашель	Одышка	Бессонница
Симптом устранён	201 (79%)	138 (54%)	171 (67%)
Цензурирование	27 (10%)	21 (8%)	24 (9%)
Симптомы отсутствовали	0	46 (18%)	11 (4%)
Данные отсутствуют	28 (11%)	51 (20%)	50 (20%)
Всего	256 (100%)	256 (100%)	256 (100%)

Таблица адаптирована из работы A.D. Hsu и соавт. [141].

14В. Для каждой из анализируемых переменных укажите количество участников с пропущенными данными.

Пример (см. табл. 3)

Пояснение

Поскольку отсутствующие данные могут быть источником систематической ошибки или влиять на обобщаемость результатов, авторы должны сообщить читателям о количестве таких данных применительно к воздействию, потенциальным вмешивающимся факторам и другим важным характеристикам пациентов (см. пункт 12С и вставку 6). В когортном исследовании необходимо указать количество участников, выбывших из-под наблюдения, и причины выбывания, поскольку незавершённая процедура наблюдения может исказить результаты (см. пункты 12D и 13) [148]. Мы советуем авторам использовать таблицы и рисунки для представления этой информации.

14С. Когортное исследование. Опишите характеристики времени наблюдения (например, укажите среднюю и общую продолжительность наблюдения).

Пример

«На протяжении 4366 человеко-лет наблюдения (медиана длительности наблюдения 5,4 года, максимум 8,3 года) у 265 участников была диагностирована

деменция, в том числе у 202 — болезнь Альцгеймера» [149].

Пояснение

Читателям необходимо знать длительность и границы периода наблюдения за участниками, в течение которого была получена информация об исходах. Авторы могут представить сводку этих данных, указав значения среднего арифметического и/или медианы времени наблюдения. Умножение средней продолжительности наблюдения на число участников исследования позволит читателям рассчитать общее количество человеко-лет. Авторы могут также привести минимальное и максимальное время наблюдения или процентиля распределения. Можно указать общую длительность наблюдения в человеко-годах или долю данных, которые были собраны [148]. Вся информация такого рода может быть представлена отдельно для участников в соответствии с 2 и более категориями воздействия. Почти половина из 132 статей в журналах по онкологии (в основном когортные исследования) не содержала описания данных о продолжительности наблюдения [37].

15. Исходы

15А. Когортное исследование. Укажите количество событий исхода или представьте сводку основных показателей за время наблюдения.

Пример (см. табл. 4)

Таблица 4. Случаи ВИЧ-1 сероконверсии по отдельным социально-демографическим переменным: 1990–1993 гг.

Переменные	Человеко-год	Кол-во участников с сероконверсией	Число событий / 1000 человеко-лет (95% ДИ)
Календарный год			
1990	2197,5	18	8,2 (4,4–12,0)
1991	3210,7	22	6,9 (4,0–9,7)
1992	3162,6	18	5,7 (3,1–8,3)
1993	2912,9	26	8,9 (5,5–12,4)
1994	1104,5	5	4,5 (0,6–8,5)
Племена			
Баганда	8433,1	48	5,7 (4,1–7,3)
Другие народы Уганды	578,4	9	15,6 (5,4–25,7)
Руандийцы	2318,6	16	6,9 (3,5–10,3)
Другие племена	866,0	12	13,9 (6,0–21,7)
Религия			
Мусульмане	3313,5	9	2,7 (0,9–4,5)
Другая религия	8882,7	76	8,6 (6,6–10,5)

Примечание. ДИ — доверительный интервал.

Таблица адаптирована из работы J.F. Kengeya-Kayondo и соавт. [150].

Таблица 5. Факторы экспозиции среди «случаев» и «контролей» с циррозом печени

	«Случаи» (n=40)	«Контроли» (n=139)
Винилхлоридный мономер (совокупное воздействие: ч/млн × годы)		
<160	7 (18%)	38 (27%)
160–500	7 (18%)	40 (29%)
500–2500	9 (23%)	37 (27%)
>2500	17 (43%)	24 (17%)
Потребление алкоголя (г/день)		
<30	1 (3%)	82 (59%)
30–60	7 (18%)	46 (33%)
>60	32 (80%)	11 (8%)
HBsAG/HCV		
Отрицательный	33 (83%)	136 (98%)
Положительный	7 (18%)	3 (2%)

Примечание. HBsAG — поверхностный антиген вируса гепатита В; HCV — вирус гепатита С.

Таблица адаптирована из работы G. Mastrangelo и соавт. [151].

15В. Исследования «случай–контроль». Укажите количество событий по каждой категории воздействия или сводные показатели воздействия.

Пример (см. табл. 5)

15С. Одномоментное исследование. Укажите количество событий исхода или описание основных показателей исследования.

Пример (см. табл. 6)

Пояснение

Прежде чем обратиться к возможной связи между воздействиями (факторами риска) и исходами, авторы должны привести некоторые относящиеся к теме описательные данные. Оценку такой связи можно представить в той же таблице, в которой представлены описательные данные (см. пункт 14А). В когортном исследовании при изучении событий в качестве исходов (*outcomes*) укажите количество событий по каждому анализируемому исходу. При необходимости опишите частоту событий в пересчёте на человеко-год наблюдений (*event rate per person-year*). Если риск событий изменялся в период наблюдения, представьте число и частоту событий, разбив период наблюдения на подходящие интервалы

с помощью таблицы дожития (смертности) или графика Kaplan–Meier. Может оказаться предпочтительным использовать график кумулятивной инцидентности (*cumulative incidence*), который возрастает от 0% быстрее, чем замедляется при приближении к 100%, особенно если частота наблюдаемых событий составляет менее 30% [153]. Рассмотрите возможность представления такой информации отдельно для участников в различных категориях воздействия. Если в когортном исследовании изучают и другие зависящие от времени исходы (например, количественные маркеры заболевания, такие как артериальное давление), укажите, как изменялись соответствующие сводные оценки (например, среднее арифметическое и стандартное отклонение) в течение времени, возможно, в виде таблицы или рисунка.

Для одномоментных исследований мы рекомендуем представлять, как описано выше, информацию о распространённости событий исходов или сводные оценки измерений. В исследованиях «случай–контроль» основное внимание должно быть уделено описанию частотных или количественных характеристик воздействий отдельно для «случаев» и «контролей» [154]. При любом

Таблица 6. Распространённость астмы и диагностированной сенной лихорадки в зависимости от среднего уровня антигена плесневого грибка *Alternaria alternate* в жилых помещениях

Категории содержания грибка <i>Alternaria</i> *	Астма		Сенная лихорадка	
	N	Распространённость† (95% ДИ)	N	Распространённость† (95% ДИ)
1-й тертиль	40	4,8 (3,3–6,9)	93	16,4 (13,0–20,5)
2-й тертиль	61	7,5 (5,2–10,6)	122	17,1 (12,8–22,5)
3-й тертиль	33	8,7 (6,7–11,3)	93	15,2 (12,1–18,9)

Примечание. * 1-й тертиль <3,90 мкг/г; 2-й тертиль от 3,90 до 6,27 мкг/г; 3-й тертиль ≥6,28 мкг/г.

Проценты (95% ДИ) взвешены для многоступенчатого дизайна формирования выборки Национального исследования свинца и аллергенов в жилых помещениях.

Таблица адаптирована из работы P.M. Salo и соавт. [152].

дизайне исследования может быть полезно представить непрерывные данные исходов или воздействий в таблице с разбивкой по категориям, даже если эти данные не анализируются как таковые.

16. Основные результаты

16А. Укажите нескорректированные оценки и, если применимо, оценки с поправкой на вмешивающиеся факторы и их точность (например, 95% доверительный интервал). Объясните, влияние каких вмешивающихся факторов было учтено, и почему именно эти факторы были выбраны.

Пример 1

«Изначально мы рассматривали включение следующих переменных в качестве возможных вмешивающихся факторов в стратифицированный анализ Mantel-Haenszel <...>. В окончательную логистическую регрессионную модель включили следующие переменные <...>, что вызвало 10% изменение отношения шансов после поправки по Mantel-Haenszel» [155].

Пример 2 (см. табл. 7)

Пояснение

Во многих случаях авторы могут представлять читателям результаты нескорректированного, минимально скорректированного и полностью скорректированного анализа. Мы советуем приводить результаты нескорректированного анализа вместе с основными данными, например указывать число «случаев» и «контролей», подвергшихся и не подвергавшихся воздействию. Это позволит читателям разобраться в данных, прежде чем переходить к анализу связей (см. также пункт 15).

Для скорректированных анализов укажите число лиц, включённых в анализ, поскольку это число может отличаться от такового при нескорректированном анализе вследствие неполных данных у некоторых переменных, используемых для коррекции (см. пункт 12С). Оценки необходимо приводить вместе с доверительными интервалами.

Читатели могут сравнивать нескорректированные показатели связи с теми, которые скорректированы с учётом потенциальных вмешивающихся факторов, и судить, насколько и в каком направлении они изменились. Читатели могут подумать, что «скорректированные» результаты представляют собой «причинную» часть показателей связи, однако скорректированные результаты не обязательно свободны от случайной ошибки выборки (*random sampling error*), систематических ошибок отбора (*selection bias*) и информации (*information bias*), остаточного влияния вмешивающихся факторов (*residual confounding*) (см. вставку 5). Именно поэтому интерпретировать скорректированные результаты надо с большой осторожностью, их достоверность часто зависит от полноты информации о важных вмешивающихся факторах, точности их измерения и выбора подходящей статистической модели (см. пункт 20) [157, 158].

Авторы должны объяснить все рассматриваемые ими потенциальные вмешивающиеся факторы, а также критерии включения или исключения переменных в статистические модели. При принятии решения о включении или исключении переменных нужно руководствоваться знаниями, чётко сформулированными предположениями

Таблица 7. Отношение частот повторной госпитализации при лечении пациентов, находящихся в стационаре по месту жительства, после первой госпитализации в связи с шизофренией и шизоаффективными расстройствами

Лечение	Кол-во рецидивов	Человеко-год	Отношение частот (95% ДИ)	Скорректированное отношение частот (95% ДИ)	Полностью скорректированное отношение частот (95% ДИ)
Перфеназин	53	187	0,41 (0,29–0,59)	0,45 (0,32–0,65)	0,32 (0,22–0,49)
Оланзапин	329	822	0,59 (0,45–0,75)	0,55 (0,43–0,72)	0,54 (0,41–0,71)
Клозапин	336	804	0,61 (0,47–0,79)	0,53 (0,41–0,69)	0,64 (0,48–0,85)
Хлорпротиксен	79	146	0,79 (0,58–1,09)	0,83 (0,61–1,15)	0,64 (0,45–0,91)
Тиоридазин	115	201	0,84 (0,63–1,12)	0,82 (0,61–1,10)	0,70 (0,51–0,96)
Перфеназин	155	327	0,69 (0,58–0,82)	0,78 (0,59–1,03)	0,85 (0,63–1,13)
Рisperидон	343	651	0,77 (0,60–0,99)	0,80 (0,62–1,03)	0,89 (0,69–1,16)
Галоперидол	73	107	1,00	1,00	1,00
Хлорпромазин	82	127	0,94 (0,69–1,29)	0,97 (0,71–1,33)	1,06 (0,76–1,47)
Левомепромазин	52	63	1,21 (0,84–1,73)	0,82 (0,58–1,18)	1,09 (0,76–1,57)
Без антипсихотической терапии	2248	3362	0,98 (0,77–1,23)	1,01 (0,80–1,27)	1,16 (0,91–1,47)

Примечание. Скорректировано с учётом пола, календарного года, возраста на момент начала наблюдения, количества предыдущих рецидивов, продолжительности первой госпитализации и последующего наблюдения (скорректированные показатели), а также дополнительно по показателю предрасположенности (*score of the propensity*) к началу лечения препаратами, отличными от галоперидола (полностью скорректированные показатели).

Таблица адаптирована из работы J. Tiihonen и соавт. [156].

или причинными отношениями. Неверные решения могут привести к систематическим ошибкам, например при включении переменных, вовлечённых в причинную цепочку между воздействием и заболеванием (за исключением случаев, когда цель заключается в оценке того, какой эффект оказывает промежуточная переменная). Если решение о включении переменной в модель основано на изменении оценки, важно указать, какие изменения было решено считать достаточно значимыми для включения переменной в анализ. Если при отборе вмешивающихся факторов использовали стратегии «обратного исключения» (*'backward deletion'*) или «прямого включения» (*'forward inclusion'*), этот процесс должен быть разъяснён, и указаны уровни значимости для отклонения нулевой гипотезы об отсутствии влияния вмешивающихся факторов. Следует отметить, что мы не советуем выбирать вмешивающиеся факторы, основываясь исключительно на проверке статистической значимости [147, 159, 160].

Согласно недавним работам, посвящённым оценке качества отчётов эпидемиологических исследований, доверительные интервалы были указаны в большинстве статей [4]. Однако лишь немногие авторы объяснили свой выбор вмешивающихся факторов [4, 5].

16В. При категоризации непрерывных переменных, укажите границы категорий.

Пример (см. табл. 8)

Пояснение

Категоризация значений непрерывной переменной имеет несколько важных последствий для анализа (см. вставку 4) и влияет на представление результатов. В таблицах исходы должны быть приведены для каждой категории фактора воздействия (*exposure category*), например с указанием числа лиц, подвергшихся воздействию (*persons at risk*), человеко-времени для периода воздействия (*person-time at risk*), если применимо — то отдельно для каждой группы (например, для «случаев» и «контролей»). Детальное описание категорий может помочь при сравнении исследований и проведении метаанализа. Если данные группировали с использованием традиционных точек разделения (*cut-points*), таких как пороговые значения индекса массы тела [162], границы групп (т.е. диапазон значений) легко представить, за исключением самой верхней и нижней категории. Если разделение на категории выполнено по значениям

квантилей, границы групп не могут быть выведены из данных, поэтому авторы как минимум должны указать границы категории; полезно также указать диапазон данных, среднее арифметическое значение или медиану для значений показателя внутри категорий.

16С. Если применимо, рассмотрите возможность преобразования относительного риска в абсолютный для значимого периода времени.

Пример

«10-летнее использование гормонально-заместительной терапии привело к 5 (95% ДИ 3–7) дополнительным случаям рака молочной железы на 1000 женщин, получавших только эстрогенные препараты, и 19 (95% ДИ 15–23) дополнительным случаям рака на 1000 женщин, получавших комбинацию эстрогенов и прогестинов» [163].

Пояснение

Результаты исследований, в которых изучалась связь (*association*) между воздействием и заболеванием, часто представляют в виде относительных величин: отношение рисков, частот или шансов (см. вставку 8). Относительные показатели отражают силу такой связи. Если относительный риск (*relative risk*) сильно отличается от 1, маловероятно, что связь обусловлена вмешивающимися факторами [164, 165]. Относительные эффекты или ассоциации, как правило, более согласованы между исследованиями и популяциями, чем абсолютные показатели, но то, что часто имеет место, может быть неуместным в конкретном случае. Например, схожий относительный риск был выявлен при изучении классических факторов риска сердечно-сосудистых заболеваний у мужчин, живущих в Северной Ирландии, Франции, Соединённых Штатах Америки и в Германии, хотя базовый уровень риска ишемической болезни сердца различается между этими странами весьма значительно [166, 167]. В противоположность этому, в исследовании артериальной гипертензии как фактора риска смерти от сердечно-сосудистых заболеваний данные были более согласованы с постоянной (во времени) разницей частот (*constant rate difference*), чем с постоянным отношением частот (*constant rate ratio*) [168].

Широко используемые статистические модели, в том числе логистической регрессии [169] и модели пропорциональных рисков Кокса [170], основываются на относительных величинах. В этих моделях легко выявить лишь отклонения относительных характеристик эффекта от постоянного уровня. Тем не менее показатели отклонения от аддитивности разницы рисков, такие как избыточный относительный риск вследствие взаимодействия (*relative excess risk from interaction*, RERI) (см. пункт 12В и вставку 8), также можно изучать с помощью моделей, основанных на относительных показателях.

Во многих случаях абсолютный риск (*absolute risk*), ассоциируемый с воздействием, более интересен, чем относительный риск. Например, если в фокусе

Таблица 8. Полихлорированные бифенилы в сыворотке пуповинной крови

Квартиль	Диапазон, нг/г	Количество
1	0,07–0,24	180
2	0,24–0,38	181
3	0,38–0,60	181
4	0,61–18,14	180

Таблица адаптирована из работы S.K. Sagiv и соавт. [161].

исследования побочные эффекты лекарственного средства, то интересным может быть число дополнительных случаев в единицу времени приёма этого средства (дни, недели или годы). В примере выше указано дополнительное число случаев рака молочной железы на 1000 женщин, использовавших заместительную гормональную терапию в течение 10 лет [163]. Такие показатели, как дополнительный риск (*attributable risk*) или дополнительная доля риска в популяции (*population attributable fraction*), могут быть полезны для измерения числа случаев заболевания, которые можно было предотвратить, устранив воздействие. Их необходимо представлять вместе с показателями статистической неопределённости (например, с доверительными интервалами, как в примере выше). Авторы должны знать о значимых допущениях, сделанных в этом контексте, включая причинно-следственные связи между фактором риска и заболеванием (см. вставку 7) [171]. Вследствие

семантической неоднозначности и сложности авторы должны подробно сообщить, какие методы использовались для расчёта дополнительных рисков, приписываемых исследуемому фактору, в идеале с указанием используемых формул [172].

Недавний обзор аннотаций 222 статей, опубликованных в ведущих медицинских журналах, показал, что в большинстве (62%) аннотаций рандомизированных исследований в качестве показателя отношения были указаны абсолютные риски. Что касается когортных исследований, такие данные были представлены в гораздо меньшем количестве статей (21%) [173]. Обзор информации в базе данных Medline за период с 1966–1997 гг. показал, что дополнительные риски (*attributable risks*) указывались в названии или аннотации статей 619 раз в сравнении с относительным риском или отношением шансов, упоминаемых 18 955 раз, т.е. в соотношении 1:31 [174].

Вставка 7. Показатели связи, эффекта и влияния

Наблюдательные исследования могут проводиться исключительно для описания масштаба и распространения отклонений в состоянии здоровья среди населения. В таких исследованиях могут изучать количество людей, у которых есть определённое заболевание в конкретный момент времени (распространённость), или у которых заболевание развивается в течение определённого периода (инцидентность, заболеваемость). Заболеваемость может быть выражена долей лиц, у которых развивается заболевание (кумулятивная инцидентность, *cumulative incidence*), или показателем частоты событий в единицах человеко-времени (относительная инцидентность, *incidence rate*). Для описания различных типов инцидентности, среди прочего, используются такие термины, как смертность (*mortality rate*), рождаемость (*birth rate*), поражённость (*attack rate*) и летальность (*case fatality rate*). Аналогичным образом при описании различных типов распространённости говорят о распространённости заболевания в конкретный момент времени (*point prevalence*), за определённый период времени (*period prevalence*), в течение года (*annual prevalence*) или в течение жизни (*lifetime prevalence*) [30].

В других наблюдательных исследованиях рассматриваются причинно-следственные связи. Их целью является сравнение рисков, частоты или распространённости интересующего события между теми, кто подвергается или не подвергается воздействию исследуемого фактора риска. В этих исследованиях часто оценивается «относительный риск», который может означать отношения рисков (отношения показателей кумулятивной инцидентности, *ratios of cumulative incidences*), а также отношения частоты событий в единицах человеко-времени (отношения показателей относительной инцидентности, *ratios of incidence rates*). В исследования «случай–контроль» включают только часть исходной популяции («контроли»). Результаты выражаются как отношение шансов воздействия предполагаемой причины заболевания между «случаями» и «контролями». Отношение шансов (*odds ratio*) позволяет оценивать отношение рисков или частот в зависимости от выборки «случаев» и «контролей» (см. вставку 1) [175, 176]. В некоторых случаях может быть полезным определение отношения распространённости (*prevalence ratio*) или отношения шансов распространённости (*prevalence odds ratio*) исходя из данных одномоментных исследований [177].

Часто бывает полезным выражение результатов как в относительных, так и в абсолютных величинах. Например, в исследовании, проведённом с участием британских врачей-мужчин, смертность от рака лёгких за 50 лет наблюдения составил 249 заболевших на 100 000 человек в год среди курильщиков по сравнению с 17 заболевшими на 100 000 человек в год среди некурящих, при этом отношение частот составляло 14,6 (249/17) [178]. Для случаев ишемической болезни сердца (ИБС) соответствующие показатели составили 1001 и 619 заболевших на 100 000 человек в год, отношение частот — 1,61 (1001/619). Выходит, что курение оказывает гораздо большее влияние на смертность от рака, чем от ИБС. Но если рассматривать абсолютный эффект курения, показатели будут другими. Разница в значениях относительной инцидентности составила 232 заболевших на 100 000 человек в год (249/17) для рака лёгких и 382 случая заболевания (1001/619) для ИБС. Таким образом, среди курящих докторов курение с большей вероятностью было причиной смерти от ИБС, чем от рака лёгких.

Какое бремя болезни в популяции можно предотвратить, устранив воздействие? Согласно глобальным оценкам, 91% всех случаев рака лёгких, 40% ИБС и 33% всех летальных исходов среди мужчин в 2000 году были вызваны курением [179]. Дополнительная доля риска в популяции (*population attributable fraction*), как правило, оценивается как доля случаев, возникших под определённым воздействием, но существует нескольких концепций определения этого показателя (без единой терминологии). Более того, иногда используются неверные подходы для коррекции величины дополнительного риска с учётом других факторов [172, 180]. Как следствие, относительные показатели указывают на значимость связей и наиболее полезны в этиологических исследованиях (*etiologic research*). Если причинно-следственная связь с воздействием задокументирована, а связи (*associations*) интерпретируются как *эффекты*, полученные оценки могут быть преобразованы из относительного риска в абсолютный, что позволит определить возможное влияние политики в области общественного здравоохранения (см. пункт 16С) [181]. Однако в этом контексте авторам следует приводить достаточно убедительные доводы [171] и с большой осторожностью подходить к выбору концепции и методов анализа данных, подходящих к конкретной ситуации.

17. Другие анализы. Если проводились другие виды анализа, например анализы в подгруппах, взаимодействий, чувствительности, укажите это.

Пример 1 (см. табл. 9)

Пример 2 (см. табл. 10)

Пояснение

В дополнение к основному анализу в наблюдаемых исследованиях часто используются и другие: анализ в специфических подгруппах, возможных

взаимодействий между факторами риска, расчёт дополнительных рисков (*attributable risks*), анализ чувствительности (*sensitivity analyses*) к смещениям при использовании альтернативных определений изучаемых переменных.

Ведётся обсуждение опасностей, связанных с анализом в подгруппах и множественным анализом в целом [4, 104]. По нашему мнению, слишком велико стремление искать доказательства в связях, обнаруживаемых

Таблица 9. Анализ связи между приёмом оральных контрацептивов, лейденской мутацией фактора V и риском возникновения венозной тромбоэмболии

Лейденская мутация фактора V	Оральные контрацептивы	Кол-во пациентов	Кол-во контролей	Отношение шансов
Да	Да	25	2	34,7
Да	Нет	10	4	6,9
Нет	Да	84	63	3,7
Нет	Нет	36	100	1 (референсная величина)

Таблица из работы J.P. Vandenbroucke и соавт. [182] с изменениями L.D. Botto и соавт. [183].

Таблица 10. Чувствительность отношения частоты сердечно-сосудистых исходов к неизмеренному вмешивающемуся фактору

Распространённость неизмеренного бинарного вмешивающегося фактора в группе воздействия (%)	Распространённость неизмеренного бинарного вмешивающегося фактора в группе сравнения (%)	Отношение частоты для неизмеренного бинарного вмешивающегося фактора	Отношение частоты сильного воздействия (95% ДИ)*
90	10	1,5	1,20 (1,01–1,42)
90	50	1,5	1,43 (1,22–1,67)
50	10	1,5	1,39 (1,18–1,63)
90	10	2	0,96 (0,81–1,13)
90	50	2	1,27 (1,11–1,45)
50	10	2	1,21 (1,03–1,42)
90	50	3	1,18 (1,01–1,38)
50	10	3	0,99 (0,85–1,16)
90	50	5	1,08 (0,85–1,26)

Примечание. ДИ — доверительный интервал. * Скорректировано с учётом возраста, пола, приёма препаратов, назначенных при болезнях сердечно-сосудистой системы, неизмеренного бинарного вмешивающегося фактора.

Таблица адаптирована из работы L. Wei и соавт. [184].

при анализе в подгруппах, когда результаты исследования в целом свидетельствуют о малом эффекте или полном его отсутствии. С другой стороны, имеет смысл изучить, сохраняется ли обнаруженная при основном анализе связь в заранее определённых подгруппах, особенно когда исследование достаточно велико, и данных по подгруппам достаточно много. Второй часто обсуждаемый вопрос — возникновение интереса к анализу в подгруппах в процессе анализа данных. Такие подгруппы могут стать источником важных находок, но иногда возникают случайно. Некоторые утверждают, что информировать читателя обо всех проведённых анализах в подгруппах не представляется ни возможным, ни целесообразным, поскольку только будущий анализ других данных покажет, в какой степени ранние захватывающие результаты выдержат испытание временем [9]. Мы советуем авторам указывать, какие анализы были запланированы, а какие нет (см. пункты 4, 12B и 20). Это позволит читателям оценить последствия множественного анализа и принять в расчёт место исследования на всём протяжении — от открытия до подтверждения или опровержения.

Третий предмет дискуссий — как оценивать совместный эффект и взаимодействие между факторами риска — с помощью аддитивной или мультипликативной шкал? Или шкала должна определяться подходящей наилучшим образом статистической моделью (см. пункт 12B и вставку 8). Целесообразно описать и отдельные эффекты каждого воздействия, и совместный эффект, если возможно, в таблице, как в первом примере выше [183], или в исследовании I. Martinelli и соавт. [185]. Такая таблица даёт читателю достаточно информации для оценки аддитивного и мультипликативного взаимодействия (см. вставку 8). Доверительные интервалы для отдельных и совместных эффектов помогут читателям судить о силе данных (*strength of the data*). Кроме того, доверительные

интервалы для показателей взаимодействия, таких как избыточный относительный риск (*relative excess risk from interaction*, RERI), относятся к тестам взаимодействия или однородности. Периодически возникающая проблема — использование авторами *p*-значений при сравнении подгрупп, что приводит к ошибочным заявлениям о модификаторе эффекта (*effect modifier*). Например, связь, статистически значимая в одной категории (скажем, у мужчин), но незначимая в другой (у женщин), сама по себе не доказывает модификацию эффекта. Точно так же доверительные интервалы для каждой точечной оценки иногда неправильно используются для вывода о том, что взаимодействие отсутствует, когда интервалы частично перекрываются. Более достоверные выводы можно сделать, если напрямую изучить, различается ли величина связи между подгруппами.

Анализ чувствительности полезен при изучении влияния выбора, сделанного в статистических анализах, или при выяснении устойчивости полученных результатов к отсутствующим данным и возможным систематическим ошибкам (см. пункт 12B). Авторам необходимо определиться, насколько подробно стоит освещать такие анализы в своих работах. Если предпринималось много анализов чувствительности, детальное их описание может не иметь практического смысла. Иногда достаточно сообщить, что такой анализ был проведён, и он согласуется с основными представленными результатами. Большая детальность уместна, если изучается очень важный вопрос или если оценки эффекта значительно варьируют [59, 186].

S.J. Rosock и соавт. обнаружили, что 43 статьи из 73 с отчётами о наблюдательных исследованиях включали результаты анализа, проведённого в подгруппах. Большинство авторов сообщали о различиях между группами, и лишь в 8 статьях была представлена формальная оценка взаимодействия (см. пункт 12B) [4].

Вставка 8. Взаимодействие (модификация эффекта): анализ совместных эффектов

Взаимодействие возникает, если связь (*association*) воздействия с риском заболевания меняется при наличии другого воздействия (*exposure*). Одна из проблем оценки и описания взаимодействий заключается в том, что эффект воздействия можно измерить двумя способами: как относительный риск (или отношение частот в единицах человека-времени) или как разность рисков (или разность частот). Использование относительного риска приводит к мультипликативной модели, в то время как использование разницы рисков соответствует аддитивной модели [187, 188]. Иногда проводится различие между «статистическим взаимодействием», которое может быть отклонением от мультипликативной или аддитивной модели, и «биологическим взаимодействием», которое измеряется отклонением от аддитивной модели [189]. Однако ни аддитивные, ни мультипликативные модели не указывают на конкретный биологический механизм взаимодействия.

Независимо от выбора модели, важно понять, в какой степени совместный эффект двух факторов воздействия отличается от эффектов каждого из них (при отсутствии другого воздействия). HuGENet (Human Genomic Epidemiology Network) предложила схему прозрачного представления отдельных и совместных эффектов, позволяющую оценить различные типы взаимодействия [183]. Данные исследования о пероральных контрацептивах и лейденской мутации фактора V [182] были использованы для объяснения этого предложения, и этот пример также используется в пункте 17. Каждый из двух представленных факторов в отдельности повышает риск возникновения

венозного тромбоза. Вычисление их отдельных и совместных эффектов приводится в таблице 2×4 (см. пример 1 в пункте 17), где отношение шансов со значением 1 означает исходный уровень для женщин без лейденской мутации, не принимающих оральные контрацептивы.

Сложность состоит в том, что некоторые дизайны, такие как исследование «случай–контроль», и некоторые статистические модели (логистическая регрессия или модель пропорциональных рисков Кокса) оценивают относительные риски (или отношение частот) и, по сути, ведут к мультипликативному моделированию. В этом случае относительные риски можно вычислить по аддитивной шкале. В примере 1 к пункту 17 отношения шансов для каждого фактора равны 3,7 и 6,9, при одновременном действии двух факторов шансы равны 34,7. Если анализировать эти данные в рамках мультипликативной модели, ожидается, что общее отношение шансов составит 25,7 ($3,7 \times 6,9$). Наблюдаемый совместный эффект (34,7) в 1,4 раза превышает ожидаемый результат, определяемый по мультипликативной шкале ($34,7/25,7$). Эта величина (1,4) представляет собой отношение шансов мультипликативного взаимодействия (*odds ratio of the multiplicative interaction*). Такое же значение было бы получено при вычислении антилогарифма коэффициента взаимодействия в модели логистической регрессии. Согласно аддитивной шкале, отношение шансов при одновременном действии двух факторов составит 9,6 ($3,7 + 6,9 - 1$). Наблюдаемый совместный эффект двух факторов значительно отличается от аддитивного эффекта: разница составляет 25,1 ($34,7 - 9,6$). Если отношения шансов интерпретируются как относительные риски (или отношения частот), последняя величина (25,1) представляет собой значение относительного избыточного риска вследствие взаимодействия (*relative excess risk from interaction, RERI*) [190]. Это легче понять, если представить, что референсное значение (отношение шансов равное 1) представляет собой исходную частоту развития венозного тромбоза, т.е., скажем, 1 случай на 10 000 человеко-лет, которая затем увеличивается при наличии отдельных и совместных воздействий.

Обсуждение

Раздел «Обсуждение» посвящён ключевым вопросам достоверности (*validity*) и смысла результатов исследования [191]. Исследования показали, что в этом разделе часто преобладают неполные и предвзятые оценки результатов исследования и их значения, а также риторика, поддерживающая находки авторов [192, 193]. Структурирование раздела «Обсуждение» поможет авторам избежать необоснованных предположений и переоценки результатов и вместе с тем направит читателей по тексту [194, 195]. К примеру, в журнале *Annals of Internal Medicine* [196] приводятся рекомендации по структурированию раздела «Обсуждение» согласно следующей схеме: 1) краткое изложение основных результатов; 2) рассмотрение возможных механизмов и объяснений; 3) сравнение с релевантными результатами других опубликованных исследований; 4) ограничения исследования; 5) краткое обобщение значения работы для практики и науки. Другие авторы внесли похожие предложения [191, 194]. Разделы статьи с рекомендациями для дальнейших исследований и обсуждением ограничений исследования должны быть тесно связаны друг с другом. Исследователи должны предложить способы, используя которые в последующих исследованиях можно достичь лучших результатов, а не вежливо заметить: «необходимы дальнейшие исследования» [197, 198]. Мы предлагаем структурировать раздел «Обсуждение», используя подходящие подзаголовки.

18. Основные результаты. Обобщите основные результаты, ссылаясь на цели исследования.

Пример

«Мы выдвинули гипотезу, согласно которой принадлежность к этническому меньшинству может быть связана с относительно высоким уровнем факторов риска сердечно-сосудистых заболеваний, и эта связь в значительной мере может быть приписана социально-экономическому статусу. Наша гипотеза не подтвердилась. После поправки на возраст и социально-экономический статус сохранялась высокая статистически значимая разница по индексу массы тела, артериальному давлению, наличию сахарного диабета и низкой физической активности между белыми женщинами, афроамериканками и американками мексиканского происхождения. Кроме того, мы обнаружили значительную разницу уровней факторов риска сердечно-сосудистых заболеваний в зависимости от социально-экономического статуса. Эти показатели иллюстрируют высокий риск как для женщин, относящихся к этническим меньшинствам, так и для белых женщин с низким социально-экономическим статусом» [199].

Пояснение

Хорошая практика — начинать обсуждение с резюмирования основных результатов исследования. Резюме напоминает читателям основные находки и помогает им оценить, подкреплены ли дальнейшие интерпретации и разъяснения авторов результатами работы.

19. Ограничения. Обсудите ограничения исследования с анализом источников потенциальных систематических ошибок или неточностей. Обсудите направление действия и значимость потенциальных систематических ошибок.

Пример

«Поскольку распространённость консультирования возрастает с увеличением степени ожирения, мы можем переоценивать истинную распространённость. Телефонный опрос также может приводить к завышению истинной распространённости консультирования. Хотя среди лиц, не имеющих доступа к телефонной связи, уровень ожирения схож с тем, который наблюдается среди тех, у кого такой доступ есть; отсутствие телефона имеет тенденцию быть связанным с более низким уровнем образования. Связь этого фактора с более низким уровнем консультирования продемонстрирована в нашем исследовании. Кроме того, возможно возникновение систематической ошибки из-за отказа некоторых лиц участвовать в опросе или отвечать на вопросы о своей массе тела. Более того, поскольку данные собирали одновременно (*cross-sectionally*), мы не можем сделать вывод, что консультирование предшествовало попыткам пациентов снизить массу тела» [200].

Пояснение

Выявление и обсуждение ограничений исследования — необходимая часть научной публикации. Важно не только выявить источники систематических ошибок и искажений, вызванных действием вмешивающихся факторов (*confounding*), способные повлиять на результаты, но и обсудить относительную важность различных систематических ошибок, в том числе наиболее вероятное направление и величину любых возможных смещений (см. пункт 9 и вставку 3).

Авторам следует также обсудить любые неточности результатов. Неточность возрастает в связи с различными аспектами исследования: размер исследования (*study size*) (пункт 10); измерение воздействия, вмешивающихся факторов и исходов (пункт 8). Неспособность точно измерить истинную величину воздействия ведёт к возникновению смещения в направлении отсутствия эффекта (*bias towards unity*): чем менее точно измерены факторы риска, тем больше систематическая ошибка. Этот эффект описан как «аттенуация» (*'attenuation'*) [201, 202] или недавно как «систематическая ошибка регрессионного разведения» (*'regression dilution bias'*) [203]. Однако если коррелирующие факторы риска измеряются с разной степенью неточности, связанный с ними скорректированный относительный риск может быть смещён как в направлении единицы, так и от неё [204–206].

При обсуждении ограничений авторы могут сравнить свою работу с данными литературы с точки зрения достоверности (*validity*), обобщаемости (*generalisability*) и точности (*precision*). При таком подходе каждое исследование может рассматриваться как вклад в литературу, а не только как самостоятельная основа для выводов и действий [207]. Удивительно, но обсуждение важных ограничений исследования иногда не включается в опубликованные отчёты. Опрос авторов, опубликовавших результаты оригинальных исследований в журнале *The Lancet*, показал, что о важных недостатках исследования

они сообщали в опросниках, но не в опубликованных статьях [192].

20. Интерпретация результатов. Дайте осторожную, но достаточно полную интерпретацию результатов с учётом целей, ограничений, многочисленных анализов, результатов схожих исследований и других уместных свидетельств.

Пример

«Любые объяснения связи между смертью от инфаркта миокарда и приёмом оральных контрацептивов второго поколения должны носить характер гипотез. Не опубликованы данные, свидетельствующие о прямом биологическом механизме такой связи, и не было других эпидемиологических исследований с подобными результатами. <...> Повышение абсолютного риска очень невелико и, возможно, объясняется главным образом курением. Из-за отсутствия прямых доказательств и из-за того, что анализ основан на относительно небольшом количестве данных, необходимо больше доказательств по этому вопросу. Мы не рекомендуем изменять существующую ныне практику назначения лекарственных препаратов на основании наших результатов» [120].

Пояснение

Суть раздела «Обсуждение» — интерпретация результатов исследования. Слишком смелая интерпретация — частое явление, свойственное человеческой природе: даже когда мы прикладываем большие усилия, стараясь дать объективную оценку, рецензенты справедливо указывают нам, что в некоторых аспектах мы зашли слишком далеко. При объяснении полученных результатов авторы должны принимать во внимание характер исследования на протяжении от открытия до его проверки, возможные источники систематических ошибок, в том числе выбывание участников из-под наблюдения и отказ от участия (см. пункты 9, 12 и 19). Должное внимание должно быть уделено влиянию вмешивающихся факторов (пункт 16А), результатам анализов чувствительности, вопросам множественного анализа и анализов в подгруппах (пункт 17). Авторы должны также уделить внимание остаточному искажению (*residual confounding*) с учётом неизмеренных переменных или неточностей измерения вмешивающихся факторов. Например, социально-экономический статус связан со многими важными для здоровья исходами и часто различается между сравниваемыми группами. Переменные, используемые для измерения этого статуса (доход, образование, профессия) — всего лишь суррогаты других, неопределённых и неизмеренных воздействий, и истинный вмешивающийся фактор по определению будет измерен с ошибкой [208]. Авторы должны обратить внимание на действительный диапазон неопределённости оценки (*real range of uncertainty*), который шире, чем статистическая неопределённость (*statistical uncertainty*), отражаемая в доверительном интервале. Последний не принимает в расчёт неточности, связанные

с дизайном исследования, его проведением и методами измерения [209].

A.B. Hill предложил в 1965 г. ряд критериев, которые могут быть полезными для размышлений и выводов о причинных связях [164]. Насколько сильна связь с воздействием? Предшествует ли воздействие развитию заболевания? Наблюдается ли выявленная связь в других исследованиях и условиях? Поддерживаются ли полученные результаты данными экспериментальных исследований, в том числе лабораторных, и исследований на животных? Насколько специфичен вызываемый воздействием эффект и является ли он дозозависимым? Является ли связь биологически правдоподобной? Однако эти критерии не должны применяться механически. Некоторые, например, считают, что относительный риск ниже 2 или 3 необходимо игнорировать [210, 211]. Это суждение противоречит точке зрения, высказанной J. Cornfield и соавт., о значении больших относительных рисков (см. пункт 12B) [127]. Хотя причинная связь более вероятна, если относительный риск составляет 9, из этого не следует, что всё, что ниже 3, неправдоподобно. Например, небольшое повышение риска лейкемии у детей после внутриутробного облучения заслуживает доверия, поскольку объясняет возникновение побочного эффекта медицинской процедуры, для которого нет иных очевидных причин [212]. Более того, канцерогенное влияние радиации считается доказанным. Удвоение риска рака яичников при съедании от 2 до 4 яиц в неделю не должно сразу же считаться убедительным, поскольку диета связана с большим числом факторов риска, обусловленных образом жизни и социально-экономическим статусом [213]. В противоположность этому, правдоподобие широко обсуждаемых эпидемиологических данных о разнице в риске тромбоза между различными типами оральных контрацептивов стало ещё более убедительным при обнаружении различий в коагуляции в рандомизированном перекрёстном испытании (*randomised cross-over trial*) [214]. Всегда следует включать обсуждение уже существующих доказательств, полученных в ходе других исследований, что особенно важно при обнаружении небольшого повышения риска. Авторы должны поместить свою работу в контекст схожих исследований и объяснить, как новые данные влияют на существующий комплекс доказательств, обращаясь, в идеале, к результатам систематического обзора.

21. Обобщаемость. Обсудите обобщаемость (внешнюю достоверность) результатов исследования.

Пример

«Насколько применимы наши оценки к другим ВИЧ-1-инфицированным пациентам? Это важный вопрос, поскольку точность прогностической модели снижается, когда в неё включаются иные данные, отличающиеся от использованных при разработке. Для решения этой проблемы мы ограничили сложность модели и выбрали такую, которая обладает наилучшей обобщаемостью

с учётом когорт, исключённых из процедуры оценки. Наша база данных включала информацию о пациентах из множества стран Европы и Северной Америки, получающих лечение в различных условиях. Спектр пациентов широк: мужчины и женщины, от подростков до пожилых людей. Основные категории воздействия представлены в достаточной мере. Тяжесть иммунодефицита в начале наблюдения варьирует от неподдающейся измерению до очень выраженной, вирусная нагрузка — от неопределимой до очень высокой» [215].

Пояснение

Обобщаемость (*generalisability*), называемая также внешней достоверностью (*external validity*), — мера применимости результатов исследования в других условиях [216]. Она не существует сама по себе: этот термин имеет смысл только в отношении чётко определённых состояний [217]. Могут ли результаты быть применены к индивидуумам, группам или популяциям, отличающимся от участников исследования по полу, возрасту, этнической принадлежности, тяжести заболевания и наличию сопутствующих болезней? Сравнимы ли природа и уровень воздействия, а также определения исходов с теми, что наблюдаются в других условиях или популяциях? Если сбор данных в продолжительных когортных исследованиях (*longitudinal studies*) происходил на протяжении многих лет, остаются ли эти данные актуальными сегодня? Применимы ли результаты исследований системы здравоохранения в одной стране к системам здравоохранения других стран?

Вопрос внешней достоверности результатов исследования часто может быть решён с учётом сведений об условиях исследования, характеристиках участников исследования, воздействиях и исходах. Таким образом, крайне важно, чтобы авторы предоставляли читателям адекватную информацию об условиях и месте проведения исследования, критериях отбора, воздействиях и способах их измерения, об определениях исходов, а также о периоде набора и последующего наблюдения. Степень неучастия и доля не подвергшихся воздействию участников, у которых развились исходы, также важны. Знание абсолютного риска и распространённости воздействия, которая часто варьирует от популяции к популяции, полезно при оценке применимости результатов в других условиях и популяциях (см. вставку 7).

Дополнительная информация

22. Финансирование. Укажите источники финансирования и роль спонсоров в настоящем исследовании и, если применимо, в оригинальном исследовании, на котором основывается представляемая статья.

Пояснение

Некоторые журналы требуют от авторов раскрытия информации о наличии или отсутствии финансовых и иных конфликтов интересов [100, 218]. В нескольких работах продемонстрирована чёткая связь между

источником финансирования и выводами научных статей [219–222]. Если рандомизированное испытание финансируется коммерческой организацией, в выводах таких исследований значительно чаще (отношение шансов 5,3) рекомендуется использовать тестируемое лекарственное средство как препарат выбора [223]. Другие работы подтвердили влияние табачной и телекоммуникационной индустрии на финансируемые ими исследования [224–227]. Имеются также примеры неуместного влияния на выводы исследований спонсоров, в роли которых выступали правительства или некоммерческие организации.

Авторы или спонсоры могут иметь конфликт интересов, который может влиять на дизайн исследования [228], выбор воздействий [228, 229], исходов [230], статистических методов [231], приводить к избирательной публикации информации об исходах [230] и исследований в целом [232]. Следовательно, роль спонсоров должна быть детально описана: к какой части исследования они имели непосредственное отношение (например, выбор дизайна, сбор данных, анализ результатов, написание рукописи, принятие решения о публикации) [100]. Другие источники неуместного влияния включают работодателей (например, администрация университета для аффилированных научных работников и правительственные контролирующие органы, в частности политические деятели для исследователей из государственных научных организаций), надзирающие органы, противники в судебной тяжбе, особые группы заинтересованных лиц.

ЗАКЛЮЧЕНИЕ

Цель STROBE — дать полезные рекомендации по составлению отчётов о наблюдательных исследованиях. Надлежащая отчётность выявляет сильные и слабые стороны исследования и способствует правильной интерпретации и применению результатов исследования. Рекомендации STROBE также могут быть полезны при планировании наблюдательных исследований и служить руководством для рецензентов и редакторов при работе над рукописью.

Наша пояснительная статья подчёркивает важность прозрачного и полного представления результатов исследований, включает в себя обоснование различных пунктов, включённых в контрольный перечень, а также примеры из опубликованных статей, иллюстрирующие нашу точку зрения о том, что мы считаем качественными отчётами. Мы надеемся, что представленный здесь материал поможет авторам и редакторам в дальнейшем применении рекомендаций STROBE.

Мы подчёркиваем, что STROBE и другие рекомендации по описанию исследований [13, 233, 234] — документы, требующие постоянной оценки, уточнений и, если необходимо, изменений [235, 236]. Например,

руководство CONSORT для описания рандомизированных исследований, проводимых в параллельных группах, было впервые представлено в середине 1990-х годов [237]. С тех пор члены группы регулярно собираются для рассмотрения необходимости пересмотра рекомендаций; в 2001 г. появилась их обновлённая версия [233], и в настоящее время ещё одна версия находится на стадии разработки. Аналогичным образом, принципы, представленные в этой статье и в контрольном перечне STROBE, могут меняться по мере накопления новых доказательств и критических замечаний. Веб-сайт STROBE (<http://www.strobe-statement.org/>) предоставляет форум для обсуждения предложений по совершенствованию рекомендаций, настоящего пояснительного документа и другой дополнительной информации, способствующей повышению качества отчётов об эпидемиологических исследованиях.

Некоторые журналы в своих инструкциях для авторов просят следовать рекомендациям STROBE (актуальный список см. на <http://www.strobe-statement.org/>). Мы предлагаем другим журналам включить в инструкции для авторов руководство STROBE и сообщить об этом, связавшись с нами через наш веб-сайт. Журналы, публикующие рекомендации STROBE, предоставляют свободный доступ для читателей и всего биомедицинского сообщества.

ДОПОЛНИТЕЛЬНО

Источники финансирования. Первую рабочую встречу в рамках инициативы STROBE финансировал Европейский научный фонд (ESF). Дополнительное финансирование получено от Medical Research Council Health Services Research Collaboration и National Health Services Research & Development Methodology Programme. Спонсоры не влияли на планирование работы, сбор и анализ данных, подготовку рукописи или решение о её публикации.

Конфликт интересов. Авторы заявляют об отсутствии конкурирующих интересов.

Вклад авторов. Все авторы внесли вклад в написание статьи. Jan P. Vandenbroucke, Erik von Elm, Douglas G. Altman, Peter C. Gøtzsche, Stuart J. Pocock, Matthias Egger — подготовка разделов первой версии рукописи; Erik von Elm — координация инициативы STROBE; Matthias Egger инициировал STROBE и совместно с Erik von Elm организовал первый семинар для обсуждения STROBE.

Участники инициативы STROBE. Следующие лица внесли вклад в создание и совершенствование руководства STROBE: Douglas G. Altman, Maria Blettner, Paolo Boffetta, Hermann Brenner, Genevieve Chene, Cyrus Cooper, George Davey-Smith, Erik von Elm, Matthias Egger, France Gagnon, Peter C. Gøtzsche, Philip Greenland, Sander Greenland, Claire Infante-Rivard, John Ioannidis, Astrid James, Giselle Jones, Bruno Ledergerber, Julian Little, Margaret May, David Moher, Hooman Momen, Alfredo Morabia, Hal Morgenstern, Cynthia D. Mulrow, Fred Paccaud, Stuart J. Pocock, Charles Poole, Martin Roosli, Dietrich Rothenbacher, Kenneth Rothman, Caroline Sabin, Willi Sauerbrei, Lale Say, James J. Schlesselman,

Jonathan Sterne, Holly Syddall, Jan P. Vandenbroucke, Ian White, Susan Wieland, Hywel Williams, Guang Yong Zou.

Благодарности. Мы благодарны Gerd Antes, Kay Dickersin, Shah Ebrahim, Richard Lilford за поддержку инициативы STROBE. Мы благодарны следующим институтам, на базе которых мы провели рабочие встречи: Institute of Social and Preventive Medicine, University of Bern (Берн, Швейцария); Department

of Social Medicine, University of Bristol (Бристоль, Великобритания); London School of Hygiene & Tropical Medicine (Лондон, Великобритания); Nordic Cochrane Centre (Копенгаген, Дания); Centre for Statistics in Medicine (Оксфорд, Великобритания). Мы также благодарим четырёх анонимных рецензентов, которые сопроводили рабочую версию этой статьи полезными комментариями.

СПИСОК ЛИТЕРАТУРЫ

1. Glasziou P., Vandenbroucke J.P., Chalmers I. Assessing the quality of research // *BMJ*. 2004. Vol. 328, N 7430. P. 39–41. doi: 10.1136/bmj.328.7430.39
2. Funai E.F., Rosenbush E.J., Lee M.J., Del Priore G. Distribution of study designs in four major US journals of obstetrics and gynecology // *Gynecol Obstet Invest*. 200. Vol. 151, N 1. P. 8–11. doi: 10.1159/000052882
3. Scales C.D., Norris R.D., Peterson B.L., et al. Clinical research and statistical methods in the urology literature // *J Urol*. 2005. Vol. 174, N (4 Pt 1). P. 1374–1379. doi: 10.1097/01.ju.0000173640.91654.b5
4. Pocock S.J., Collier T.J., Dandreo K.J., et al. Issues in the reporting of epidemiological studies: a survey of recent practice // *BMJ*. 2004. Vol. 329, N 7471. P. 883. doi: 10.1136/bmj.38250.571088.55
5. Tooth L., Ware R., Bain C., et al. Quality of reporting of observational longitudinal research // *Am J Epidemiol*. 2005. Vol. 161, N 3. P. 280–288. doi: 10.1093/aje/kwi042
6. Von Elm E., Altman D.G., Egger M., et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies // *Epidemiology*. 2007. Vol. 18, N 6. P. 800–804. doi: 10.1097/EDE.0b013e3181577654
7. Mihailovic A., Bell C.M., Urbach D.R. Users' guide to the surgical literature. Case-control studies in surgical journals // *Can J Surg*. 2005. Vol. 48, N 2. P. 148–151.
8. Rushton L. Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods // *Occup Environ Med*. 2000. Vol. 57, N 1. P. 1–9. doi: 10.1136/oem.57.1.1
9. Rothman K.J. No adjustments are needed for multiple comparisons // *Epidemiology*. 1990. Vol. 1, N 1. P. 43–46. doi: 10.1097/00001648-199001000-00010
10. Moonesinghe R., Khoury M.J., Janssens A.C. Most published research findings are false-but a little replication goes a long way // *PLoS Med*. 2007. Vol. 4, N 2. P. e28. doi: 10.1371/journal.pmed.0040028
11. Jenicek M. Clinical Case Reporting. Evidence-Based Medicine. Oxford: Butterworth-Heinemann; 1999. 117 p.
12. Vandenbroucke J.P. In defense of case reports and case series // *Ann Intern Med*. 2001. Vol. 134, N 4. P. 330–334. doi: 10.7326/0003-4819-134-4-200102200-00017
13. Bossuyt P.M., Reitsma J.B., Bruns D.E., et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative // *Ann Intern Med*. 2003. Vol. 138, N 1. P. 40–44. doi: 10.7326/0003-4819-138-1-200301070-00010
14. McShane L.M., Altman D.G., Sauerbrei W., et al. REporting recommendations for tumour MARKer prognostic studies (REMARK) // *Br J Cancer*. 2005. Vol. 93, N 4. P. 387–391. doi: 10.1038/sj.bjc.6602678
15. Ioannidis J.P., Gwinn M., Little J., et al. A road map for efficient and reliable human genome epidemiology // *Nat Genet*. 2006. Vol. 38, N 1. P. 3–5. doi: 10.1038/ng0106-3
16. Rodrigues L., Kirkwood B.R. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls // *Int J Epidemiol*. 1990. Vol. 19, N 1. P. 205–213. doi: 10.1093/ije/19.1.205
17. Rothman K.J., Greenland S. Case-Control Studies. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincott Raven, 1998. pp. 93–114.
18. Forand S.P. Leukaemia incidence among workers in the shoe and boot manufacturing industry: a case-control study // *Environ Health*. 2004. Vol. 3, N 1. P. 7. doi: 10.1186/1476-069X-3-7
19. Benson K., Hartz A.J. A comparison of observational studies and randomized, controlled trials // *N Engl J Med*. 2000. Vol. 342, N 25. P. 1878–1886. doi: 10.1056/NEJM200006223422506
20. Gøtzsche P.C., Harden A. Searching for non-randomised studies. Draft chapter 3. Cochrane Non-Randomised Studies Methods Group, 2002. Available from: <http://www.cochrane.dk/nrsmg>. Accessed 10 September 2007.
21. Lohse N., Hansen A.B., Pedersen G., et al. Survival of persons with and without HIV infection in Denmark, 1995–2005 // *Ann Intern Med*. 2007. Vol. 146, N 2. P. 87–95. doi: 10.7326/0003-4819-146-2-200701160-00003
22. American Journal of Epidemiology. 2007. Information for authors. Available from: http://www.oxfordjournals.org/aje/for_authors/index.html. Accessed 10 September 2007.
23. Haynes R.B., Mulrow C.D., Huth E.J., et al. More informative abstracts revisited // *Ann Intern Med*. 1990. Vol. 113, N 1. P. 69–76. doi: 10.7326/0003-4819-113-1-69
24. Taddio A., Pain T., Fassos F.F., et al. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association // *CMAJ*. 1994. Vol. 150, N 10. P. 1611–1615.
25. Hartley J., Sydes M. Which layout do you prefer? An analysis of readers' preferences for different typographic layouts of structured abstracts // *J Inform Sci*. 1996. Vol. 22, N 1. P. 27–37. doi: 10.1177/016555159602200103
26. Viner R.M., Cole T.J. Adult socioeconomic, educational, social, and psychological outcomes of childhood obesity: a national birth cohort study // *BMJ*. 2005. Vol. 330, N 7504. P. 1354. doi: 10.1136/bmj.38453.422049.E0
27. McCauley J., Kern D.E., Kolodner K., et al. The "battering syndrome": prevalence and clinical characteristics of domestic violence in primary care internal medicine practices // *Ann Intern Med*. 1995. Vol. 123, N 10. P. 737–746. doi: 10.7326/0003-4819-123-10-199511150-00001
28. McEvoy S.P., Stevenson M.R., McCartt A.T., et al. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study // *BMJ*. 2005. Vol. 331, N 7514. P. 428. doi: 10.1136/bmj.38537.397512.55

29. Vandenbroucke J.P. Prospective or retrospective: what's in a name? // *BMJ*. 1991. Vol. 302, N 6771. P. 249–250. doi: 10.1136/bmj.302.6771.249
30. Last J.M. *A Dictionary of Epidemiology*. New York: Oxford University Press, 2000.
31. Miettinen O.S. *Theoretical Epidemiology: principles of occurrence research in medicine*. New York: Wiley, 1985. P. 64–66.
32. Rothman K.J., Greenland S. Types of Epidemiologic Studies. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed. Lippincott Raven; 1998. P. 74–75.
33. MacMahon B., Trichopoulos D. *Epidemiology, principles and methods*. 2nd ed. Boston: Little, Brown; 1996. 81 p. doi: 10.1016/S0033-3506(97)00047-4
34. Lilienfeld A.M. *Foundations of Epidemiology*. New York: Oxford University Press, 1976.
35. Ridker P.M., Hennekens C.H., Lindpaintner K., et al. Mutation in the gene coding for coagulation factor V and the risk of myocardial infarction, stroke, and venous thrombosis in apparently healthy men // *N Engl J Med*. 1995. Vol. 332, N 14. P. 912–917. doi: 10.1056/NEJM199504063321403
36. Goodman K.J., O'Rourke K., Day R.S., et al. Dynamics of *Helicobacter pylori* infection in a US-Mexico cohort during the first two years of life // *Int J Epidemiol*. 2005. Vol. 34, N 6. P. 1348–1355. doi: 10.1093/ije/dyi152
37. Altman D.G., De Stavola B.L., Love S.B., Stepniowska K.A. Review of survival analyses published in cancer journals // *Br J Cancer*. 1995. Vol. 72, N 2. P. 511–518. doi: 10.1038/bjc.1995.364
38. Cerhan J.R., Wallace R.B., Folsom A.R., et al. Transfusion history and cancer risk in older women // *Ann Intern Med*. 1993. Vol. 119, N 1. P. 8–15. doi: 10.7326/0003-4819-119-1-199307010-00002
39. Freeman L.E., Dennis L.K., Lynch C.F., et al. Toenail arsenic content and cutaneous melanoma in Iowa // *Am J Epidemiol*. 2004. Vol. 160, N 7. P. 679–687. doi: 10.1093/aje/kwh267
40. Canto J.G., Allison J.J., Kiefe C.I., et al. Relation of race and sex to the use of reperfusion therapy in Medicare beneficiaries with acute myocardial infarction // *N Engl J Med*. 2000. Vol. 342, N 15. P. 1094–1100. doi: 10.1056/NEJM200004133421505
41. Metzker-Cotter E., Kletter Y., Avidor B., et al. Long-term serological analysis and clinical follow-up of patients with cat scratch disease // *Clin Infect Dis*. 2003. Vol. 37, N 9. P. 1149–1154. doi: 10.1086/378738
42. Johnson E.S. Bias on withdrawing lost subjects from the analysis at the time of loss, in cohort mortality studies, and in follow-up methods // *J Occup Med*. 1990. Vol. 32, N 3. P. 250–254. doi: 10.1097/00043764-199003000-00013
43. Berkson J. Limitations of the application of fourfold table analysis to hospital data // *Biom Bull*. 1946. Vol. 2, N 3. P. 47. doi: 10.2307/3002000
44. Feinstein A.R., Walter S.D., Horwitz R.I. An analysis of Berkson's bias in case-control studies // *J Chronic Dis*. 1986. Vol. 39, N 7. P. 495–504. doi: 10.1016/0021-9681(86)90194-3
45. Jick H., Vessey M.P. Case-control studies in the evaluation of drug-induced illness // *Am J Epidemiol*. 1978. Vol. 107, N 1. P. 1–7. doi: 10.1093/oxfordjournals.aje.a112502
46. Hackam D.G., Mamdani M., Li P., Redelmeier D.A. Statins and sepsis in patients with cardiovascular disease: a population-based cohort analysis // *Lancet*. 2006. Vol. 367, N 9508. P. 413–418. doi: 10.1016/S0140-6736(06)68041-0
47. Smeeth L., Cook C., Fombonne E., et al. MMR vaccination and pervasive developmental disorders: a case-control study // *Lancet*. 2004. Vol. 364, N 9438. P. 963–969. doi: 10.1016/S0140-6736(04)17020-7
48. Costanza M.C. Matching // *Prev Med*. 1995. Vol. 24, N 5. P. 425–433. doi: 10.1006/pmed.1995.1069
49. Sturmer T., Brenner H. Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case-control studies // *Am J Epidemiol*. 2002. Vol. 155, N 7. P. 593–602. doi: 10.1093/aje/155.7.593
50. Rothman K.J., Greenland S. Matching. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. Lippincott Raven; 1998. P. 147–161.
51. Szklo M.F., Nieto J. *Epidemiology, Beyond the Basics*. Sudbury (MA): Jones and Bartlett; 2000. P. 40–51.
52. Cole P., MacMahon B. Attributable risk percent in case-control studies // *Br J Prev Soc Med*. 1971. Vol. 25, N 4. P. 242–244. doi: 10.1136/jech.25.4.242
53. Gissler M., Hemminki E. The danger of overmatching in studies of the perinatal mortality and birthweight of infants born after assisted conception // *Eur J Obstet Gynecol Reprod Biol*. 1996. Vol. 69, N 2. P. 73–75. doi: 10.1016/0301-2115(95)02517-0
54. Gefeller O., Pfahlerberg A., Brenner H., Windeler J. An empirical investigation on matching in published case-control studies // *Eur J Epidemiol*. 1998. Vol. 14, N 4. P. 321–325. doi: 10.1023/A:1007497104800
55. Artama M., Ritvanen A., Gissler M., et al. Congenital structural anomalies in offspring of women with epilepsy—a population-based cohort study in Finland // *Int J Epidemiol*. 2006. Vol. 35, N 2. P. 280–287. doi: 10.1093/ije/dyi234
56. Ebrahim S. Cohorts, infants and children // *Int J Epidemiol*. 2004. Vol. 33, N 6. P. 1165–1166. doi: 10.1093/ije/dyh368
57. Walker M., Whincup P.H., Shaper A.G. The British Regional Heart Study 1975–2004 // *Int J Epidemiol*. 2004. Vol. 33, N 6. P. 1185–1192. doi: 10.1093/ije/dyh295
58. Wieland S., Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives // *J Clin Epidemiol*. 2005. Vol. 58, N 6. P. 560–567. doi: 10.1016/j.jclinepi.2004.11.018
59. Anderson H.R., Atkinson R.W., Peacock J.L., et al. Ambient particulate matter and health effects: publication bias in studies of short-term associations // *Epidemiology*. 2005. Vol. 16, N 2. P. 155–163. doi: 10.1097/01.ede.0000152528.22746.0f
60. Winkelmayr W.C., Stampfer M.J., Willett W.C., Curhan G.C. Habitual caffeine intake and the risk of hypertension in women // *JAMA*. 2005. Vol. 294, N 18. P. 2330–2335. doi: 10.1001/jama.294.18.2330
61. Lukanova A., Soderberg S., Kaaks R., et al. Serum adiponectin is not associated with risk of colorectal cancer // *Cancer Epidemiol Biomarkers Prev*. 2006. Vol. 15, N 2. P. 401–402. doi: 10.1158/1055-9965.EPI-05-0836
62. Becher H. The concept of residual confounding in regression models and some applications // *Stat Med*. 1992. Vol. 11, N 13. P. 1747–1758. doi: 10.1002/sim.4780111308
63. Brenner H., Blettner M. Controlling for continuous confounders in epidemiologic research // *Epidemiology*. 1997. Vol. 8, N 4. P. 429–434. doi: 10.1097/00001648-199707000-00014
64. Phillips M.R., Yang G., Zhang Y., et al. Risk factors for suicide in China: a national case-control psychological autopsy study // *Lancet*. 2002. Vol. 360, N 9347. P. 1728–1736. doi: 10.1016/S0140-6736(02)11681-3

65. Pasquale L.R., Kang J.H., Manson J.E., et al. Prospective study of type 2 diabetes mellitus and risk of primary open-angle glaucoma in women // *Ophthalmology*. 2006. Vol. 113, N 7. P. 1081–1086. doi: 10.1016/j.ophtha.2006.01.066
66. Craig S.L., Feinstein A.R. Antecedent therapy versus detection bias as causes of neoplastic multimorbidity // *Am J Clin Oncol*. 1999. Vol. 22, N 1. P. 51–56. doi: 10.1097/00000421-199902000-00013
67. Rogler L.H., Mroczek D.K., Fellows M., Loftus S.T. The neglect of response bias in mental health research // *J Nerv Ment Dis*. 2001. Vol. 189, N 3. P. 182–187. doi: 10.1097/00005053-200103000-00007
68. Murphy E.A. The logic of medicine. Baltimore: Johns Hopkins University Press, 1976.
69. Sackett D.L. Bias in analytic research // *J Chronic Dis*. 1979. Vol. 32, N 1–2. P. 51–63. doi: 10.1016/0021-9681(79)90012-2
70. Johannes C.B., Crawford S.L., McKinlay J.B. Interviewer effects in a cohort study. Results from the Massachusetts Women's Health Study // *Am J Epidemiol*. 1997. Vol. 146, N 5. P. 429–438. doi: 10.1093/oxfordjournals.aje.a009296
71. Bloemenkamp K.W., Rosendaal F.R., Buller H.R., et al. Risk of venous thrombosis with use of current low-dose oral contraceptives is not explained by diagnostic suspicion and referral bias // *Arch Intern Med*. 1999. Vol. 159, N 1. P. 65–70. doi: 10.1001/archinte.159.1.65
72. Feinstein A.R. Clinical epidemiology: the architecture of clinical research. Philadelphia: W.B. Saunders, 1985.
73. Yadon Z.E., Rodrigues L.C., Davies C.R., Quigley M.A. Indoor and peridomestic transmission of American cutaneous leishmaniasis in northwestern Argentina: a retrospective case-control study // *Am J Trop Med Hyg*. 2003. Vol. 68, N 5. P. 519–526. doi: 10.4269/ajtmh.2003.68.519
74. Anoop S., Saravanan B., Joseph A., et al. Maternal depression and low maternal intelligence as risk factors for malnutrition in children: a community based case-control study from South India // *Arch Dis Child*. 2004. Vol. 89, N 4. P. 325–329. doi: 10.1136/adc.2002.009738
75. Carlin J.B., Doyle L.W. Sample size // *J Paediatr Child Health*. 2002. Vol. 38, N 3. P. 300–304. doi: 10.1046/j.1440-1754.2002.00855.x
76. Rigby A.S., Vail A. Statistical methods in epidemiology. II: A commonsense approach to sample size estimation // *Disabil Rehabil*. 1998. Vol. 20, N 11. P. 405–410. doi: 10.3109/09638289809166102
77. Schulz K.F., Grimes D.A. Sample size calculations in randomised trials: mandatory and mystical // *Lancet*. 2005. Vol. 365, N 9467. P. 1348–1353. doi: 10.1016/S0140-6736(05)61034-3
78. Drescher K., Timm J., Jockel K.H. The design of case-control studies: the effect of confounding on sample size requirements // *Stat Med*. 1990. Vol. 9, N 7. P. 765–776. doi: 10.1002/sim.4780090706
79. Devine O.J., Smith J.M. Estimating sample size for epidemiologic studies: the impact of ignoring exposure measurement uncertainty // *Stat Med*. 1998. Vol. 17, N 12. P. 1375–1389. doi: 10.1002/(SICI)1097-0258(19980630)17:12<1375::AID-SIM857>3.0.CO;2-D
80. Linn S., Levi L., Grunau P.D., et al. Effect measure modification and confounding of severe head injury mortality by age and multiple organ injury severity // *Ann Epidemiol*. 2007. Vol. 17, N 2. P. 142–147. doi: 10.1016/j.annepidem.2006.08.004
81. Altman D.G., Lausen B., Sauerbrei W., Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors // *J Natl Cancer Inst*. 1994. Vol. 86, N 11. P. 829–835. doi: 10.1093/jnci/86.11.829
82. Royston P., Altman D.G., Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea // *Stat Med*. 2006. Vol. 25, N 1. P. 127–141. doi: 10.1002/sim.2331
83. Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis // *Epidemiology*. 1995. Vol. 6, N 4. P. 450–454. doi: 10.1097/00001648-199507000-00025
84. Royston P., Ambler G., Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology // *Int J Epidemiol*. 1999. Vol. 28, N 5. P. 964–974. doi: 10.1093/ije/28.5.964
85. MacCallum R.C., Zhang S., Preacher K.J., Rucker D.D. On the practice of dichotomization of quantitative variables // *Psychol Methods*. 2002. Vol. 7, N 1. P. 19–40. doi: 10.1037/1082-989X.7.1.19
86. Altman D.G. Categorizing continuous variables. In: Armitage P, Colton T, editors. *Encyclopedia of biostatistics*. 2nd ed. Chichester: John Wiley; 2005. P. 708–711. doi: 10.1002/0470011815.b2a10012
87. Cohen J. The cost of dichotomization // *Applied Psychological Measurement*. 1983. Vol. 7, N 3. P. 249–253. doi: 10.1177/014662168300700301
88. Zhao L.P., Kolonel L.N. Efficiency loss from categorizing quantitative exposures into qualitative exposures in case-control studies // *Am J Epidemiol*. 1992. Vol. 136, N 4. P. 464–474. doi: 10.1093/oxfordjournals.aje.a116520
89. Cochran W.G. The effectiveness of adjustment by subclassification in removing bias in observational studies // *Biometrics*. 1968. Vol. 24, N 2. P. 295–313. doi: 10.2307/2528036
90. Clayton D., Hills M. Models for dose-response (Chapter 25). *Statistical Models in Epidemiology*. Oxford: Oxford University Press, 1993. P. 249–260.
91. Cox D.R. Note on grouping // *J Am Stat Assoc*. 1957. Vol. 52, N 280. P. 543–547. doi: 10.1080/01621459.1957.10501411
92. Il'yasova D., Hertz-Picciotto I., Peters U., et al. Choice of exposure scores for categorical regression in meta-analysis: a case study of a common problem // *Cancer Causes Control*. 2005. Vol. 16, N 4. P. 383–388. doi: 10.1007/s10552-004-5025-x
93. Berglund A., Alfredsson L., Cassidy J.D., et al. The association between exposure to a rear-end collision and future neck or shoulder pain: a cohort study // *J Clin Epidemiol*. 2000. Vol. 53, N 11. P. 1089–1094. doi: 10.1016/S0895-4356(00)00225-0
94. Slama R., Werwatz A. Controlling for continuous confounding factors: non- and semiparametric approaches // *Rev Epidemiol Sante Publique*. 2005. Vol. 53, N Spec No 2. P. 2S65–80. doi: 10.1016/S0398-7620(05)84769-8
95. Greenland S. Introduction to regression modelling (Chapter 21). In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed: Lippincott Raven, 1998. P. 401–432.
96. Thompson W.D. Statistical analysis of case-control studies // *Epidemiol Rev*. 1994. Vol. 16, N 1. P. 33–50. doi: 10.1093/oxfordjournals.epirev.a036143
97. Schlesselman J.J. Logistic regression for case-control studies (Chapter 8.2). *Case-control studies Design, conduct, analysis*. New York, Oxford: Oxford University Press, 1982. P. 235–241.
98. Clayton D., Hills M. Choice and interpretation of models (Chapter 27). *Statistical Models in Epidemiology*. Oxford: Oxford University Press, 1993. P. 271–281.
99. Altman D.G., Gore S.M., Gardner M.J., Pocock S.J. Statistical guidelines for contributors to medical journals // *Br Med J*. 1983. Vol. 286, N 6376. P. 1489–1493. doi: 10.1136/bmj.286.6376.1489

100. International Committee of Medical Journal Editors Uniform requirements for manuscripts submitted to biomedical journals // *N Engl J Med*. 1997. Vol. 336. P. 309–315. [Electronic version updated February 2006, Available from at <http://www.icmje.org/>. doi: 10.1056/NEJM199701233360422
101. Mullner M., Matthews H., Altman D.G. Reporting on statistical methods to adjust for confounding: a cross-sectional survey // *Ann Intern Med*. 2002. Vol. 136, N 2. P. 122–126. doi: 10.7326/0003-4819-136-2-200201150-00009
102. Olsen J., Basso O. Re: Residual confounding // *Am J Epidemiol*. 1999. Vol. 149, N 3. P. 290. doi: 10.1093/oxfordjournals.aje.a009805
103. Hallan S., de Mutsert R., Carlsen S., et al. Obesity, smoking, and physical inactivity as risk factors for CKD: are men more vulnerable? // *Am J Kidney Dis*. 2006. Vol. 47, N 3. P. 396–405. doi: 10.1053/j.ajkd.2005.11.027
104. Gotzsche P.C. Believability of relative risks and odds ratios in abstracts: cross sectional study // *BMJ*. 2006. Vol. 333, N 7561. P. 231–234. doi: 10.1136/bmj.38895.410451.79
105. Szklo M.F., Nieto J. Communicating Results of Epidemiologic Studies (Chapter 9). *Epidemiology, Beyond the Basics*. Sudbury (MA): Jones and Bartlett, 2000. P. 408–430.
106. Chandola T., Brunner E., Marmot M. Chronic stress at work and the metabolic syndrome: prospective study // *BMJ*. 2006. Vol. 332, N 7540. P. 521–525. doi: 10.1136/bmj.38693.435301.80
107. Vach W., Blettner M. Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables // *Am J Epidemiol*. 1991. Vol. 134, N 8. P. 895–907. doi: 10.1093/oxfordjournals.aje.a116164
108. Little R.J., Rubin D.B. A taxonomy of missing-data methods (Chapter 1.4.). *Statistical Analysis with Missing Data*. New York: Wiley, 2002. P. 19–23. doi: 10.1002/9781119013563
109. Ware J.H. Interpreting incomplete data in studies of diet and weight loss // *N Engl J Med*. 2003. Vol. 348, N 21. P. 2136–2137. doi: 10.1056/NEJMe030054
110. Rubin D.B. Inference and missing data // *Biometrika*. 1976. Vol. 63, N 3. P. 581–592. doi: 10.1093/biomet/63.3.581
111. Schafer J.L. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall, 1997. doi: 10.1201/9781439821862
112. Lipsitz S.R., Ibrahim J.G., Chen M.H., Peterson H. Non-ignorable missing covariates in generalized linear models // *Stat Med*. 1999. Vol. 18, N 17–18. P. 2435–2448. doi: 10.1002/(SICI)1097-0258(19990915/30)18:17/18<2435::AID-SIM267>3.0.CO;2-B
113. Rotnitzky A., Robins J. Analysis of semi-parametric regression models with non-ignorable non-response // *Stat Med*. 1997. Vol. 16, N 1–3. P. 81–102. doi: 10.1002/(SICI)1097-0258(19970115)16:1<81::AID-SIM473>3.0.CO;2-0
114. Rubin D.B. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley; 1987. doi: 10.1002/9780470316696
115. Barnard J., Meng X.L. Applications of multiple imputation in medical studies: from AIDS to NHANES // *Stat Methods Med Res*. 1999. Vol. 8, N 1. P. 17–36. doi: 10.1177/096228029900800103
116. Braitstein P., Brinkhof M.W., Dabis F., et al. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries // *Lancet*. 2006. Vol. 367, N 9513. P. 817–824. doi: 10.1016/S0140-6736(06)68337-2
117. Purandare N., Burns A., Daly K.J., et al. Cerebral emboli as a potential cause of Alzheimer's disease and vascular dementia: case-control study // *BMJ*. 2006. Vol. 332, N 7550. P. 1119–1124. doi: 10.1136/bmj.38814.696493.AE
118. Steyn K., Gaziano T.A., Bradshaw D., et al. Hypertension in South African adults: results from the Demographic and Health Survey, 1998 // *J Hypertens*. 2001. Vol. 19, N 10. P. 1717–1725. doi: 10.1097/00004872-200110000-00004
119. Lohr S.L. *Design Effects* (Chapter 7.5). *Sampling: Design and Analysis*. Pacific Grove (CA): Duxbury Press, 1999.
120. Dunn N.R., Arscott A., Thorogood M. The relationship between use of oral contraceptives and myocardial infarction in young women with fatal outcome, compared to those who survive: results from the MICA case-control study // *Contraception*. 2001. Vol. 63, N 2. P. 65–69. doi: 10.1016/S0010-7824(01)00172-X
121. Rothman K.J., Greenland S. *Basic Methods for Sensitivity Analysis and External Adjustment*. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed. Lippincott Raven; 1998. P. 343–357.
122. Custer B., Longstreth W.T., Phillips L.E., et al. Hormonal exposures and the risk of intracranial meningioma in women: a population-based case-control study // *BMC Cancer*. 2006. Vol. 6. P. 152. doi: 10.1186/1471-2407-6-152
123. Wakefield M.A., Chaloupka F.J., Kaufman N.J., et al. Effect of restrictions on smoking at home, at school, and in public places on teenage smoking: cross sectional study // *BMJ*. 2000. Vol. 321, N 7257. P. 333–337. doi: 10.1136/bmj.321.7257.333
124. Greenland S. The impact of prior distributions for uncontrolled confounding and response bias: a case study of the relation of wire codes and magnetic fields to childhood leukemia // *J Am Stat Assoc*. 2003. Vol. 98, N 461. P. 47–54. doi: 10.1198/01621450338861905
125. Lash T.L., Fink A.K. Semi-automated sensitivity analysis to assess systematic errors in observational data // *Epidemiology*. 2003. Vol. 14, N 4. P. 451–458. doi: 10.1097/01.EDE.0000071419.41011.cf
126. Phillips C.V. Quantifying and reporting uncertainty from systematic errors // *Epidemiology*. 2003. Vol. 14, N 4. P. 459–466. doi: 10.1097/01.ede.0000072106.65262.ae
127. Cornfield J., Haenszel W., Hammond E.C., et al. Smoking and lung cancer: recent evidence and a discussion of some questions // *J Natl Cancer Inst*. 1959. Vol. 22, N 1. P. 173–203.
128. Langholz B. Factors that explain the power line configuration wiring code-childhood leukemia association: what would they look like? // *Bioelectromagnetics*. 2001. Suppl 5. P. S19–31. doi: 10.1002/1521-186x(2001)22:5+::aid-bem1021>3.3.co;2-9
129. Eisner M.D., Smith A.K., Blanc P.D. Bartenders' respiratory health after establishment of smoke-free bars and taverns // *JAMA*. 1998. Vol. 280, N 22. P. 1909–1914. doi: 10.1001/jama.280.22.1909
130. Dunne M.P., Martin N.G., Bailey J.M., et al. Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders // *Int J Epidemiol*. 1997. Vol. 26, N 4. P. 844–854. doi: 10.1093/ije/26.4.844
131. Schuz J., Kaatsch P., Kaletsch U., et al. Association of childhood cancer with factors related to pregnancy and birth // *Int J Epidemiol*. 1999. Vol. 28, N 4. P. 631–639. doi: 10.1093/ije/28.4.631
132. Cnattingius S., Zack M., Ekblom A., et al. Prenatal and neonatal risk factors for childhood myeloid leukemia // *Cancer Epidemiol Biomarkers Prev*. 1995. Vol. 4, N 5. P. 441–445.
133. Schuz J. Non-response bias as a likely cause of the association between young maternal age at the time of delivery and the risk of cancer in the offspring // *Paediatr Perinat Epidemiol*. 2003. Vol. 17, N 1. P. 106–112. doi: 10.1046/j.1365-3016.2003.00460.x

134. Slattery M.L., Edwards S.L., Caan B.J., et al. Response rates among control subjects in case-control studies // *Ann Epidemiol.* 1995. Vol. 5, N 3. P. 245–249. doi: 10.1016/1047-2797(94)00113-8
135. Schulz K.F., Grimes D.A. Case-control studies: research in reverse // *Lancet.* 2002. Vol. 359, N 9304. P. 431–434. doi: 10.1016/S0140-6736(02)07605-5
136. Olson S.H., Voigt L.F., Begg C.B., Weiss N.S. Reporting participation in case-control studies // *Epidemiology.* 2002. Vol. 13, N 2. P. 123–126. doi: 10.1097/00001648-200203000-00004
137. Morton L.M., Cahill J., Hartge P. Reporting participation in epidemiologic studies: a survey of practice // *Am J Epidemiol.* 2006. Vol. 163, N 3. P. 197–203. doi: 10.1093/aje/kwj036
138. Olson S.H. Reported participation in case-control studies: changes over time // *Am J Epidemiol.* 2001. Vol. 154, N 6. P. 574–581. doi: 10.1093/aje/154.6.574
139. Sandler D.P. On revealing what we'd rather hide: the problem of describing study participation // *Epidemiology.* 2002. Vol. 13, N 2. P. 117. doi: 10.1097/00001648-200203000-00001
140. Hepworth S.J., Schoemaker M.J., Muir K.R., et al. Mobile phone use and risk of glioma in adults: case-control study // *BMJ.* 2006. Vol. 332, N 7546. P. 883–887. doi: 10.1136/bmj.38720.687975.55
141. Hay A.D., Wilson A., Fahey T., Peters T.J. The duration of acute cough in pre-school children presenting to primary care: a prospective cohort study // *Fam Pract.* 2003. Vol. 20, N 6. P. 696–705. doi: 10.1093/fampra/cm613
142. Egger M., Juni P., Bartlett C. Value of flow diagrams in reports of randomized controlled trials // *JAMA.* 2001. Vol. 285, N 15. P. 1996–1999. doi: 10.1001/jama.285.15.1996
143. Osella A.R., Misciagna G., Guerra V.M., et al. Hepatitis C virus (HCV) infection and liver-related mortality: a population-based cohort study in southern Italy. The Association for the Study of Liver Disease in Puglia // *Int J Epidemiol.* 2000. Vol. 29, N 5. P. 922–927. doi: 10.1093/ije/29.5.922
144. Dales L.G., Ury H.K. An improper use of statistical significance testing in studying covariables // *Int J Epidemiol.* 1978. Vol. 7, N 4. P. 373–375. doi: 10.1093/ije/7.4.373
145. Maldonado G., Greenland S. Simulation study of confounder-selection strategies // *Am J Epidemiol.* 1993. Vol. 138, N 11. P. 923–936. doi: 10.1093/oxfordjournals.aje.a116813
146. Tanis B.C., van den Bosch M.A., Kemmeren J.M., et al. Oral contraceptives and the risk of myocardial infarction // *N Engl J Med.* 2001. Vol. 345, N 25. P. 1787–1793. doi: 10.1056/NEJMoa003216
147. Rothman K.J., Greenland S. Precision and Validity in Epidemiologic Studies. In: Rothman KJ, Greenland S, editors. *Modern epidemiology.* 2nd ed. Lippincott Raven; 1998. P. 120–125.
148. Clark T.G., Altman D.G., De Stavola B.L. Quantification of the completeness of follow-up // *Lancet.* 2002. Vol. 359, N 9314. P. 1309–1310. doi: 10.1016/S0140-6736(02)08272-7
149. Qiu C., Fratiglioni L., Karp A., et al. Occupational exposure to electromagnetic fields and risk of Alzheimer's disease // *Epidemiology.* 2004. Vol. 15, N 6. P. 687–694. doi: 10.1097/01.ede.0000142147.49297.9d
150. Kengeya-Kayondo J.F., Kamali A., Nunn A.J., et al. Incidence of HIV-1 infection in adults and socio-demographic characteristics of seroconverters in a rural population in Uganda: 1990–1994 // *Int J Epidemiol.* 1996. Vol. 25, N 5. P. 1077–1082. doi: 10.1093/ije/25.5.1077
151. Mastrangelo G., Fedeli U., Fadda E., et al. Increased risk of hepatocellular carcinoma and liver cirrhosis in vinyl chloride workers: synergistic effect of occupational exposure with alcohol intake // *Environ Health Perspect.* 2004. Vol. 112, N 11. P. 1188–1192. doi: 10.1289/ehp.6972
152. Salo P.M., Arbes S.J., Sever M., et al. Exposure to *Alternaria alternata* in US homes is associated with asthma symptoms // *J Allergy Clin Immunol.* 2006. Vol. 118, N 4. P. 892–898. doi: 10.1016/j.jaci.2006.07.037
153. Pocock S.J., Clayton T.C., Altman D.G. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls // *Lancet.* 2002. Vol. 359, N 9318. P. 1686–1689. doi: 10.1016/S0140-6736(02)08594-X
154. Sasieni P. A note on the presentation of matched case-control data // *Stat Med.* 1992. Vol. 11, N 5. P. 617–620. doi: 10.1002/sim.4780110506
155. Lee G.M., Neutra R.R., Hristova L., et al. A nested case-control study of residential and personal magnetic field measures and miscarriages // *Epidemiology.* 2002. Vol. 13, N 1. P. 21–31. doi: 10.1097/00001648-200201000-00005
156. Tiihonen J., Wahlbeck K., Lonnqvist J., et al. Effectiveness of antipsychotic treatments in a nationwide cohort of patients in community care after first hospitalisation due to schizophrenia and schizoaffective disorder: observational follow-up study // *BMJ.* 2006. Vol. 333, N 7561. P. 224. doi: 10.1136/bmj.38881.382755.2F
157. Christenfeld N.J., Sloan R.P., Carroll D., Greenland S. Risk factors, confounding, and the illusion of statistical control // *Psychosom Med.* 2004. Vol. 66, N 6. P. 868–875. doi: 10.1097/01.psy.0000140008.70959.41
158. Smith G.D., Phillips A. Declaring independence: why we should be cautious // *J Epidemiol Community Health.* 1990. Vol. 44, N 4. P. 257–258. doi: 10.1136/jech.44.4.257
159. Greenland S., Neutra R. Control of confounding in the assessment of medical technology // *Int J Epidemiol.* 1980. Vol. 9, N 4. P. 361–367. doi: 10.1093/ije/9.4.361
160. Robins J.M. Data, design, and background knowledge in etiologic inference // *Epidemiology.* 2001. Vol. 12, N 3. P. 313–320. doi: 10.1097/00001648-200105000-00011
161. Sagiv S.K., Tolbert P.E., Altshul L.M., Korrick S.A. Organochlorine exposures during pregnancy and infant size at birth // *Epidemiology.* 2007. Vol. 18, N 1. P. 120–129. doi: 10.1097/01.ede.0000249769.15001.7c
162. World Health Organization. Body Mass Index (BMI). 2007. Available from: http://www.euro.who.int/nutrition/20030507_1
163. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study // *Lancet.* 2003. Vol. 362, N 9382. P. 419–427. doi: 10.1016/S0140-6736(03)14065-2
164. Hill A.B. The environment and disease: Association or causation? // *Proc R Soc Med.* 1965. Vol. 58, N 5. P. 295–300. doi: 10.1177/003591576505800503
165. Vineis P. Causality in epidemiology // *Soz Präventivmed.* 2003. Vol. 48, N 2. P. 80–87. doi: 10.1007/s00038-003-1029-7
166. Empana J.P., Ducimetiere P., Arveiler D., et al. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study // *Eur Heart J.* 2003. Vol. 24, N 21. P. 1903–1911. doi: 10.1016/j.ehj.2003.09.002
167. Tunstall-Pedoe H., Kuulasmaa K., Mahonen M., et al. Contribution of trends in survival and coronary-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA project populations. Monitoring trends and determinants in cardiovascular disease // *Lancet.* 1999. Vol. 353, N 9164. P. 1547–1557. doi: 10.1016/S0140-6736(99)04021-0

168. Cambien F., Chretien J.M., Ducimetiere P., et al. Is the relationship between blood pressure and cardiovascular risk dependent on body mass index? // *Am J Epidemiol.* 1985. Vol. 122, N 3. P. 434–442. doi: 10.1093/oxfordjournals.aje.a114124
169. Hosmer D.W., Taber S., Lemeshow S. The importance of assessing the fit of logistic regression models: a case study // *Am J Public Health.* 1991. Vol. 81, N 12. P. 1630–1635. doi: 10.2105/AJPH.81.12.1630
170. Tibshirani R. A plain man's guide to the proportional hazards model // *Clin Invest Med.* 1982. Vol. 5, N 1. P. 63–68.
171. Rockhill B., Newman B., Weinberg C. Use and misuse of population attributable fractions // *Am J Public Health.* 1998. Vol. 88, N 1. P. 15–19. doi: 10.2105/AJPH.88.1.15
172. Uter W., Pfahlberg A. The application of methods to quantify attributable risk in medical practice // *Stat Methods Med Res.* 2001. Vol. 10, N 3. P. 231–237. doi: 10.1177/096228020101000305
173. Schwartz L.M., Woloshin S., Dvorin E.L., Welch H.G. Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks // *BMJ.* 2006. Vol. 333, N 7581. P. 1248. doi: 10.1136/bmj.38985.564317.7C
174. Nakayama T., Zaman M.M., Tanaka H. Reporting of attributable and relative risks, 1966–97 // *Lancet.* 1998. Vol. 351, N 9110. P. 1179. doi: 10.1016/S0140-6736(05)79123-6
175. Cornfield J. A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix // *J Natl Cancer Ins.* 1951. Vol. 11, N 6. P. 1269–1275.
176. Pearce N. What does the odds ratio estimate in a case-control study? // *Int J Epidemiol.* 1993. Vol. 22, N 6. P. 1189–1192. doi: 10.1093/ije/22.6.1189
177. Rothman K.J., Greenland S., Rothman K.J., Greenland S. Measures of Disease Frequency. *Modern epidemiology.* 2nd ed. Lippincott Raven, 1998. P. 44–45.
178. Doll R., Hill A.B. The mortality of doctors in relation to their smoking habits: a preliminary report. // *BMJ.* 1954. Vol. 1, N 4877. P. 1451–1455. doi: 10.1136/bmj.1.4877.1451
179. Ezzati M., Lopez A.D. Estimates of global mortality attributable to smoking in 2000 // *Lancet.* 2003. Vol. 362, N 9387. P. 847–852. doi: 10.1016/S0140-6736(03)14338-3
180. Greenland S. Applications of Stratified Analysis Methods. In: Rothman KJ, Greenland S, editors. *Modern epidemiology.* 2nd ed. Lippincott Raven, 1998. P. 295–297.
181. Rose G. Sick individuals and sick populations // *Int J Epidemiol.* 2001. Vol. 30, N 3. P. 427–432. discussion. doi: 10.1093/ije/30.3.427
182. Vandenbroucke J.P., Koster T., Briet E., et al. Increased risk of venous thrombosis in oral-contraceptive users who are carriers of factor V Leiden mutation // *Lancet.* 1994. Vol. 344, N 8935. P. 1453–1457. doi: 10.1016/S0140-6736(94)90286-0
183. Botto L.D., Khoury M.J. Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond // *Am J Epidemiol.* 2001. Vol. 153, N 10. P. 1016–1020. doi: 10.1093/aje/153.10.1016
184. Wei L., MacDonald T.M., Walker B.R. Taking glucocorticoids by prescription is associated with subsequent cardiovascular disease // *Ann Intern Med.* 2004. Vol. 141, N 10. P. 764–770. doi: 10.7326/0003-4819-141-10-200411160-00007
185. Martinelli I., Taioli E., Battaglioli T., et al. Risk of venous thromboembolism after air travel: interaction with thrombophilia and oral contraceptives // *Arch Intern Med.* 2003. Vol. 163, N 22. P. 2771–2774. doi: 10.1001/archinte.163.22.2771
186. Kyzas P.A., Loizou K.T., Ioannidis J.P. Selective reporting biases in cancer prognostic factor studies // *J Natl Cancer Inst.* 2005. Vol. 97, N 14. P. 1043–1055. doi: 10.1093/jnci/dji184
187. Rothman K.J., Greenland S., Walker A.M. Concepts of interaction // *Am J Epidemiol.* 1980. Vol. 112, N 4. P. 467–470. doi: 10.1093/oxfordjournals.aje.a113015
188. Saracci R. Interaction and synergism // *Am J Epidemiol.* 1980. Vol. 112, N 4. P. 465–466. doi: 10.1093/oxfordjournals.aje.a113014
189. Rothman K.J. *Epidemiology. An introduction.* Oxford: Oxford University Press, 2002. P. 168–180.
190. Rothman K.J. *Interactions Between Causes. Modern epidemiology.* Boston: Little Brown, 1986. P. 311–326.
191. Hess D.R. How to write an effective discussion // *Respir Care.* 2004. Vol. 49, N 10. P. 1238–1241.
192. Horton R. The hidden research paper // *JAMA.* 2002. Vol. 287, N 21. P. 2775–2778. doi: 10.1001/jama.287.21.2775
193. Horton R. The rhetoric of research // *BMJ.* 1995. Vol. 310, N 6985. P. 985–987. doi: 10.1136/bmj.310.6985.985
194. Docherty M., Smith R. The case for structuring the discussion of scientific papers // *BMJ.* 1999. Vol. 318, N 7193. P. 1224–1225. doi: 10.1136/bmj.318.7193.1224
195. Perneger T.V., Hudelson P.M. Writing a research article: advice to beginners // *Int J Qual Health Care.* 2004. Vol. 16, N 3. P. 191–192. doi: 10.1093/intqhc/mzh053
196. *Annals of Internal Medicine.* Information for authors. Available from: http://www.annals.org/shared/author_info.html. Accessed 10 September 2007.
197. Maldonado G., Poole C. More research is needed // *Ann Epidemiol.* 1999. Vol. 9, N 1. P. 17–18. doi: 10.1016/S1047-2797(98)00050-7
198. Phillips C.V. The economics of 'more research is needed' // *Int J Epidemiol.* 2001. Vol. 30, N 4. P. 771–776. doi: 10.1093/ije/30.4.771
199. Winkleby M.A., Kraemer H.C., Ahn D.K., Varady A.N. Ethnic and socioeconomic differences in cardiovascular disease risk factors: findings for women from the Third National Health and Nutrition Examination Survey, 1988–1994 // *JAMA.* 1998. Vol. 280, N 4. P. 356–362. doi: 10.1001/jama.280.4.356
200. Galuska D.A., Will J.C., Serdula M.K., Ford E.S. Are health care professionals advising obese patients to lose weight? // *JAMA.* 1999. Vol. 282, N 16. P. 1576–1578. doi: 10.1001/jama.282.16.1576
201. Spearman C. The proof and measurement of association between two things // *Am J Psychol.* 1904. Vol. 15, N 1. P. 72–101. doi: 10.2307/1412159
202. Fuller W.A., Hidioglou M.A. Regression estimates after correcting for attenuation // *J Am Stat Assoc.* 1978. Vol. 73, N 361. P. 99–104. doi: 10.1080/01621459.1978.10480011
203. MacMahon S., Peto R., Cutler J., et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias // *Lancet.* 1990. Vol. 335, N 8692. P. 765–774. doi: 10.1016/0140-6736(90)90878-9
204. Phillips A.N., Smith G.D. How independent are "independent" effects? Relative risk estimation when correlated exposures are measured imprecisely // *J Clin Epidemiol.* 1991. Vol. 44, N 11. P. 1223–1231. doi: 10.1016/0895-4356(91)90155-3
205. Phillips A.N., Smith G.D. Bias in relative odds estimation owing to imprecise measurement of correlated exposures // *Stat Med.* 1992. Vol. 11, N 7. P. 953–961. doi: 10.1002/sim.4780110712

- 206.** Greenland S. The effect of misclassification in the presence of covariates // *Am J Epidemiol*. 1980. Vol. 112, N 4. P. 564–569. doi: 10.1093/oxfordjournals.aje.a113025
- 207.** Poole C., Peters U., Il'yasova D., Arab L. Commentary: This study failed? // *Int J Epidemiol*. 2003. Vol. 32, N 4. P. 534–535. doi: 10.1093/ije/dyg197
- 208.** Kaufman J.S., Cooper R.S., McGee D.L. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race // *Epidemiology*. 1997. Vol. 8, N 6. P. 621–628. doi: 10.1097/00001648-199710000-00002
- 209.** Greenland S. Randomization, statistics, and causal inference // *Epidemiology*. 1990. Vol. 1, N 6. P. 421–429. doi: 10.1097/00001648-199011000-00003
- 210.** Taubes G. Epidemiology faces its limits // *Science*. 1995. Vol. 269, N 5221. P. 164–169. doi: 10.1126/science.7618077
- 211.** Temple R. Meta-analysis and epidemiologic studies in drug development and postmarketing surveillance // *JAMA*. 1999. Vol. 281, N 9. P. 841–844. doi: 10.1001/jama.281.9.841
- 212.** Greenberg R.S., Shuster J.L. Epidemiology of cancer in children // *Epidemiol Rev*. 1985. Vol. 7. P. 22–48. doi: 10.1093/oxfordjournals.epirev.a036284
- 213.** Kushi L.H., Mink P.J., Folsom A.R., et al. Prospective study of diet and ovarian cancer // *Am J Epidemiol*. 1999. Vol. 149, N 1. P. 21–31. doi: 10.1093/oxfordjournals.aje.a009723
- 214.** Kemmeren J.M., Algra A., Meijers J.C., et al. Effect of second- and third-generation oral contraceptives on the protein C system in the absence or presence of the factor V Leiden mutation: a randomized trial // *Blood*. 2004. Vol. 103, N 3. P. 927–933. doi: 10.1182/blood-2003-04-1285
- 215.** Egger M., May M., Chene G., et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies // *Lancet*. 2002. Vol. 360, N 9327. P. 119–129. doi: 10.1016/S0140-6736(02)09411-4
- 216.** Campbell D.T. Factors relevant to the validity of experiments in social settings // *Psychol Bull*. 1957. Vol. 54, N 4. P. 297–312. doi: 10.1037/h0040950
- 217.** Justice A.C., Covinsky K.E., Berlin J.A. Assessing the generalizability of prognostic information // *Ann Intern Med*. 1999. Vol. 130, N 6. P. 515–524. doi: 10.7326/0003-4819-130-6-199903160-00016
- 218.** Krimsky S., Rothenberg L.S. Conflict of interest policies in science and medical journals: editorial practices and author disclosures // *Sci Eng Ethics*. 2001. Vol. 7, N 2. P. 205–218. doi: 10.1007/s11948-001-0041-7
- 219.** Bekelman J.E., Li Y., Gross C. Scope and impact of financial conflicts of interest in biomedical research: a systematic review // *JAMA*. 2003. Vol. 289, N 4. P. 454–465. doi: 10.1001/jama.289.4.454
- 220.** Davidson R.A. Source of funding and outcome of clinical trials // *J Gen Intern Med*. 1986. Vol. 1, N 3. P. 155–158. doi: 10.1007/BF02602327
- 221.** Stelfox H.T., Chua G., O'Rourke K., Detsky A.S. Conflict of interest in the debate over calcium-channel antagonists // *N Engl J Med*. 1998. Vol. 338, N 2. P. 101–106. doi: 10.1056/NEJM199801083380206
- 222.** Lexchin J., Bero L.A., Djulbegovic B., Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review // *BMJ*. 2003. Vol. 326, N 7400. P. 1167–1170. doi: 10.1136/bmj.326.7400.1167
- 223.** Als-Nielsen B., Chen W., Gluud C., Kjaergard L.L. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? // *JAMA*. 2003. Vol. 290, N 7. P. 921–928. doi: 10.1001/jama.290.7.921
- 224.** Barnes D.E., Bero L.A. Why re s on the health effects of passive smoking reach different conclusions // *JAMA*. 1998. Vol. 279, N 19. P. 1566–1570. doi: 10.1001/jama.279.19.1566
- 225.** Barnes D.E., Bero L.A. Industry-funded research and conflict of interest: an analysis of research sponsored by the tobacco industry through the Center for Indoor Air Research // *J Health Polit Policy Law*. 1996. Vol. 21, N 3. P. 515–542. doi: 10.1215/03616878-21-3-515
- 226.** Glantz S.A., Barnes D.E., Bero L., et al. Looking through a keyhole at the tobacco industry. The Brown and Williamson documents // *JAMA*. 1995. Vol. 274, N 3. P. 219–224. doi: 10.1001/jama.1995.03530030039032
- 227.** Huss A., Egger M., Hug K., et al. Source of funding and results of studies of health effects of mobile phone use: systematic review of experimental studies // *Environ Health Perspect*. 2007. Vol. 115, N 1. P. 1–4. doi: 10.1289/ehp.9149
- 228.** Safer D.J. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials // *J Nerv Ment Dis*. 2002. Vol. 190, N 9. P. 583–592. doi: 10.1097/00005053-200209000-00002
- 229.** Aspinall R.L., Goodman N.W. Denial of effective treatment and poor quality of clinical information in placebo controlled trials of ondansetron for postoperative nausea and vomiting: a review of published trials // *BMJ*. 1995. Vol. 311, N 7009. P. 844–846. doi: 10.1136/bmj.311.7009.844
- 230.** Chan A.W., Hrobjartsson A., Haahr M.T., et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles // *JAMA*. 2004. Vol. 291, N 20. P. 2457–2465. doi: 10.1001/jama.291.20.2457
- 231.** Melander H., Ahlqvist-Rastad J., Meijer G., Beermann B. Evidence b(i)ased medicine-selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications // *BMJ*. 2003. Vol. 326, N 7400. P. 1171–1173. doi: 10.1136/bmj.326.7400.1171
- 232.** Scherer R.W., Langenberg P., von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews*. (Issue 2). Art. No.: MR000005. Available from: <http://www.cochrane.org/reviews/en/mr000005.html>. Accessed 10 September 2005. doi: 10.1002/14651858.MR000005.pub2
- 233.** Moher D., Schulz K.F., Altman D.G. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials // *Lancet*. 2001. Vol. 357, N 9263. P. 1191–1194. doi: 10.1016/S0140-6736(00)04337-3
- 234.** Stroup D.F., Berlin J.A., Morton S.C., et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group // *JAMA*. 2000. Vol. 283, N 15. P. 2008–2012. doi: 10.1001/jama.283.15.2008
- 235.** Altman D.G., Schulz K.F., Moher D., et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration // *Ann Intern Med*. 2001. Vol. 134, N 8. P. 663–694. doi: 10.7326/0003-4819-134-8-200104170-00012
- 236.** Moher D. CONSORT: an evolving tool to help improve the quality of reports of randomized controlled trials. *Consolidated Standards of Reporting Trials* // *JAMA*. 1998. Vol. 279, N 18. P. 1489–1491. doi: 10.1001/jama.279.18.1489
- 237.** Begg C., Cho M., Eastwood S., et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement // *JAMA*. 1996. Vol. 276, N 8. P. 637–639. doi: 10.1001/jama.276.8.637

REFERENCES

1. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ*. 2004;328(7430):39–41. doi: 10.1136/bmj.328.7430.39
2. Funai EF, Rosenbush EJ, Lee MJ, Del Priore G. Distribution of study designs in four major US journals of obstetrics and gynecology. *Gynecol Obstet Invest*. 200;151(1):8–11. doi: 10.1159/000052882
3. Scales CD, Norris RD, Peterson BL, et al. Clinical research and statistical methods in the urology literature. *J Urol*. 2005;174(4 Pt 1):1374–1379. doi: 10.1097/01.ju.0000173640.91654.b5
4. Pocock SJ, Collier TJ, Dandreo KJ, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ*. 2004;329(7471):883. doi: 10.1136/bmj.38250.571088.55
5. Tooth L, Ware R, Bain C, et al. Quality of reporting of observational longitudinal research. *Am J Epidemiol*. 2005;161(3):280–288. doi: 10.1093/aje/kwi042
6. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Epidemiology*. 2007;18(6):800–804. doi: 10.1097/EDE.0b013e3181577654
7. Mihailovic A, Bell CM, Urbach DR. Users' guide to the surgical literature. Case-control studies in surgical journals. *Can J Surg*. 2005;48(2):148–151.
8. Rushton L. Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods. *Occup Environ Med*. 2000;57(1):1–9. doi: 10.1136/oem.57.1.1
9. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990;1(1):43–46. doi: 10.1097/00001648-199001000-00010
10. Moonesinghe R, Khoury MJ, Janssens AC. Most published research findings are false—but a little replication goes a long way. *PLoS Med*. 2007;4(2):e28. doi: 10.1371/journal.pmed.0040028
11. Jenicek M. Clinical Case Reporting. Evidence-Based Medicine. Oxford: Butterworth-Heinemann; 1999. 117 p.
12. Vandenbroucke JP. In defense of case reports and case series. *Ann Intern Med*. 2001;134(4):330–334. doi: 10.7326/0003-4819-134-4-200102200-00017
13. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med*. 2003;138(1):40–44. doi: 10.7326/0003-4819-138-1-200301070-00010
14. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer*. 2005;93(4):387–391. doi: 10.1038/sj.bjc.6602678
15. Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet*. 2006;38(1):3–5. doi: 10.1038/ng0106-3
16. Rodrigues L, Kirkwood BR. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. *Int J Epidemiol*. 1990;19(1):205–213. doi: 10.1093/ije/19.1.205
17. Rothman KJ, Greenland S. Case-Control Studies. In: Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Raven; 1998. P. 93–114.
18. Forand SP. Leukaemia incidence among workers in the shoe and boot manufacturing industry: a case-control study. *Environ Health*. 2004;3(1):7. doi: 10.1186/1476-069X-3-7
19. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342(25):1878–1886. doi: 10.1056/NEJM200006223422506
20. Gøtzsche PC, Harden A. Searching for non-randomised studies. Draft chapter 3. Cochrane Non-Randomised Studies Methods Group; 2002. Available from: <http://www.cochrane.dk/nrsmg>. Accessed 10 September 2007.
21. Lohse N, Hansen AB, Pedersen G, et al. Survival of persons with and without HIV infection in Denmark, 1995–2005. *Ann Intern Med*. 2007;146(2):87–95. doi: 10.7326/0003-4819-146-2-200701160-00003
22. American Journal of Epidemiology. 2007. Information for authors. Available from: http://www.oxfordjournals.org/aje/for_authors/index.html. Accessed 10 September 2007.
23. Haynes RB, Mulrow CD, Huth EJ, et al. More informative abstracts revisited. *Ann Intern Med*. 1990;113(1):69–76. doi: 10.7326/0003-4819-113-1-69
24. Taddio A, Pain T, Fassos FF, et al. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. *CMAJ*. 1994;150(10):1611–1615.
25. Hartley J, Sydes M. Which layout do you prefer? An analysis of readers' preferences for different typographic layouts of structured abstracts. *J Inform Sci*. 1996;22(1):27–37. doi: 10.1177/016555159602200103
26. Viner RM, Cole TJ. Adult socioeconomic, educational, social, and psychological outcomes of childhood obesity: a national birth cohort study. *BMJ*. 2005;330(7504):1354. doi: 10.1136/bmj.38453.422049.E0
27. McCauley J, Kern DE, Kolodner K, et al. The "battering syndrome": prevalence and clinical characteristics of domestic violence in primary care internal medicine practices. *Ann Intern Med*. 1995;123(10):737–746. doi: 10.7326/0003-4819-123-10-199511150-00001
28. McEvoy SP, Stevenson MR, McCartt AT, et al. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *BMJ*. 2005;331(7514):428. doi: 10.1136/bmj.38537.397512.55
29. Vandenbroucke JP. Prospective or retrospective: what's in a name? *BMJ*. 1991;302(6771):249–250. doi: 10.1136/bmj.302.6771.249
30. Last JM. A Dictionary of Epidemiology. New York: Oxford University Press; 2000.
31. Miettinen OS. Theoretical Epidemiology: principles of occurrence research in medicine. New York: Wiley; 1985. P. 64–66.
32. Rothman KJ, Greenland S. Types of Epidemiologic Studies. In: Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd ed. Lippincott Raven; 1998. P. 74–75.
33. MacMahon B, Trichopoulos D. Epidemiology, principles and methods. 2nd ed. Boston: Little, Brown; 1996. 81 p. doi: 10.1016/S0033-3506(97)00047-4
34. Lilienfeld AM. Foundations of Epidemiology. New York: Oxford University Press; 1976.
35. Ridker PM, Hennekens CH, Lindpaintner K, et al. Mutation in the gene coding for coagulation factor V and the risk of myocardial infarction, stroke, and venous thrombosis in apparently healthy men. *N Engl J Med*. 1995;332(14):912–917. doi: 10.1056/NEJM199504063321403
36. Goodman KJ, O'Rourke K, Day RS, et al. Dynamics of *Helicobacter pylori* infection in a US-Mexico cohort during the first two years of life. *Int J Epidemiol*. 2005;34(6):1348–1355. doi: 10.1093/ije/dyi152

37. Altman DG, De Stavola BL, Love SB, Stepniowska KA. Review of survival analyses published in cancer journals. *Br J Cancer*. 1995;72(2): 511–518. doi: 10.1038/bjc.1995.364
38. Cerhan JR, Wallace RB, Folsom AR, et al. Transfusion history and cancer risk in older women. *Ann Intern Med*. 1993;119(1):8–15. doi: 10.7326/0003-4819-119-1-199307010-00002
39. Freeman LE, Dennis LK, Lynch CF, et al. Toenail arsenic content and cutaneous melanoma in Iowa. *Am J Epidemiol*. 2004;160(7):679–687. doi: 10.1093/aje/kwh267
40. Canto JG, Allison JJ, Kiefe C, et al. Relation of race and sex to the use of reperfusion therapy in Medicare beneficiaries with acute myocardial infarction. *N Engl J Med*. 2000;342(15):1094–1100. doi: 10.1056/NEJM200004133421505
41. Metzker-Cotter E, Kletter Y, Avidor B, et al. Long-term serological analysis and clinical follow-up of patients with cat scratch disease. *Clin Infect Dis*. 2003;37(9):1149–1154. doi: 10.1086/378738
42. Johnson ES. Bias on withdrawing lost subjects from the analysis at the time of loss, in cohort mortality studies, and in follow-up methods. *J Occup Med*. 1990;32(3):250–254. doi: 10.1097/00043764-199003000-00013
43. Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biom Bull*. 1946;2(3):47. doi: 10.2307/3002000
44. Feinstein AR, Walter SD, Horwitz RJ. An analysis of Berkson's bias in case-control studies. *J Chronic Dis*. 1986;39(7):495–504. doi: 10.1016/0021-9681(86)90194-3
45. Jick H, Vessey MP. Case-control studies in the evaluation of drug-induced illness. *Am J Epidemiol*. 1978;107(1):1–7. doi: 10.1093/oxfordjournals.aje.a112502
46. Hackam DG, Mamdani M, Li P, Redelmeier DA. Statins and sepsis in patients with cardiovascular disease: a population-based cohort analysis. *Lancet*. 2006;367(9508):413–418. doi: 10.1016/S0140-6736(06)68041-0
47. Smeeth L, Cook C, Fombonne E, et al. MMR vaccination and pervasive developmental disorders: a case-control study. *Lancet*. 2004;364(9438):963–969. doi: 10.1016/S0140-6736(04)17020-7
48. Costanza MC. Matching. *Prev Med*. 1995;24(5):425–433. doi: 10.1006/pmed.1995.1069
49. Sturmer T, Brenner H. Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case-control studies. *Am J Epidemiol*. 2002;155(7):593–602. doi: 10.1093/aje/k155.7.593
50. Rothman KJ, Greenland S. Matching. In: Rothman KJ, Greenland S, editors. 2nd ed. *Modern epidemiology*. Lippincott Raven; 1998. P. 147–161.
51. Szklo MF, Nieto J. *Epidemiology, Beyond the Basics*. Sudbury (MA): Jones and Bartlett; 2000. P. 40–51.
52. Cole P, MacMahon B. Attributable risk percent in case-control studies. *Br J Prev Soc Med*. 1971;25(4):242–244. doi: 10.1136/jech.25.4.242
53. Gissler M, Hemminki E. The danger of overmatching in studies of the perinatal mortality and birthweight of infants born after assisted conception. *Eur J Obstet Gynecol Reprod Biol*. 1996;69(2):73–75. doi: 10.1016/0301-2115(95)02517-0
54. Gefeller O, Pfahlerberg A, Brenner H, Windeler J. An empirical investigation on matching in published case-control studies. *Eur J Epidemiol*. 1998;14(4):321–325. doi: 10.1023/A:1007497104800
55. Artama M, Ritvanen A, Gissler M, et al. Congenital structural anomalies in offspring of women with epilepsy—a population-based cohort study in Finland. *Int J Epidemiol*. 2006;35(2):280–287. doi: 10.1093/ije/dyi234
56. Ebrahim S. Cohorts, infants and children. *Int J Epidemiol*. 2004;33(6):1165–1166. doi: 10.1093/ije/dyh368
57. Walker M, Whincup PH, Shaper AG. The British Regional Heart Study 1975–2004. *Int J Epidemiol*. 2004;33(6):1185–1192. doi: 10.1093/ije/dyh295
58. Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J Clin Epidemiol*. 2005;58(6):560–567. doi: 10.1016/j.jclinepi.2004.11.018
59. Anderson HR, Atkinson RW, Peacock JL, et al. Ambient particulate matter and health effects: publication bias in studies of short-term associations. *Epidemiology*. 2005;16(2):155–163. doi: 10.1097/01.ede.0000152528.22746.0f
60. Winkelmayr WC, Stampfer MJ, Willett WC, Curhan GC. Habitual caffeine intake and the risk of hypertension in women. *JAMA*. 2005;294(18):2330–2335. doi: 10.1001/jama.294.18.2330
61. Lukanova A, Soderberg S, Kaaks R, et al. Serum adiponectin is not associated with risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev*. 2006;15(2):401–402. doi: 10.1158/1055-9965.EPI-05-0836
62. Becher H. The concept of residual confounding in regression models and some applications. *Stat Med*. 1992;11(13):1747–1758. doi: 10.1002/sim.4780111308
63. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. *Epidemiology*. 1997;8(4):429–434. doi: 10.1097/00001648-199707000-00014
64. Phillips MR, Yang G, Zhang Y, et al. Risk factors for suicide in China: a national case-control psychological autopsy study. *Lancet*. 2002;360(9347):1728–1736. doi: 10.1016/S0140-6736(02)11681-3
65. Pasquale LR, Kang JH, Manson JE, et al. Prospective study of type 2 diabetes mellitus and risk of primary open-angle glaucoma in women. *Ophthalmology*. 2006;113(7):1081–1086. doi: 10.1016/j.opht.2006.01.066
66. Craig SL, Feinstein AR. Antecedent therapy versus detection bias as causes of neoplastic multimorbidity. *Am J Clin Oncol*. 1999;22(1):51–56. doi: 10.1097/00000421-199902000-00013
67. Rogler LH, Mroczek DK, Fellows M, Loftus ST. The neglect of response bias in mental health research. *J Nerv Ment Dis*. 2001;189(3):182–187. doi: 10.1097/00005053-200103000-00007
68. Murphy EA. *The logic of medicine*. Baltimore: Johns Hopkins University Press; 1976.
69. Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32(1-2):51–63. doi: 10.1016/0021-9681(79)90012-2
70. Johannes CB, Crawford SL, McKinlay JB. Interviewer effects in a cohort study. Results from the Massachusetts Women's Health Study. *Am J Epidemiol*. 1997;146(5):429–438. doi: 10.1093/oxfordjournals.aje.a009296
71. Bloemenkamp KW, Rosendaal FR, Buller HR, et al. Risk of venous thrombosis with use of current low-dose oral contraceptives is not explained by diagnostic suspicion and referral bias. *Arch Intern Med*. 1999;159(1):65–70. doi: 10.1001/archinte.159.1.65
72. Feinstein AR. *Clinical epidemiology: the architecture of clinical research*. Philadelphia: W.B. Saunders; 1985.
73. Yadon ZE, Rodrigues LC, Davies CR, Quigley MA. Indoor and peridomestic transmission of American cutaneous leishmaniasis in northwestern Argentina: a retrospective case-control study. *Am J Trop Med Hyg*. 2003;68(5): 519–526. doi: 10.4269/ajtmh.2003.68.519

74. Anoop S, Saravanan B, Joseph A, et al. Maternal depression and low maternal intelligence as risk factors for malnutrition in children: a community based case-control study from South India. *Arch Dis Child*. 2004;89(4):325–329. doi: 10.1136/adc.2002.009738
75. Carlin JB, Doyle LW. Sample size. *J Paediatr Child Health*. 2002;38(3):300–304. doi: 10.1046/j.1440-1754.2002.00855.x
76. Rigby AS, Vail A. Statistical methods in epidemiology. II: A commonsense approach to sample size estimation. *Disabil Rehabil*. 1998;20(11):405–410. doi: 10.3109/09638289809166102
77. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet*. 2005;365(9467):1348–1353. doi: 10.1016/S0140-6736(05)61034-3
78. Drescher K, Timm J, Jockel KH. The design of case-control studies: the effect of confounding on sample size requirements. *Stat Med*. 1990;9(7):765–776. doi: 10.1002/sim.4780090706
79. Devine OJ, Smith JM. Estimating sample size for epidemiologic studies: the impact of ignoring exposure measurement uncertainty. *Stat Med*. 1998;17(12):1375–1389. doi: 10.1002/(SICI)1097-0258(19980630)17:12<1375::AID-SIM857>3.0.CO;2-D
80. Linn S, Levi L, Grunau PD, et al. Effect measure modification and confounding of severe head injury mortality by age and multiple organ injury severity. *Ann Epidemiol*. 2007;17(2):142–147. doi: 10.1016/j.annepidem.2006.08.004
81. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using “optimal” cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst*. 1994;86(11):829–835. doi: 10.1093/jnci/86.11.829
82. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006;25(1):127–141. doi: 10.1002/sim.2331
83. Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. *Epidemiology*. 1995;6(4):450–454. doi: 10.1097/00001648-199507000-00025
84. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol*. 1999;28(5):964–974. doi: 10.1093/ije/28.5.964
85. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Methods*. 2002;7(1):19–40. doi: 10.1037/1082-989X.7.1.19
86. Altman DG. Categorizing continuous variables. In: Armitage P, Colton T, editors. *Encyclopedia of biostatistics*. 2nd ed. Chichester: John Wiley; 2005. P. 708–711. doi: 10.1002/0470011815.b2a10012
87. Cohen J. The cost of dichotomization. *Applied Psychological Measurement*. 1983;7(3):249–253. doi: 10.1177/014662168300700301
88. Zhao LP, Kolonel LN. Efficiency loss from categorizing quantitative exposures into qualitative exposures in case-control studies. *Am J Epidemiol*. 1992;136(4):464–474. doi: 10.1093/oxfordjournals.aje.a116520
89. Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 1968;24(2):295–313. doi: 10.2307/2528036
90. Clayton D, Hills M. Models for dose-response (Chapter 25). *Statistical Models in Epidemiology*. Oxford: Oxford University Press; 1993. P. 249–260.
91. Cox DR. Note on grouping. *J Am Stat Assoc*. 1957;52(280):543–547. doi: 10.1080/01621459.1957.10501411
92. Il'yasova D, Hertz-Picciotto I, Peters U, et al. Choice of exposure scores for categorical regression in meta-analysis: a case study of a common problem. *Cancer Causes Control*. 2005;16(4):383–388. doi: 10.1007/s10552-004-5025-x
93. Berglund A, Alfredsson L, Cassidy JD, et al. The association between exposure to a rear-end collision and future neck or shoulder pain: a cohort study. *J Clin Epidemiol*. 2000;53(11):1089–1094. doi: 10.1016/S0895-4356(00)00225-0
94. Srama R, Werwatz A. Controlling for continuous confounding factors: non- and semiparametric approaches. *Rev Epidemiol Sante Publique*. 2005;53(2):2S65–80. doi: 10.1016/S0398-7620(05)84769-8
95. Greenland S. Introduction to regression modelling (Chapter 21). In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed: Lippincott Raven; 1998. P. 401–432.
96. Thompson WD. Statistical analysis of case-control studies. *Epidemiol Rev*. 1994;16(1):33–50. doi: 10.1093/oxfordjournals.epirev.a036143
97. Schlesselman JJ. Logistic regression for case-control studies (Chapter 8.2). *Case-control studies Design, conduct, analysis*. New York, Oxford: Oxford University Press; 1982. P. 235–241.
98. Clayton D, Hills M. Choice and interpretation of models (Chapter 27). *Statistical Models in Epidemiology*. Oxford: Oxford University Press; 1993. P. 271–281.
99. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *Br Med J*. 1983;286(6376):1489–1493. doi: 10.1136/bmj.286.6376.1489
100. International Committee of Medical Journal Editors Uniform requirements for manuscripts submitted to biomedical journals. *N Engl J Med*. 1997;336:309–315. [Electronic version updated February 2006, Available from at <http://www.icmje.org/>]. doi: 10.1056/NEJM199701233360422
101. Mullner M, Matthews H, Altman DG. Reporting on statistical methods to adjust for confounding: a cross-sectional survey. *Ann Intern Med*. 2002;136(2):122–126. doi: 10.7326/0003-4819-136-2-200201150-00009
102. Olsen J, Basso O. Re: Residual confounding. *Am J Epidemiol*. 1999;149(3):290. doi: 10.1093/oxfordjournals.aje.a009805
103. Hallan S, de Mutsert R, Carlsen S, et al. Obesity, smoking, and physical inactivity as risk factors for CKD: are men more vulnerable? *Am J Kidney Dis*. 2006;47(3):396–405. doi: 10.1053/j.ajkd.2005.11.027
104. Gotzsche PC. Believability of relative risks and odds ratios in abstracts: cross sectional study. *BMJ*. 2006;333(7561):231–234. doi: 10.1136/bmj.38895.410451.79
105. Szklo MF, Nieto J. Communicating Results of Epidemiologic Studies (Chapter 9). *Epidemiology, Beyond the Basics*. Sudbury (MA): Jones and Bartlett; 2000. P. 408–430.
106. Chandola T, Brunner E, Marmot M. Chronic stress at work and the metabolic syndrome: prospective study. *BMJ*. 2006;332(7540):521–525. doi: 10.1136/bmj.38693.435301.80
107. Vach W, Blettner M. Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables. *Am J Epidemiol*. 1991;134(8):895–907. doi: 10.1093/oxfordjournals.aje.a116164
108. Little RJ, Rubin DB. A taxonomy of missing-data methods (Chapter 1.4.). *Statistical Analysis with Missing Data*. New York: Wiley; 2002. P. 19–23. doi: 10.1002/9781119013563
109. Ware JH. Interpreting incomplete data in studies of diet and weight loss. *N Engl J Med*. 2003;348(21):2136–2137. doi: 10.1056/NEJMe030054
110. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581–592. doi: 10.1093/biomet/63.3.581
111. Schafer JL. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall; 1997. doi: 10.1201/9781439821862

112. Lipsitz SR, Ibrahim JG, Chen MH, Peterson H. Non-ignorable missing covariates in generalized linear models. *Stat Med*. 1999;18(17-18):2435–2448. doi: 10.1002/(SICI)1097-0258(19990915/30)18:17/18<2435::AID-SIM267>3.0.CO;2-B
113. Rotnitzky A, Robins J. Analysis of semi-parametric regression models with non-ignorable non-response. *Stat Med*. 1997;16(1-3):81–102. doi: 10.1002/(SICI)1097-0258(19970115)16:1<81::AID-SIM473>3.0.CO;2-0
114. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley; 1987. doi: 10.1002/9780470316696
115. Barnard J, Meng XL. Applications of multiple imputation in medical studies: from AIDS to NHANES. *Stat Methods Med Res*. 1999;8(1):17–36. doi: 10.1177/096228029900800103
116. Braitstein P, Brinkhof MW, Dabis F, et al. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet*. 2006;367(9513):817–824. doi: 10.1016/S0140-6736(06)68337-2
117. Purandare N, Burns A, Daly KJ, et al. Cerebral emboli as a potential cause of Alzheimer's disease and vascular dementia: case-control study. *BMJ*. 2006;332(7550):1119–1124. doi: 10.1136/bmj.38814.696493.AE
118. Steyn K, Gaziano TA, Bradshaw D, et al. Hypertension in South African adults: results from the Demographic and Health Survey, 1998. *J Hypertens*. 2001;19(10):1717–1725. doi: 10.1097/00004872-200110000-00004
119. Lohr SL. Design Effects (Chapter 7.5). Sampling: Design and Analysis. Pacific Grove (CA): Duxbury Press; 1999.
120. Dunn NR, Arscott A, Thorogood M. The relationship between use of oral contraceptives and myocardial infarction in young women with fatal outcome, compared to those who survive: results from the MICA case-control study. *Contraception*. 2001;63(2):65–69. doi: 10.1016/S0010-7824(01)00172-X
121. Rothman KJ, Greenland S. Basic Methods for Sensitivity Analysis and External Adjustment. In: Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd ed. Lippincott Raven; 1998. P. 343–357.
122. Custer B, Longstreth WT, Phillips LE, et al. Hormonal exposures and the risk of intracranial meningioma in women: a population-based case-control study. *BMC Cancer*. 2006;6:152. doi: 10.1186/1471-2407-6-152
123. Wakefield MA, Chaloupka FJ, Kaufman NJ, et al. Effect of restrictions on smoking at home, at school, and in public places on teenage smoking: cross sectional study. *BMJ*. 2000;321(7257):333–337. doi: 10.1136/bmj.321.7257.333
124. Greenland S. The impact of prior distributions for uncontrolled confounding and response bias: a case study of the relation of wire codes and magnetic fields to childhood leukemia. *J Am Stat Assoc*. 2003;98(461):47–54. doi: 10.1198/01621450338861905
125. Lash TL, Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology*. 2003;14(4):451–458. doi: 10.1097/01.EDE.0000071419.41011.cf
126. Phillips CV. Quantifying and reporting uncertainty from systematic errors. *Epidemiology*. 2003;14(4):459–466. doi: 10.1097/01.ede.0000072106.65262.ae
127. Cornfield J, Haenszel W, Hammond EC, et al. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst*. 1959;22(1):173–203.
128. Langholz B. Factors that explain the power line configuration wiring code-childhood leukemia association: what would they look like? *Bioelectromagnetics*. 2001;Suppl 5:S19–31. doi: 10.1002/1521-186x(2001)22:5+<::aid-bem1021>3.3.co;2-9.
129. Eisner MD, Smith AK, Blanc PD. Bartenders' respiratory health after establishment of smoke-free bars and taverns. *JAMA*. 1998;280(22):1909–1914. doi: 10.1001/jama.280.22.1909
130. Dunne MP, Martin NG, Bailey JM, et al. Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *Int J Epidemiol*. 1997;26(4):844–854. doi: 10.1093/ije/26.4.844
131. Schuz J, Kaatsch P, Kaletsch U, et al. Association of childhood cancer with factors related to pregnancy and birth. *Int J Epidemiol*. 1999;28(4):631–639. doi: 10.1093/ije/28.4.631
132. Cnattingius S, Zack M, Ekblom A, et al. Prenatal and neonatal risk factors for childhood myeloid leukemia. *Cancer Epidemiol Biomarkers Prev*. 1995;4(5):441–445.
133. Schuz J. Non-response bias as a likely cause of the association between young maternal age at the time of delivery and the risk of cancer in the offspring. *Paediatr Perinat Epidemiol*. 2003;17(1):106–112. doi: 10.1046/j.1365-3016.2003.00460.x
134. Slaterry ML, Edwards SL, Caan BJ, et al. Response rates among control subjects in case-control studies. *Ann Epidemiol*. 1995;5(3):245–249. doi: 10.1016/1047-2797(94)00113-8
135. Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet*. 2002;359(9304):431–434. doi: 10.1016/S0140-6736(02)07605-5
136. Olson SH, Voigt LF, Begg CB, Weiss NS. Reporting participation in case-control studies. *Epidemiology*. 2002;13(2):123–126. doi: 10.1097/00001648-200203000-00004
137. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol*. 2006;163(3):197–203. doi: 10.1093/aje/kwj036
138. Olson SH. Reported participation in case-control studies: changes over time. *Am J Epidemiol*. 2001;154(6):574–581. doi: 10.1093/aje/154.6.574
139. Sandler DP. On revealing what we'd rather hide: the problem of describing study participation. *Epidemiology*. 2002;13(2):117. doi: 10.1097/00001648-200203000-00001
140. Hepworth SJ, Schoemaker MJ, Muir KR, et al. Mobile phone use and risk of glioma in adults: case-control study. *BMJ*. 2006;332(7546):883–887. doi: 10.1136/bmj.38720.687975.55
141. Hay AD, Wilson A, Fahey T, Peters TJ. The duration of acute cough in pre-school children presenting to primary care: a prospective cohort study. *Fam Pract*. 2003;20(6):696–705. doi: 10.1093/fampra/cmg613
142. Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA*. 2001;285(15):1996–1999. doi: 10.1001/jama.285.15.1996
143. Osella AR, Misciagna G, Guerra VM, et al. Hepatitis C virus (HCV) infection and liver-related mortality: a population-based cohort study in southern Italy. The Association for the Study of Liver Disease in Puglia. *Int J Epidemiol*. 2000;29(5):922–927. doi: 10.1093/ije/29.5.922
144. Dales LG, Ury HK. An improper use of statistical significance testing in studying covariables. *Int J Epidemiol*. 1978;7(4):373–375. doi: 10.1093/ije/7.4.373
145. Maldonado G, Greenland S. Simulation study of confounder-selection strategies. *Am J Epidemiol*. 1993;138(11):923–936. doi: 10.1093/oxfordjournals.aje.a116813
146. Tanis BC, van den Bosch MA, Kemmeren JM, et al. Oral contraceptives and the risk of myocardial infarction. *N Engl J Med*. 2001;345(25):1787–1793. doi: 10.1056/NEJMoa003216

147. Rothman KJ, Greenland S. Precision and Validity in Epidemiologic Studies. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed. Lippincott Raven; 1998. P. 120–125.
148. Clark TG, Altman DG, De Stavola BL. Quantification of the completeness of follow-up. *Lancet*. 2002;359(9314):1309–1310. doi: 10.1016/S0140-6736(02)08272-7
149. Qiu C, Fratiglioni L, Karp A, et al. Occupational exposure to electromagnetic fields and risk of Alzheimer's disease. *Epidemiology*. 2004;15(6):687–694. doi: 10.1097/01.ede.0000142147.49297.9d
150. Kengeya-Kayondo JF, Kamali A, Nunn AJ, et al. Incidence of HIV-1 infection in adults and socio-demographic characteristics of seroconverters in a rural population in Uganda: 1990–1994. *Int J Epidemiol*. 1996;25(5):1077–1082. doi: 10.1093/ije/25.5.1077
151. Mastrangelo G, Fedeli U, Fadda E, et al. Increased risk of hepatocellular carcinoma and liver cirrhosis in vinyl chloride workers: synergistic effect of occupational exposure with alcohol intake. *Environ Health Perspect*. 2004;112(11):1188–1192. doi: 10.1289/ehp.6972
152. Salo PM, Arbes SJ, Sever M, et al. Exposure to *Alternaria alternata* in US homes is associated with asthma symptoms. *J Allergy Clin Immunol*. 2006;118(4):892–898. doi: 10.1016/j.jaci.2006.07.037
153. Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet*. 2002;359(9318):1686–1689. doi: 10.1016/S0140-6736(02)08594-X
154. Sasieni P. A note on the presentation of matched case-control data. *Stat Med*. 1992;11(5):617–620. doi: 10.1002/sim.4780110506
155. Lee GM, Neutra RR, Hristova L, et al. A nested case-control study of residential and personal magnetic field measures and miscarriages. *Epidemiology*. 2002;13(1):21–31. doi: 10.1097/00001648-200201000-00005
156. Tiihonen J, Walhbeck K, Lonnqvist J, et al. Effectiveness of antipsychotic treatments in a nationwide cohort of patients in community care after first hospitalisation due to schizophrenia and schizoaffective disorder: observational follow-up study. *BMJ*. 2006;333(7561):224. doi: 10.1136/bmj.38881.382755.2F
157. Christenfeld NJ, Sloan RP, Carroll D, Greenland S. Risk factors, confounding, and the illusion of statistical control. *Psychosom Med*. 2004;66(6):868–875. doi: 10.1097/01.psy.0000140008.70959.41
158. Smith GD, Phillips A. Declaring independence: why we should be cautious. *J Epidemiol Community Health*. 1990;44(4):257–258. doi: 10.1136/jech.44.4.257
159. Greenland S, Neutra R. Control of confounding in the assessment of medical technology. *Int J Epidemiol*. 1980;9(4):361–367. doi: 10.1093/ije/9.4.361
160. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiology*. 2001;12(3):313–320. doi: 10.1097/00001648-200105000-00011
161. Sagiv SK, Tolbert PE, Altshul LM, Korrick SA. Organochlorine exposures during pregnancy and infant size at birth. *Epidemiology*. 2007;18(1):120–129. doi: 10.1097/01.ede.0000249769.15001.7c
162. World Health Organization. Body Mass Index (BMI). 2007. Available from: http://www.euro.who.int/nutrition/20030507_1
163. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet*. 2003;362(9382):419–427. doi: 10.1016/S0140-6736(03)14065-2.
164. Hill AB. The environment and disease: Association or causation? *Proc R Soc Med*. 1965;58(5):295–300. doi: 10.1177/003591576505800503
165. Vineis P. Causality in epidemiology. *Soz Präventivmed*. 2003;48(2):80–87. doi: 10.1007/s00038-003-1029-7
166. Empana JP, Ducimetiere P, Arveiler D, et al. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. *Eur Heart J*. 2003;24(21):1903–1911. doi: 10.1016/j.ehj.2003.09.002
167. Tunstall-Pedoe H, Kuulasmaa K, Mahonen M, et al. Contribution of trends in survival and coronary-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA project populations. Monitoring trends and determinants in cardiovascular disease. *Lancet*. 1999;353(9164):1547–1557. doi: 10.1016/S0140-6736(99)04021-0
168. Cambien F, Chretien JM, Ducimetiere P, et al. Is the relationship between blood pressure and cardiovascular risk dependent on body mass index? *Am J Epidemiol*. 1985;122(3):434–442. doi: 10.1093/oxfordjournals.aje.a114124
169. Hosmer DW, Taber S, Lemeshow S. The importance of assessing the fit of logistic regression models: a case study. *Am J Public Health*. 1991;81(12):1630–1635. doi: 10.2105/AJPH.81.12.1630
170. Tibshirani R. A plain man's guide to the proportional hazards model. *Clin Invest Med*. 1982;5(1):63–68.
171. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health*. 1998;88(1):15–19. doi: 10.2105/AJPH.88.1.15
172. Uter W, Pfahlberg A. The application of methods to quantify attributable risk in medical practice. *Stat Methods Med Res*. 2001;10(3):231–237. doi: 10.1177/096228020101000305
173. Schwartz LM, Woloshin S, Dvorin EL, Welch HG. Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks. *BMJ*. 2006;333(7581):1248. doi: 10.1136/bmj.38985.564317.7C
174. Nakayama T, Zaman MM, Tanaka H. Reporting of attributable and relative risks, 1966–97. *Lancet*. 1998;351(9110):1179. doi: 10.1016/S0140-6736(05)79123-6
175. Cornfield J. A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. *J Natl Cancer Ins*. 1951;11(6):1269–1275.
176. Pearce N. What does the odds ratio estimate in a case-control study? *Int J Epidemiol*. 1993;22(6):1189–1192. doi: 10.1093/ije/22.6.1189
177. Rothman KJ, Greenland S, Rothman KJ, Greenland S. Measures of Disease Frequency. *Modern epidemiology*. 2nd ed. Lippincott Raven; 1998. P. 44–45.
178. Doll R, Hill AB. The mortality of doctors in relation to their smoking habits: a preliminary report. *BMJ*. 1954;1(4877):1451–1455. doi: 10.1136/bmj.1.4877.1451
179. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet*. 2003;362(9387):847–852. doi: 10.1016/S0140-6736(03)14338-3
180. Greenland S. Applications of Stratified Analysis Methods. In: Rothman KJ, Greenland S, editors. *Modern epidemiology*. 2nd ed. Lippincott Raven; 1998. P. 295–297.
181. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 2001;30(3):427–432. discussion. doi: 10.1093/ije/30.3.427
182. Vandenbroucke JP, Koster T, Briet E, et al. Increased risk of venous thrombosis in oral-contraceptive users who are carriers of factor V Leiden mutation. *Lancet*. 1994;344(8935):1453–1457. doi: 10.1016/S0140-6736(94)90286-0
183. Botto LD, Khoury MJ. Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond. *Am J Epidemiol*. 2001;153(10):1016–1020. doi: 10.1093/aje/153.10.1016

184. Wei L, MacDonald TM, Walker BR. Taking glucocorticoids by prescription is associated with subsequent cardiovascular disease. *Ann Intern Med.* 2004;141(10):764–770. doi: 10.7326/0003-4819-141-10-200411160-00007
185. Martinelli I, Taioli E, Battaglioli T, et al. Risk of venous thromboembolism after air travel: interaction with thrombophilia and oral contraceptives. *Arch Intern Med.* 2003;163(22):2771–2774. doi: 10.1001/archinte.163.22.2771
186. Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst.* 2005;97(14):1043–1055. doi: 10.1093/jnci/dji184
187. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol.* 1980;112(4):467–470. doi: 10.1093/oxfordjournals.aje.a113015
188. Saracci R. Interaction and synergism. *Am J Epidemiol.* 1980;112(4):465–466. doi: 10.1093/oxfordjournals.aje.a113014
189. Rothman KJ. *Epidemiology. An introduction.* Oxford: Oxford University Press; 2002. P. 168–180.
190. Rothman KJ. *Interactions Between Causes. Modern epidemiology.* Boston: Little Brown; 1986. P. 311–326.
191. Hess DR. How to write an effective discussion. *Respir Care.* 2004;49(10):1238–1241.
192. Horton R. The hidden research paper. *JAMA.* 2002;287(21):2775–2778. doi: 10.1001/jama.287.21.2775
193. Horton R. The rhetoric of research. *BMJ.* 1995;310(6985):985–987. doi: 10.1136/bmj.310.6985.985
194. Docherty M, Smith R. The case for structuring the discussion of scientific papers. *BMJ.* 1999;318(7193):1224–1225. doi: 10.1136/bmj.318.7193.1224
195. Perneger TV, Hudelson PM. Writing a research article: advice to beginners. *Int J Qual Health Care.* 2004;16(3):191–192. doi: 10.1093/intqhc/mzh053
196. *Annals of Internal Medicine.* Information for authors. Available from: http://www.annals.org/shared/author_info.html. Accessed 10 September 2007.
197. Maldonado G, Poole C. More research is needed. *Ann Epidemiol.* 1999;9(1):17–18. doi: 10.1016/s1047-2797(98)00050-7
198. Phillips CV. The economics of 'more research is needed'. *Int J Epidemiol.* 2001;30(4):771–776. doi: 10.1093/ije/30.4.771
199. Winkleby MA, Kraemer HC, Ahn DK, Varady AN. Ethnic and socioeconomic differences in cardiovascular disease risk factors: findings for women from the Third National Health and Nutrition Examination Survey, 1988–1994. *JAMA.* 1998;280(4):356–362. doi: 10.1001/jama.280.4.356
200. Galuska DA, Will JC, Serdula MK, Ford ES. Are health care professionals advising obese patients to lose weight? *JAMA.* 1999;282(16):1576–1578. doi: 10.1001/jama.282.16.1576
201. Spearman C. The proof and measurement of association between two things. *Am J Psychol.* 1904;15(1):72–101. doi: 10.2307/1412159
202. Fuller WA, Hidioglou MA. Regression estimates after correcting for attenuation. *J Am Stat Assoc.* 1978;73(361):99–104. doi: 10.1080/01621459.1978.10480011
203. MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet.* 1990;335(8692):765–774. doi: 10.1016/0140-6736(90)90878-9
204. Phillips AN, Smith GD. How independent are “independent” effects? Relative risk estimation when correlated exposures are measured imprecisely. *J Clin Epidemiol.* 1991;44(11):1223–1231. doi: 10.1016/0895-4356(91)90155-3
205. Phillips AN, Smith GD. Bias in relative odds estimation owing to imprecise measurement of correlated exposures. *Stat Med.* 1992;11(7):953–961. doi: 10.1002/sim.4780110712
206. Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol.* 1980;112(4):564–569. doi: 10.1093/oxfordjournals.aje.a113025
207. Poole C, Peters U, Il'yasova D, Arab L. Commentary: This study failed? *Int J Epidemiol.* 2003;32(4):534–535. doi: 10.1093/ije/dyg197
208. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology.* 1997;8(6):621–628. doi: 10.1097/00001648-199710000-00002
209. Greenland S. Randomization, statistics, and causal inference. *Epidemiology.* 1990;1(6):421–429. doi: 10.1097/00001648-199011000-00003
210. Taubes G. Epidemiology faces its limits. *Science.* 1995;269(5221):164–169. doi: 10.1126/science.7618077
211. Temple R. Meta-analysis and epidemiologic studies in drug development and postmarketing surveillance. *JAMA.* 1999;281(9):841–844. doi: 10.1001/jama.281.9.841
212. Greenberg RS, Shuster JL. Epidemiology of cancer in children. *Epidemiol Rev.* 1985;7:22–48. doi: 10.1093/oxfordjournals.epirev.a036284
213. Kushi LH, Mink PJ, Folsom AR, et al. Prospective study of diet and ovarian cancer. *Am J Epidemiol.* 1999;149(1):21–31. doi: 10.1093/oxfordjournals.aje.a009723
214. Kemmeren JM, Algra A, Meijers JC, et al. Effect of second- and third-generation oral contraceptives on the protein C system in the absence or presence of the factor V Leiden mutation: a randomized trial. *Blood.* 2004;103(3):927–933. doi: 10.1182/blood-2003-04-1285
215. Egger M, May M, Chene G, et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet.* 2002;360(9327):119–129. doi: 10.1016/S0140-6736(02)09411-4
216. Campbell DT. Factors relevant to the validity of experiments in social settings. *Psychol Bull.* 1957;54(4):297–312. doi: 10.1037/h0040950
217. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med.* 1999;130(6):515–524. doi: 10.7326/0003-4819-130-6-199903160-00016
218. Krinsky S, Rothenberg LS. Conflict of interest policies in science and medical journals: editorial practices and author disclosures. *Sci Eng Ethics.* 2001;7(2):205–218. doi: 10.1007/s11948-001-0041-7
219. Bekelman JE, Li Y, Gross C. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA.* 2003;289(4):454–465. doi: 10.1001/jama.289.4.454
220. Davidson RA. Source of funding and outcome of clinical trials. *J Gen Intern Med.* 1986;1(3):155–158. doi: 10.1007/BF02602327
221. Stelfox HT, Chua G, O'Rourke K, Detsky AS. Conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med.* 1998;338(2):101–106. doi: 10.1056/NEJM199801083380206
222. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ.* 2003;326(7400):1167–1170. doi: 10.1136/bmj.326.7400.1167

- 223.** Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA*. 2003;290(7):921–928. doi: 10.1001/jama.290.7.921
- 224.** Barnes DE, Bero LA. Why results on the health effects of passive smoking reach different conclusions. *JAMA*. 1998;279(19):1566–1570. doi: 10.1001/jama.279.19.1566
- 225.** Barnes DE, Bero LA. Industry-funded research and conflict of interest: an analysis of research sponsored by the tobacco industry through the Center for Indoor Air Research. *J Health Polit Policy Law*. 1996;21(3):515–542. doi: 10.1215/03616878-21-3-515
- 226.** Glantz SA, Barnes DE, Bero L, et al. Looking through a keyhole at the tobacco industry. The Brown and Williamson documents. *JAMA*. 1995;274(3):219–224. doi: 10.1001/jama.1995.03530030039032
- 227.** Huss A, Egger M, Hug K, et al. Source of funding and results of studies of health effects of mobile phone use: systematic review of experimental studies. *Environ Health Perspect*. 2007;115(1):1–4. doi: 10.1289/ehp.9149
- 228.** Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *J Nerv Ment Dis*. 2002;190(9):583–592. doi: 10.1097/00005053-200209000-00002
- 229.** Aspinall RL, Goodman NW. Denial of effective treatment and poor quality of clinical information in placebo controlled trials of ondansetron for postoperative nausea and vomiting: a review of published trials. *BMJ*. 1995;311(7009):844–846. doi: 10.1136/bmj.311.7009.844
- 230.** Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457–2465. doi: 10.1001/jama.291.20.2457
- 231.** Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B. Evidence biased medicine-selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ*. 2003;326(7400):1171–1173. doi: 10.1136/bmj.326.7400.1171
- 232.** Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews*. (Issue 2). Art. No.: MR000005. Available from: <http://www.cochrane.org/reviews/en/mr000005.html>. Accessed 10 September. 2005. doi: 10.1002/14651858.MR000005.pub2
- 233.** Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001;357(9263):1191–1194. doi: 10.1016/S0140-6736(00)04337-3
- 234.** Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008–2012. doi: 10.1001/jama.283.15.2008
- 235.** Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134(8):663–694. doi: 10.7326/0003-4819-134-8-200104170-00012
- 236.** Moher D. CONSORT: an evolving tool to help improve the quality of reports of randomized controlled trials. *Consolidated Standards of Reporting Trials*. *JAMA*. 1998;279(18):1489–1491. doi: 10.1001/jama.279.18.1489
- 237.** Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996;276(8):637–639. doi: 10.1001/jama.276.8.637

Для корреспонденции:E-mail: strobe@ispm.unibe.ch

DOI: <https://doi.org/10.17816/DD60393>

Роль системы контроля качества лучевой диагностики онкологических заболеваний в радиомике

А.Н. Хоружая, Е.С. Ахмад, Д.С. Семенов

Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Российская Федерация

АННОТАЦИЯ

Современные методы медицинской визуализации дают возможность качественно и количественно оценить как ткани опухоли, так и пространство вокруг неё. Прогресс в информатике, особенно с участием методов машинного обучения в анализе медицинских изображений, позволяет преобразовывать любые радиологические исследования в поддающиеся анализу наборы данных. Среди этих наборов данных затем можно искать статистически значимые корреляции с клиническими событиями, чтобы впоследствии оценивать их прогностическую значимость и способность предсказывать тот или иной клинический исход. Эта концепция впервые была описана в 2012 г. и получила название «радиомика». Особую значимость она представляет для онкологии, поскольку известно, что каждый тип опухоли может подразделяться на множество различных молекулярно-генетических подтипов, и просто визуальной характеристики сейчас уже недостаточно. А радиомика при абсолютной неинвазивности способна обеспечить врача-радиолога информацией, которую порой может дать только гистологическое исследование биопсийного материала. Однако, как и в любой методике, основанной на использовании больших данных, здесь остро встаёт вопрос о качестве исходной информации данных, потому как это прямым образом может повлиять на исход анализа и дать неверную диагностическую информацию.

В литературном обзоре мы анализируем возможные подходы к обеспечению качества исследований на всех этапах — от технического контроля за состоянием диагностического оборудования до извлечения маркеров визуализации в онкологии и вычисления их корреляции с клиническими данными.

Ключевые слова: радиомика; лучевая диагностика; контроль качества; стандартизация; опухоли; онкологические заболевания.

Как цитировать

Хоружая А.Н., Ахмад Е.С., Семенов Д.С. Роль системы контроля качества лучевой диагностики онкологических заболеваний в радиомике // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 170–184. DOI: <https://doi.org/10.17816/DD60393>

DOI: <https://doi.org/10.17816/DD60393>

The role of the quality control system for diagnostics of oncological diseases in radiomics

Anna N. Khoruzhaya, Ekaterina S. Akhmad, Dmitry S. Semenov

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

ABSTRACT

Modern medical imaging methods allow for both qualitative and quantitative evaluations of tumors and issues surrounding them. Advances in computer science and big data processing are transforming any radiological study into analytic datasets, especially with the use of machine learning in medical image analysis. Among these datasets, statistically significant correlations with clinical events can then be searched for to subsequently assess their predictive value and ability to predict a particular clinical outcome. This concept, known as “radiomics,” was first described in 2012. It is particularly important in oncology because each type of tumor can be subdivided into many different molecular genetic subtypes, and simple visual characteristics are no longer sufficient. Moreover, as an absolutely noninvasive method, radiomics can provide a radiologist with additional information that would otherwise be unavailable without a histological examination of biopsy material. However, as with any methodology based on the use of big data, the question of the quality of the initial data becomes critical, because this can directly affect the outcome of the analysis and provide incorrect diagnostic information.

In this literature review, we examine potential approaches to ensuring the quality of research at all stages, from technical control of the state of diagnostic equipment to the extraction of imaging markers in oncology and the calculation of their correlation with clinical data.

Keywords: radiomics; radiology; quality assurance; quality control; tumors; cancer; standardization.

To cite this article

Khoruzhaya AN, Akhmad ES, Semenov DS. The role of the quality control system for diagnostics of oncological diseases in radiomics. *Digital Diagnostics*. 2021;2(2):170–184. DOI: <https://doi.org/10.17816/DD60393>

Received: 09.02.2021

Accepted: 31.05.2021

Published: 04.07.2021

DOI: <https://doi.org/10.17816/DD60393>

肿瘤疾病放射诊断质量控制系统在放射组学中的作用

Anna N. Khoruzhaya, Ekaterina S. Akhmad, Dmitry S. Semenov

Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

简评

现代医学成像方法可以定性和定量地评估肿瘤组织及其周围的空间。计算机科学的进步，特别是机器学习方法在医学图像分析中的参与，允许将任何放射学研究转变为可分析的数据集。在这些数据集中，可以寻找有统计学意义的相关性与临床事件，以便随后评估其预后意义和预测不同临床结果的能力。这个概念在2012年首次被描述并称为“放射组学”。这对于肿瘤学特别重要，因为已知每种类型的肿瘤可以分为许多不同的分子遗传亚型，而仅仅是视觉特征已经不够了。在绝对非侵入性的情况下，放射组学能够为放射科医生提供有时只有活检材料的组织学检查才能提供的信息。然而，正如在任何基于使用大数据的方法中一样，存在关于初始数据信息的质量的尖锐问题，因为这可能直接影响分析的结果并给出不正确的诊断信息。

在文献综述中，我们分析了确保各个阶段研究质量的可能方法 – 从诊断设备状态的技术控制到提取肿瘤学中的成像标记并计算其与临床数据的相关性。

关键词：放射学；放射诊断；质量控制；标准化；肿瘤；肿瘤疾病。

引用本文：

Khoruzhaya AN, Akhmad ES, Semenov DS. 肿瘤疾病放射诊断质量控制系统在放射组学中的作用. *Digital Diagnostics*. 2021;2(2):170–184.

DOI: <https://doi.org/10.17816/DD60393>

收到: 09.02.2021

接受: 31.05.2021

发布日期: 04.07.2021

ВВЕДЕНИЕ

Прогресс в области лучевых методов визуализации привёл к тому, что их роль среди всего спектра методов управления опухолевыми процессами значительно расширилась — от диагностики первичных очагов и обнаружения метастазов до мониторинга реакции на лечение и прогнозирования индивидуальных исходов пациентов. Однако простого визуального анализа опухоли методами лучевой диагностики сейчас недостаточно, поскольку известно, что каждый тип опухоли может подразделяться на множество различных молекулярно-генетических подтипов. Соответственно, к каждому из них нужен свой лечебно-диагностический подход. Здесь со стороны диагностики большую помощь может оказать радиомика.

Радиомика представляет собой направление не просто визуального анализа медицинских изображений, а извлечения из них большого числа количественных признаков, которые позволят провести более глубокий анализ и комплексно оценить, к примеру, опухолевые фенотипы и другие патанатомические свойства поражённых тканей, оценить биологические характеристики опухоли и спрогнозировать ответ на лечение [1, 2]. Например, солидный рак неоднороден во времени и пространстве, что ограничивает использование молекулярного анализа, основанного на инвазивной биопсии, но даёт огромный потенциал для медицинской визуализации, которая позволяет неинвазивно обнаружить внутриопухолевую гетерогенность [3–5].

Переход к количественному анализу требует разработки автоматизированных и воспроизводимых методологий анализа для извлечения дополнительной информации из изображений [6]. И здесь остро встаёт вопрос о качестве исходных данных, так как это может повлиять

на исход анализа и дать неверную диагностическую информацию, что скажется на клинической значимости выявляемых показателей и здоровье пациентов [7, 8].

Таким образом, целью данного литературного обзора стал анализ возможных подходов к обеспечению качества исследований лучевой диагностики на всех этапах — от технического контроля за состоянием диагностического оборудования до извлечения маркеров визуализации в онкологии и вычисления их корреляции с клиническими данными.

Поиск литературы проводили в базах данных PubMed, GoogleScholar и eLibrary на английском и русском языках. Для поиска в PubMed и GoogleScholar использовали запросы «radiomics», «cancer and tumors», «standardization», «quality assurance or quality control».

МЕТОДОЛОГИЯ РАДИОМИКИ

Получение изображения

Первый шаг радиомики заключается в получении изображений с использованием методов лучевой диагностики: магнитно-резонансной томографии (МРТ), компьютерной томографии (КТ), позитронно-эмиссионной томографии, совмещённой с компьютерной томографией (ПЭТ/КТ) (рис. 1). Методы лучевой диагностики дают различную и часто дополняющую информацию о физических, кинетических свойствах тканей, метаболизме и др. Например, анализ на основе размера или объёма патологической структуры можно получить с помощью анатомической МРТ или КТ. Перфузию можно определить по серии динамических МРТ или КТ с контрастным усилением. Диффузионно-взвешенная МРТ может использоваться для оценки микроциркуляции в тканях и оценки целлюлярности. Метаболические изменения, такие как скорость метаболизма глюкозы, могут быть

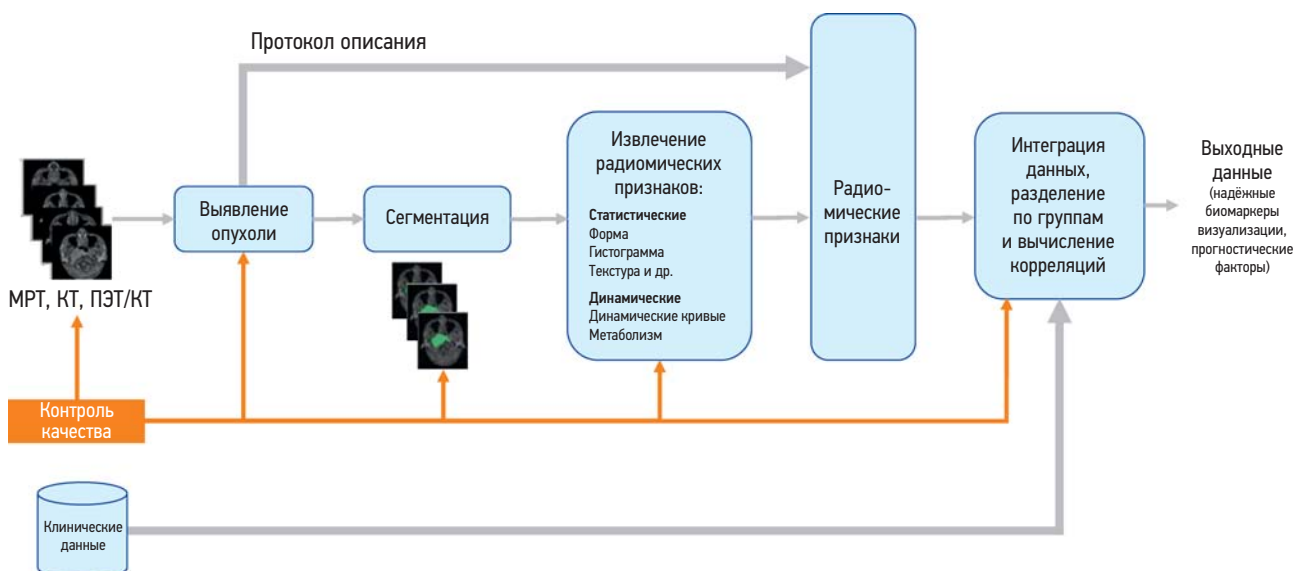


Рис. 1. Схема радиомического анализа изображений лучевой диагностики с указанием роли системы контроля качества.

измерены с использованием ПЭТ/КТ и фтордезоксиглюкозы. Кроме того, в ходе клинических испытаний могут быть предложены и другие, дополнительные биомаркеры [9, 10].

Исторически устройства визуализации были разработаны для субъективной интерпретации изображений, позволяя клиницистам определять, например, наличие и местоположение поражения. Последующие технические инновации были сосредоточены на улучшении качества изображения, сокращении времени сканирования или интеграции с обрабатывающими машинами. Эти устройства не были предназначены в первую очередь для обеспечения количественного измерения воспроизводимым образом. Протоколы для стандартизации получения изображений часто отсутствуют. Кроме того, могут существовать большие различия в параметрах реконструкции. Н. Kim и соавт. [11] изучили влияние фильтров реконструкции на радиомические признаки, выделенные по КТ-изображениям пациентов с раком лёгких, и пришли к выводу, что взаимосвязь статистически значима и не следует использовать настройки реконструкции взаимозаменяемо. N. Ohri и соавт. [12] оценили изменчивость радиомических характеристик, извлечённых из ПЭТ/КТ при различных режимах сбора данных, алгоритмов реконструкции, постфильтрации и числа итераций. Было показано, что 40 из 50 признаков имеют существенную — до 30% — изменчивость. При выполнении МРТ из-за амплитуды градиентного магнитного поля сканера, используемой импульсной последовательности, метода введения контрастного вещества, выборки траектории в k-пространстве и других факторов изменчивость признаков может варьировать более значительно [13]. Поскольку качество данных зависит от надёжности протоколов сбора данных, используемых в клинических центрах, влияние этих изменений на стабильность радиомических признаков необходимо тщательно изучить и проанализировать в будущих исследованиях.

Новые методы обработки изображений

Следующим этапом выделения радиомических признаков является обработка изображений. Так, выделение области интереса (region of interest, ROI) и объёма является фундаментальной задачей, например, в онкологической практике [14]. Ручное описание опытными рентгенологами или радиологами рассматривается как золотой стандарт, но оно требует больших временных затрат и имеет высокую степень межоператорской или даже внутриоператорской изменчивости. Для определения ROI чаще используются автоматизированные или полуавтоматизированные методы, такие как определение пороговых значений, классификаторы, кластеризация, марковские модели случайных полей, искусственные нейронные сети, деформируемые модели и некоторые другие [15].

Хотя автоматизация может предоставить новые возможности для стандартизации методов сегментации, проблемы, связанные со сложной анатомией или областями с низким контрастом мягких тканей, сохраняются, и часто требуется ручная корректировка, которую проводит опытный врач. Одним из методов полуавтоматической сегментации, позволяющим избежать ошибок, является использование цифровой биопсии, при которой производится выборка только определённых сегментов на основе значений интенсивности и текстуры [16]. Для сегментации или отбора изображений появились и стали использоваться также передовые методы машинного обучения [17].

Существует несколько крупных инициатив, направленных на разработку решений автоматической сегментации с использованием глубокого обучения. К ним относятся, например, Google DeepMind, Microsoft Project InnerEye, Mirada DLCExpert. Было показано, что данные автоматизированные инструменты сегментации повышают эффективность при воссоздании структуры, особенно для органов, подверженных риску [18, 19]. В ближайшем будущем инструменты сегментации, основанные на глубоком обучении, могут стать достаточно надёжными для рутинных исследований.

Извлечение признаков, объединение по группам и интеграция данных

Основным этапом радиомики является извлечение многомерных наборов данных — радиомических признаков — для количественной оценки выделенных на изображении объёмов интереса (ROI или VOI) [20]. Признаки, извлечённые из изображений, можно разделить на статические и динамические группы.

Особенности статических признаков. Среди множества статических признаков выделяют две категории: морфологические и статистические [21]. Морфологические признаки используются для определения трёхмерных (3D) характеристик формы, таких как объём и площадь поверхности, а также сферичность (насколько близко трёхмерный объём напоминает сферу). Статистические признаки используются для математической оценки распределения градаций серого в пределах ROI или объёма интереса. Так, к признакам первого порядка относятся среднее значение, стандартное отклонение, перцентили, эксцесс и асимметрия. Они используются для характеристики общей изменчивости интенсивности. Признаки второго порядка характеризуют текстуру выделенной области путём анализа взаимосвязи между отдельными вокселями внутри ROI или области, т.е. проявляют локальное распределение.

Особенности динамических признаков. Для количественной оценки динамического распределения контрастного вещества или другого индикатора в пределах области (которая может быть одним или несколькими

вокселями) обычно используется фармакокинетическое моделирование. В целом фармакокинетическое моделирование рассматривает концентрацию контраста как функцию артериального ввода и остаточного спада контраста внутри ROI. Внутрисосудистое и интерстициальное пространство можно моделировать при разных допущениях. Например, наиболее широко используемая кинетическая модель, модель Тофта, предполагает мгновенное смешивание контраста во внутрисосудистом и интерстициальном пространстве, а расширенная модель Тофта учитывает эффект задержки концентрации контрастного вещества в ткани. Модель гомогенности адиабатической ткани объясняется тем фактом, что концентрация контрастного вещества в объёме распределения вне сосудов изменяется медленнее по сравнению с концентрацией во внутрисосудистом пространстве. Таким образом, модель предполагает, что существует конечное время прохождения контрастных веществ от артериальной до венозной фазы.

В целом существующий аналитический конвейер обычно включает в себя тысячи извлечённых радиомических характеристик, и ожидается, что их число будет расти по мере поступления новых данных. Однако к клинически значимым признакам будут относиться не все выделенные, а наиболее надёжные, коррелирующие с клиническими данными для возможности прогнозирования хода заболевания.

Вычисление корреляций, выделение прогностических факторов

Как и во многих других областях, где использован суффикс «-omics», количество входных переменных часто намного превышает количество пациентов. Для того чтобы снизить вероятность ложноположительных результатов, требуется выбор определённых признаков или уменьшение размеров зоны поиска, и обычно используются подходы ранжирования оценок, основанные на фильтрах: например, анализ Уилкоксона и анализ главных компонент. Это можно осуществить при использовании либо одномерных методов, когда критерий оценки зависит только от релевантности объекта, либо многовариантных методов, когда используют взвешенную сумму, чтобы максимизировать релевантность и минимизировать избыточность [22–25]. Выбор объекта также можно объединить с классификацией объекта в одну модель.

Как только набор характеристик получен, можно построить модель, управляемую данными. Эти модели включают в себя контролируемые и неконтролируемые подходы [21, 26]. Неуправляемый анализ предоставляет не переменную результата, а, скорее, некую сводку информации. Наиболее часто для его графического отображения используется тепловая карта, на которой

одновременно выявляются кластерные структуры в матрице данных. В отличие от этого, в ходе контролируемого анализа создаются модели, которые пытаются разделить данные относительно результата лечения. Типичные методы классификации включают в себя традиционную логистическую регрессию или более продвинутые методы машинного обучения.

Выделенные радиомические признаки, которые точно коррелируют с клиническими данными и результатами молекулярного анализа, могут быть отнесены к биомаркерам визуализации. В то время как классические биомаркеры получают путём гистологического и молекулярного исследования тканей опухоли, т.е. инвазивным методом, биомаркеры визуализации позволяют неинвазивно охарактеризовать патологию. Кроме того, они являются надёжными индикаторами нормальных или патологических процессов в тканях или ответов опухоли на какое-либо вмешательство.

КОНТРОЛЬ КАЧЕСТВА И СТАНДАРТИЗАЦИЯ ПАРАМЕТРОВ В РАДИОМИКЕ

Для обеспечения качества радиомических признаков и надёжности биомаркеров визуализации необходимо повышение точности измерений (рис. 2), которая определяется величиной смещения или абсолютной ошибкой получаемых данных и вариативностью значений (повторяемость и воспроизводимость, определяемые как разброс измеренных значений). Данные показатели достигаются путём введения в отделениях лучевой диагностики испытаний по контролю качества: приёмочных, периодических и внутреннего контроля (испытания на постоянство параметров) [27]. Приёмочные испытания проводят при монтаже оборудования для установления соответствия проверяемых характеристик предельным значениям производителя. В случае подтверждения соответствия параметров персонал медицинской организации выполняет первые испытания на постоянство параметров, в ходе которых устанавливаются базовые значения для дальнейшего контроля качества. Внутренний контроль или испытания на постоянство параметров является важным в системе контроля качества, поскольку позволяет предсказать ухудшение диагностического качества изображения. В России периодические испытания включают в себя контроль расширенного списка параметров и проводятся аккредитованными испытательными лабораториями.

В зарубежной практике распространено включение в состав штата кабинетов МРТ, КТ, ПЭТ/КТ технического персонала. Например, большую роль отводят медицинским физикам, важная задача которых состоит в оптимизации и стандартизации проведения исследования,

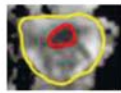
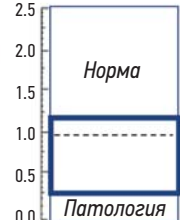
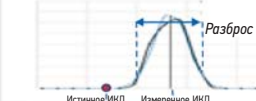
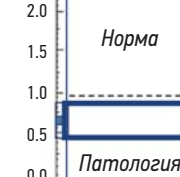
ПРОБЛЕМА	Невозможность проведения дифференциальной диагностики	Источники низкой точности измерений: 1) параметры оборудования; 2) параметры исследования; 3) методы реконструкции и постобработки и д.т.	 ADC 0.659 ± 0.434 
РЕШЕНИЕ	Повышение точности измерений	Контроль показателей: 1) смещение (bias); 2) вариативность (повторяемость, воспроизводимость)	
ДЕЙСТВИЕ	Контроль качества	Система контроля качества: 1) настройка и периодический контроль качества; 2) стандартизация протоколов исследований; 3) единая методология подготовки пациента и проведения исследования 4) контроль сегментации и анализа данных	
РЕЗУЛЬТАТ	Получение биомаркера визуализации	Обеспечение высокой точности измерений позволит выполнить радиомический анализ и получать надежные биомаркеры визуализации для повышения диагностической ценности	 ADC 0.659 ± 0.216 

Рис. 2. Обоснование внедрения системы контроля качества в радиомике.

а также на них лежит ответственность за осуществление контроля качества оборудования лучевой диагностики и организация системы безопасности при проведении исследования [28]. В России на данный момент отсутствуют требования наличия такого персонала в штате кабинетов лучевой диагностики, а медицинские персонал не обладает необходимыми компетенциями осуществления контроля качества для применения в радиомике.

Проведение мероприятий по обеспечению контроля качества оборудования лучевой диагностики необходимо для достижения достоверности и клинически приемлемой повторяемости измерений, что поддерживается Радиологическим обществом Северной Америки (Radiological Society of North America, RSNA), Европейским обществом радиологии (European Society of Radiology, ESR) и др. Так, в результате сотрудничества между членами Сети количественной визуализации (Quantitative imaging Network, QIN; США) и Национальным институтом стандартизации и технологий (National Institute of Standards and Technology, NIST) были разработаны фантомы для проведения контроля качества при выполнении клинических исследований [29, 30].

В результате радиомического анализа формируются взаимосвязи между найденными признаками и клиническими данными с целью проверки построенной модели и оценки достоверности выходной информации, для новых пациентов выполняется её валидация [31, 32]. Используют литературные данные, тестирования

на валидационном наборе данных или на данных из других медицинских организаций, чтобы получить возможность обобщения [31].

Стандартизация протоколов исследований

Так как МРТ, КТ и ПЭТ/КТ-изображения подвержены влиянию артефактов и шума, необходимо следовать стандартным методам подготовки к исследованию: исключить из области сканирования инородные предметы, которые способствуют искажениям; убедиться в соблюдении установленных правил укладки пациента для лучшей визуализации. Пациенту должно быть удобно находиться долгое время без движения.

Кроме того, на радиомические признаки большое влияние оказывают размер вокселя и интенсивность сигнала, поэтому при настройке сканирования важно обеспечить стандартизацию протоколов [32, 33]. При выполнении ПЭТ/КТ и КТ следует учесть и влияние фильтров реконструкций на качество изображения и интенсивность сигнала: должен быть выбран фильтр, не дающий потери полезного сигнала и обеспечивающий высокую воспроизводимость радиомических признаков [34].

В рамках предобработки изображения проводится масштабирование матрицы изображения и приведение его к изотропному виду (квадратному) [35]. Рекомендуется также выполнить нормирование интенсивности сигнала к одной шкале, особенно для МРТ. С этой целью используются статистические методы, например ANTsR и WhiteStripe [36]. При выполнении МРТ могут возникать явления неоднородности интенсивности сигнала,

которые вызваны не биологическими свойствами тканей, а техническими факторами. В таких случаях необходимо провести коррекцию на данную неоднородность, что должно быть включено в систему контроля качества выполняемых процедур.

Контроль постобработки

В процессе постобработки должны использоваться инструменты и алгоритмы с подтверждённой точностью их работы [36]. Например, для последующего корректного анализа радиомических признаков важно использование на этапе сегментации качественных инструментов. Если ранее применялись полуавтоматические алгоритмы с ручной коррекцией сегментации, то сейчас всё больше появляется алгоритмов на основе технологий искусственного интеллекта [37], которые должны проходить соответствующие испытания [38].

Контроль выделенных радиомических признаков и валидация биомаркеров визуализации

Указанные принципы стандартизации и контроля качества выполнения исследований и процедур пред- и постобработки изображений необходимы для обеспечения качества (смещение и вариативность) радиомических признаков, а также надёжности биомаркеров визуализации [39].

На данном этапе применяют средства контроля качества — фантомы, позволяющие оценить смещение и воспроизводимость выделяемых признаков. Фантомы могут быть как цифровые, так и физические, изготовленные с применением веществ заданных параметров [40, 41]. Например, для проведения многоцентровых исследований рака молочной железы используется соответствующий фантом, который позволяет оценивать воспроизводимость и точность проводимых исследований [42].

Для анализа влияния на вариативность параметров сканирования, методологии выполнения исследования и постобработки выполняют многократное сканирование фантома с разными условиями, после чего рассчитывается вариативность измерений и сравнивается с пороговым значением: Европейское медицинское агентство (European Medicines Agency, EMA) рекомендует устанавливать не более 15% [39].

Точность оценивается в процессе исследований на фантомах или на образцах тканей и соответствует относительной ошибке при сравнении истинного значения признаков (ground truth) и измеренных. В процессе валидации биомаркера визуализации рекомендуют устанавливать пороговое значение, успешного прохождения оценки, на уровне 15% [39].

Данное направление радиомики находится в стадии развития и может стать в ближайшем будущем эффективным методом диагностики опухолей и прогностического

анализа процесса. Мы считаем, что число исследований по данному направлению увеличится с внедрением алгоритмов искусственного интеллекта, например для построения зависимостей между выделенными признаками и клиническими данными. Однако без выполнения описанных подходов контроля качества на всех этапах невозможно будет получить надёжных данных, т.е. воспроизводимых на других популяциях, другом оборудовании со смещением показателей в рамках установленного предела. В Центре диагностики и телемедицины ранее были разработаны фантомы для контроля количественных режимов МРТ (с показателями диффузии), КТ (с показателями минеральной плотности кости). С нашей точки зрения, в данной работе важным является взаимодействие технических специалистов (медицинских физиков, инженеров) и медицинского персонала для разработки фантомов с заданными показателями точности измерений планирования исследования радиомических признаков и дальнейшего получения биомаркеров визуализации.

РОЛЬ РАЗРАБОТКИ БИОМАРКЕРОВ ВИЗУАЛИЗАЦИИ

В последние годы подходы к стандартизации радиомических показателей пытались усовершенствовать, определив стандартные протоколы сбора данных. Особые усилия для этого прилагали QIN, созданная Национальным институтом рака (National Cancer Institute, NCI), а также RSNA, Альянс биомаркеров количественной визуализации (Quantitative Imaging Biomarkers Alliance, QIBA) и др. NCI в 2010 г. запустил инициативу Центров количественного превосходства в области визуализации (Cancer Institute Centers for Quantitative Image Excellence), и создание национальной сети клинических исследований (National Clinical Trials Network, NCTN) стало ключевым направлением этих усилий [43]. Центры для количественного улучшения качества изображений создают ПЭТ/КТ, КТ и МР-фантомы, а также протоколы для стандартизации, а QIBA обеспечивает консенсусные решения в отношении точности измерения количественных изображений биомаркеров и требований/процедур, необходимых для достижения этого уровня точности [29, 35, 36, 44, 45].

С тех пор как термин «радиомика» появился в научной литературе, были опубликованы сотни радиомических исследований, направленных на улучшение качества диагностики, лечения рака и его прогнозирования. Всё больше работ демонстрируют ценность биомаркеров визуализации как дополнительного инструмента для принятия решений клиницистами и роль в этом алгоритмов машинного обучения [46].

Одним из самых ранних применений метода, основанного на радиомике, является успешное обнаружение опухолей при визуализации рака лёгких и молочных желёз.

Рак молочных желёз — патология, которая чаще всего встречается у женщин во всём мире. Поскольку это известное гетерогенное заболевание, точный диагноз и раннее предсказание реакции на лечение являются ключевыми моментами клинической практики [47]. В нескольких исследованиях использовалась радиомика для прогнозирования подтипа рака груди или статуса ER, PR, Ki67 и HER2 на маммографии [48], ПЭТ/КТ [49, 50] и МРТ [51, 52]. В дополнение к характеристике рака груди радиомика может также обеспечить неинвазивный подход для прогнозирования метастазов в сторожевые лимфатические узлы [53].

Большинство радиологических исследований рака груди сосредоточены на оценке ответа на терапию. Н.М. Chan и соавт. [54] разработали автоматизированный метод прогнозирования отсутствия или недостаточного ответа на лечение у пациенток с ранним раком груди при помощи МРТ. В большинстве других исследований предпринимались попытки получить биомаркер полной морфологической ремиссии (pCR) при неоадъювантной химиотерапии, что является горячей темой обсуждения в исследованиях, посвящённых раку молочных желёз. Так, N.M. Braman и соавт. [55] обнаружили, что внутри- и перипухолевые характеристики, обнаруживаемые на МРТ с динамическим контрастированием, могут спрогнозировать pCR до начавшегося лечения. Другие исследования также показали, что T1WI, T2WI и DWI могут помочь в обнаружении pCR [56, 57].

Всё чаще встречаются радиологические исследования, посвящённые прогнозированию рака груди. Например, Н. Park и соавт. [58] разработали алгоритм, сочетающий биомаркеры визуализации МРТ и клиническую информацию для индивидуальной оценки выживаемости пациенток с раком груди.

Рак лёгкого — это наиболее опасный вид рака, и его распространённость также продолжает расти во всём мире. Одним из наиболее важных диагностических приложений радиомики является скрининг рака лёгких. N. Nasrullah и соавт. [59] предложили модель глубокого обучения, основанную на исследованиях КТ грудной клетки из датасета LIDC-IDRI, и достигли хороших результатов выявления злокачественных узелков в лёгких с чувствительностью 94% и специфичностью 91%. B.W. Carter и соавт. [60] проводили скрининговое исследование пациентов, у которых был обнаружен рак лёгких в наборе данных Национального исследования рака лёгких (NLST), с применением низкодозной КТ. Им удалось получить точность прогнозирования в 80 и 79% для узелков, которые разовьются в злокачественные новообразования через один или два года соответственно.

Радиомика позволяет на предоперационном этапе определять стадийность рака лёгких по метастазированию опухолевых узлов (tumor nodules metastasis, TNM) [61, 62], что важно для принятия решения об оперативном вмешательстве. Помимо этого, методика может

применяться для обнаружения специфических генетических мутаций рака лёгких, например статуса по гену *EGFR* [63], что может помочь врачам избрать оптимальную тактику терапии. X. Fave и соавт. [64] использовали дельта-радиомические характеристики для прогнозирования исходов у пациентов с немелкоклеточным раком лёгких III стадии во время лучевой терапии. Их результаты предполагают, что изменение радиомических характеристик вследствие лучевой терапии будет показателем ответа опухоли. T.P. Coroller и соавт. [65] обнаружили, что радиомические особенности КТ перед лечением могут прогнозировать патологический ответ после неоадъювантной химиолучевой терапии у пациентов с распространённым немелкоклеточным раком лёгких.

В последние годы радиомика всё чаще применяется для диагностики, прогнозирования ответа на лечение и отдалённых исходов опухолей нервной системы [26, 66, 67], головы и шеи [68, 69], желудочно-кишечного тракта [70, 71], рака простаты [72, 73] и некоторых других форм онкологических заболеваний [74].

ЗАКЛЮЧЕНИЕ

Раннее выявление и идентификация опухолей, их гетерогенности и фенотипических признаков могут оказать неоценимую пользу для стратификации пациентов, определения вариантов последующего лечения и прогнозирования его эффектов. Радиомический анализ диагностических исследований предоставляет необходимую для этого информацию, но только при условии качественно собранных и обработанных данных. Все эти процессы необходимо стандартизировать и оптимизировать с использованием различных методов контроля качества, причём на каждом из этапов — от получения изображений до валидации биомаркеров визуализации. Кроме того, для установления прогностической значимости биомаркеров нужно обязательно учитывать клиническую информацию, на основе которой проводится поиск клинических корреляций. Только качественное выполнение всех этих критериев может сделать инструмент биомаркерной визуализации действительно полезным для врачей и нужным для пациентов.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении поисково-аналитической работы и подготовке статьи.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. А.Н. Хоружая — сбор и анализ литературы, написание текста; Е.С. Ахмад — анализ литературы, формирование вопроса исследования; Д.С. Семенов — обработка полученных результатов, систематизация и финальное редактирование

обзора. Авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Funding source. The authors declare that there is no external funding for the exploration and analysis work.

Competing interests. The authors declare no obvious and potential conflicts of interest related to the publication of this article.

Authors' contribution. A.N. Khoruzhaya — collecting and analyzing literature, writing text; E.S. Akhmad — analysis of literature, formation of a research question; D.S. Semenov — processing of the obtained results, systematization and final editing of the review. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

СПИСОК ЛИТЕРАТУРЫ

1. Kumar V., Gu Y., Basu S., et al. Radiomics: The process and the challenges // *Magn Reson Imaging*. 2012. Vol. 30, N 9. P. 1234–1248. doi: 10.1016/j.mri.2012.06.010
2. Papanikolaou N., Matos C., Koh D.M. How to develop a meaningful radiomic signature for clinical use in oncologic patients // *Cancer Imaging*. 2020. Vol. 20, N 1. P. 33. doi: 10.1186/s40644-020-00311-4
3. Aerts H.J., Grossmann P., Tan Y., et al. Defining a radiomic response phenotype: A pilot study using targeted therapy in NSCLC // *Sci Rep*. 2016. Vol. 6. P. 33860. doi: 10.1038/srep33860
4. Coroller T.P., Grossmann P., Hou Y., et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma // *Radiother Oncol*. 2015. Vol. 114, N 3. P. 345–350. doi: 10.1016/j.radonc.2015.02.015
5. Lopez C.J., Nagornaya N., Parra N.A., et al. Association of radiomics and metabolic tumor volumes in radiation treatment of glioblastoma multiforme // *Int J Radiat Oncol Biol Phys*. 2017. Vol. 97, N 3. P. 586–595. doi: 10.1016/j.ijrobp.2016.11.011
6. De Souza N.M., Achten E., Alberich-Bayarri A., et al. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR) // *Insights Imaging*. 2019. Vol. 10, N 1. P. 87. doi: 10.1186/s13244-019-0764-0
7. Jones E.F., Buatti J.M., Shu H.K., et al. Clinical trial design and development work group within the quantitative imaging network // *Tomography*. 2020. Vol. 6, N 2. P. 60–64. doi: 10.18383/j.tom.2019.00022
8. European Society of Radiology (ESR). ESR Statement on the validation of imaging biomarkers // *Insights Imaging*. 2020. Vol. 11, N 1. P. 76. doi: 10.1186/s13244-020-00872-9
9. Grimm L.J., Zhang J., Mazurowski M.A. Computational approach to radiogenomics of breast cancer: Luminal A and luminal B molecular subtypes are associated with imaging features on routine breast MRI extracted using computer vision algorithms // *J Magn Reson Imaging*. 2015. Vol. 42, N 4. P. 902–907. doi: 10.1002/jmri.24879
10. Nie K., Shi L., Chen Q., et al. Rectal cancer: Assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI // *Clin Cancer Res*. 2016. Vol. 22, N 21. P. 5256–5264. doi: 10.1158/1078-0432.CCR-15-2997
11. Kim H., Park C.M., Lee M., et al. Impact of reconstruction algorithms on ct radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability // *PLoS One*. 2016. Vol. 11, N 10. P. e0164924. doi: 10.1371/journal.pone.0164924
12. Ohri N., Duan F., Snyder B.S., et al. Pretreatment 18F-FDG PET textural features in locally advanced non-small cell lung cancer: Secondary analysis of ACRIN 6668/RTOG 0235 // *J Nucl Med*. 2016. Vol. 57, N 6. P. 842–848. doi: 10.2967/jnumed.115.166934
13. Zhang B., Tian J., Dong D., et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma // *Clin Cancer Res*. 2017. Vol. 23, N 15. P. 4259–4269. doi: 10.1158/1078-0432.CCR-16-2910
14. Nakatsugawa M., Cheng Z., Goatman K.A., et al. Radiomic analysis of salivary glands and its role for predicting xerostomia in irradiated head and neck cancer patients // *Int J Radiat Oncol Biol Phys*. 2016. Vol. 96, N 2, Suppl. P. S217. doi: 10.1016/j.ijrobp.2016.06.539
15. Shafiee M.J., Chung A.G., Khalvati F., et al. Discovery radiomics via evolutionary deep radiomic sequencer discovery for pathologically proven lung cancer detection // *J Med Imaging*. 2016. Vol. 4, N 4. P. 041305. doi: 10.1117/1.JMI.4.4.041305
16. Echegaray S., Nair V., Kadoch M., et al. A rapid segmentation-insensitive “Digital Biopsy” method for radiomic feature extraction: method and pilot study using ct images of non-small cell lung cancer // *Tomography*. 2016. Vol. 2, N 4. P. 283–294. doi: 10.18383/j.tom.2016.00163
17. Li H., Galperin-Aizenberg M., Pryma D., et al. Unsupervised machine learning of radiomic features for predicting treatment response and overall survival of early stage non-small cell lung cancer patients treated with stereotactic body radiation therapy // *Radiother Oncol*. 2018. Vol. 129, N 2. P. 218–226. doi: 10.1016/j.radonc.2018.06.025
18. Tajbakhsh N., Shin J.Y., Gurudu S.R., et al. Convolutional neural networks for medical image analysis: full training or fine tuning? // *IEEE Transactions on Medical Imaging*. 2016. Vol. 35, N 5. P. 1299–1312. doi: 10.1109/TMI.2016.2535302
19. Elguindi S., Zelefsky M.J., Jiang J., et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy // *Phys Imaging Radiat Oncol*. 2019. Vol. 12. P. 80–86. doi: 10.1016/j.phro.2019.11.006
20. Gillies R.J., Kinahan P.E., Hricak H. Radiomics: images are more than pictures, they are data // *Radiology*. 2016. Vol. 278, N 2. P. 563–577. doi: 10.1148/radiol.2015151169
21. Buckler A.J., Bresolin L., Dunnick N.R., et al. Quantitative imaging test approval and biomarker qualification: Interrelated but distinct activities // *Radiology*. 2011. Vol. 259, N 3. P. 875–884. doi: 10.1148/radiol.10100800
22. Alobaidli S., McQuaid S., South C., et al. The role of texture analysis in imaging as an outcome predictor and potential tool in radio-

- therapy treatment planning // *Br J Radiol*. 2014. Vol. 87, N 1042. P. 20140369. doi: 10.1259/bjr.20140369
23. Li H., Giger M.L., Lan L., et al. Comparative analysis of image-based phenotypes of mammographic density and parenchymal patterns in distinguishing between BRCA1/2 cases, unilateral cancer cases, and controls // *J Med Imaging*. 2014. Vol. 1, N 3. P. 031009. doi: 10.1117/1.JMI.1.3.031009
24. Goh V., Ganeshan B., Nathan P., et al. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker // *Radiology*. 2011. Vol. 261, N 1. P. 165–171. doi: 10.1148/radiol.11110264
25. Yip C., Davnall F., Kozarski R., et al. Assessment of changes in tumor heterogeneity following neoadjuvant chemotherapy in primary esophageal cancer // *Dis Esophagus*. 2015. Vol. 28, N 2. P. 172–179. doi: 10.1111/dote.12170
26. Park J.E., Kim H.S. Radiomics as a quantitative imaging biomarker: practical considerations and the current standpoint in neuro-oncologic studies // *Nucl Med Mol Imaging*. 2018. Vol. 52, N 2. P. 99–108. doi: 10.1007/s13139-017-0512-7
27. Сергунова К.А., Ахмад Е.С., Семенов Д.С., и др. Участие медицинских физиков в обеспечении контроля качества оборудования и безопасности пациентов при магнитно-резонансной томографии // *Медицинская физика*. 2020. № 3. С. 78–85.
28. Clements J.B., Baird C.T., de Boer S.F., et al. AAPM medical physics practice guideline 10.a: Scope of practice for clinical medical physics // *J Appl Clin Med Phys*. 2018. Vol. 19, N 6. P. 11–25. doi: 10.1002/acm2.12469
29. Shukla-Dave A., Obuchowski N.A., Chenevert T.L., et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials // *J Magn Reson Imaging*. 2019. Vol. 49, N 7. P. e101–e121. doi: 10.1002/jmri.26518
30. Russek S.E., Boss M., Jackson E.F., et al. Characterization of NIST/ISMRM MRI System Phantom // *Proc Intl Soc Mag Reson Med*. 2012. Vol. 20. P. 2456.
31. Kuo M.D., Jamshidi N. Behind the numbers: Decoding molecular phenotypes with radiogenomics –guiding principles and technical considerations // *Radiology*. 2014. Vol. 270, N 2. P. 320–325. doi: 10.1148/radiol.13132195
32. Narang S., Lehrer M., Yang D., et al. Radiomics in glioblastoma: current status, challenges and potential opportunities // *Transl Cancer Res*. 2016. Vol. 5, N 4. P. 383–397. doi: 10.21037/tcr.2016.06.31
33. O'Connor J.P., Aboagye E.O., Adams J.E., et al. Imaging biomarker roadmap for cancer studies // *Nat Rev Clin Oncol*. 2017. Vol. 14, N 3. P. 169–186. doi: 10.1038/nrclinonc.2016.162
34. Buizza G., Toma-Dasu I., Lazzeroni M., et al. Early tumor response prediction for lung cancer patients using novel longitudinal pattern features from sequential PET/CT image scans // *Phys Med*. 2018. Vol. 54. P. 21–29. doi: 10.1016/j.ejmp.2018.09.003
35. Raunig D.L., McShane L.M., Pennello G., et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment // *Stat Methods Med Res*. 2015. Vol. 24, N 1. P. 27–67. doi: 10.1177/0962280214537344
36. Obuchowski N.A., Reeves A.P., Huang E.P., et al. Quantitative imaging biomarkers: a review of statistical methods for computer algorithm comparisons // *Stat Methods Med Res*. 2015. Vol. 24, N 1. P. 68–106. doi: 10.1177/0962280214537390
37. Elguindi S., Zelefsky M.J., Jiang J., et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy // *Phys Imaging Radiat Oncol*. 2019. Vol. 12. P. 80–86. doi: 10.1016/j.phro.2019.11.006
38. Морозов С.П., Владимирский А.В., Кляшторный В.Г., и др. Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика). Методические рекомендации. Москва, 2019. 33 с.
39. Sullivan D.C., Obuchowski N.A., Kessler L.G., et al. Metrology standards for quantitative imaging biomarkers // *Radiology*. 2015. Vol. 277, N 3. P. 813–825. doi: 10.1148/radiol.2015142202
40. Shur J., Blackledge M., D'Arcy J., et al. MRI texture feature repeatability and image acquisition factor robustness, a phantom study and in silico study // *Eur Radiol Exp*. 2021. Vol. 5, N 1. P. 2. doi: 10.1186/s41747-020-00199-6
41. Bane O., Hectors S.J., Wagner M., et al. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study // *Magn Reson Med*. 2018. Vol. 79, N 5. P. 2564–2575. doi: 10.1002/mrm.26903
42. He Y., Liu Y., Dyer B.A., et al. 3D-printed breast phantom for multi-purpose and multi-modality imaging // *Quant Imaging Med Surg*. 2019. Vol. 9, N 1. P. 63–74. doi: 10.21037/qims.2019.01.05
43. Scheuermann J.S., Reddin J.S., Opanowski A., et al. Qualification of national cancer institute-designated cancer centers for quantitative PET/CT imaging in clinical trials // *J Nucl Med*. 2017. Vol. 58, N 7. P. 1065–1071. doi: 10.2967/jnumed.116.186759
44. Obuchowski N.A., Barnhart H.X., Buckler A.J., et al. Statistical issues in the comparison of quantitative imaging biomarker algorithms using pulmonary nodule volume as an example // *Stat Methods Med Res*. 2015. Vol. 24, N 1. P. 107–140. doi: 10.1177/0962280214537392
45. Kessler L.G., Barnhart H.X., Buckler A.J., et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions // *Stat Methods Med Res*. 2015. Vol. 24, N 1. P. 9–26. doi: 10.1177/0962280214537333
46. Napel S., Mu W., Jardim-Perassi B.V., et al. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats // *Cancer*. 2018. Vol. 124, N 24. P. 4633–4649. doi: 10.1002/cnrc.31630
47. Рожкова Н.И., Боженко В.К., Бурдина И.И., и др. Радиогеномика рака молочной железы — новый вектор междисциплинарной интеграции лучевых и молекулярно-биологических технологий (обзор литературы) // *Медицинский алфавит*. 2020. № 20. С. 21–29. doi: 10.33667/2078-5631-2020-20-21-29
48. Antropova N., Huynh B.Q., Giger M.L. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets // *Med Phys*. 2017. Vol. 44, N 10. P. 5162–5171. doi: 10.1002/mp.12453
49. Antunovic L., Gallivanone F., Sollini M., et al. [¹⁸F]FDG PET/CT features for the molecular characterization of primary breast tumors // *Eur J Nucl Med Mol Imaging*. 2017. Vol. 44, N 12. P. 1945–1954. doi: 10.1007/s00259-017-3770-9
50. Ha S., Park S., Bang J.I., et al. Metabolic radiomics for pretreatment ¹⁸F-FDG PET/CT to characterize locally advanced breast cancer: histopathologic characteristics, response to neoadjuvant chemotherapy, and prognosis // *Sci Rep*. 2017. Vol. 7, N 1. P. 1556. doi: 10.1038/s41598-017-01524-7
51. Guo W., Li H., Zhu Y., et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and

- genomics data // *J Med Imaging* (Bellingham). 2015. Vol. 2, N 4. P. 041007. doi: 10.1117/1.JMI.2.4.041007
52. Saha A., Harowicz M.R., Grimm L.J., et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features // *Br J Cancer*. 2018. Vol. 119, N 4. P. 508–516. doi: 10.1038/s41416-018-0185-8
53. Dong Y., Feng Q., Yang W., et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI // *Eur Radiol*. 2018. Vol. 28, N 2. P. 582–591. doi: 10.1007/s00330-017-5005-7
54. Chan H.M., van der Velden B.H., Loo C.E., Gilhuijs K.G. Eigentumors for prediction of treatment failure in patients with early-stage breast cancer using dynamic contrast-enhanced MRI: a feasibility study // *Phys Med Biol*. 2017. Vol. 62, N 16. P. 6467–6485. doi: 10.1088/1361-6560/aa7dc5
55. Braman N.M., Etesami M., Prasanna P., et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI // *Breast Cancer Res*. 2017. Vol. 19, N 1. P. 57. doi: 10.1186/s13058-017-0846-1
56. Chamming's F., Ueno Y., Ferré R., et al. Features from computerized texture analysis of breast cancers at pretreatment MR imaging are associated with response to neoadjuvant chemotherapy // *Radiology*. 2018. Vol. 286, N 2. P. 412–420. doi: 10.1148/radiol.2017170143
57. Partridge S.C., Zhang Z., Newitt D.C., et al. Diffusion-weighted MRI findings predict pathologic response in neoadjuvant treatment of breast cancer: the ACRIN 6698 multicenter trial // *Radiology*. 2018. Vol. 289, N 3. P. 618–627. doi: 10.1148/radiol.2018180273
58. Park H., Lim Y., Ko E.S., et al. Radiomics signature on magnetic resonance imaging: association with disease-free survival in patients with invasive breast cancer // *Clin Cancer Res*. 2018. Vol. 24, N 19. P. 4705–4714. doi: 10.1158/1078-0432.CCR-17-3783
59. Nasrullah N., Sang J., Alam M.S., et al. Automated lung nodule detection and classification using deep learning combined with multiple strategies // *Sensors* (Basel). 2019. Vol. 19, N 17. P. 3722. doi: 10.3390/s19173722
60. Carter B.W., Godoy M.C., Erasmus J.J. Predicting malignant nodules from screening CTs // *J Thorac Oncol*. 2016. Vol. 11, N 12. P. 2045–2047. doi: 10.1016/j.jtho.2016.09.117
61. Aerts H.J., Velazquez E.R., Leijenaar R.T., et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach // *Nat Commun*. 2014. Vol. 5. P. 4006. doi: 10.1038/ncomms5006
62. Zhou H., Dong D., Chen B., et al. Diagnosis of distant metastasis of lung cancer: based on clinical and radiomic features // *Transl Oncol*. 2018. Vol. 11, N 1. P. 31–36. doi: 10.1016/j.tranon.2017.10.010
63. Liu Y., Kim J., Balagurunathan Y., et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas // *Clin Lung Cancer*. 2016. Vol. 17, N 5. P. 441–448.e6. doi: 10.1016/j.clcc.2016.02.001
64. Fave X., Zhang L., Yang J., et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer // *Sci Rep*. 2017. Vol. 7, N 1. P. 588. doi: 10.1038/s41598-017-00665-z
65. Coroller T.P., Agrawal V., Narayan V., et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer // *Radiother Oncol*. 2016. Vol. 119, N 3. P. 480–486. doi: 10.1016/j.radonc.2016.04.004
66. Kickingereder P., Neuberger U., Bonekamp D., et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma // *Neuro Oncol*. 2018. Vol. 20, N 6. P. 848–857. doi: 10.1093/neuonc/nox188
67. Pérez-Beteta J., Molina-García D., Ortiz-Alhambra J.A., et al. Tumor surface regularity at MR imaging predicts survival and response to surgery in patients with glioblastoma // *Radiology*. 2018. Vol. 288, N 1. P. 218–225. doi: 10.1148/radiol.2018171051
68. Zhou Z., Chen L., Sher D., et al. Predicting lymph node metastasis in head and neck cancer by combining many-objective radiomics and 3-dimensional convolutional neural network through evidential reasoning // *Annu Int Conf IEEE Eng Med Biol Soc*. 2018. Vol. 2018. P. 1–4. doi: 10.1109/EMBC.2018.8513070
69. Wang G., He L., Yuan C., et al. Pretreatment MR imaging radiomics signatures for response prediction to induction chemotherapy in patients with nasopharyngeal carcinoma // *Eur J Radiol*. 2018. Vol. 98. P. 100–106. doi: 10.1016/j.ejrad.2017.11.007
70. Chen X., Oshima K., Schott D., et al. Assessment of treatment response during chemoradiation therapy for pancreatic cancer based on quantitative radiomic analysis of daily CTs: An exploratory study // *PLoS One*. 2017. Vol. 12. P. e0178961. doi: 10.1371/journal.pone.0178961
71. Huang Y.Q., Liang C.H., He L., et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer // *J Clin Oncol*. 2016. Vol. 34, N 18. P. 2157–2164. doi: 10.1200/JCO.2015.65.9128
72. Lin Y.C., Lin G., Hong J.H., et al. Diffusion radiomics analysis of intratumoral heterogeneity in a murine prostate cancer model following radiotherapy: Pixelwise correlation with histology // *J Magn Reson Imaging*. 2017. Vol. 46, N 2. P. 483–489. doi: 10.1002/jmri.25583
73. Chaddad A., Kucharczyk M.J., Niazi T. Multimodal radiomic features for the predicting gleason score of prostate cancer // *Cancers* (Basel). 2018. Vol. 10, N 8. P. 249. doi: 10.3390/cancers10080249
74. Огнерубов Н.А., Шатов И.А., Шатов А.В. Радиогеномика и радиомика в диагностике злокачественных опухолей: обзор литературы // *Вестник Тамбовского университета. Серия: Естественные и технические науки*. 2017. Т. 22, № 6-2. С. 1453–1460. doi: 10.20310/1810-0198-2017-22-6-1453-1460

REFERENCES

- Kumar V, Gu Y, Basu S, et al. Radiomics: The process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234–1248. doi: 10.1016/j.mri.2012.06.010
- Papanikolaou N, Matos C, Koh DM. How to develop a meaningful radiomic signature for clinical use in oncologic patients. *Cancer Imaging*. 2020;20(1):33. doi: 10.1186/s40644-020-00311-4
- Aerts HJ, Grossmann P, Tan Y, et al. Defining a radiomic response phenotype: A pilot study using targeted therapy in NSCLC. *Sci Rep*. 2016;6:33860. doi: 10.1038/srep33860
- Coroller TP, Grossmann P, Hou Y, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol*. 2015;114(3):345–350. doi: 10.1016/j.radonc.2015.02.015

5. Lopez CJ, Nagornaya N, Parra NA, et al. Association of radiomics and metabolic tumor volumes in radiation treatment of glioblastoma multiforme. *Int J Radiat Oncol Biol Phys.* 2017;97(3):586–595. doi: 10.1016/j.ijrobp.2016.11.011
6. De Souza NM, Achten E, Alberich-Bayarri A, et al. Validated imaging biomarkers as decision-making tools in clinical trials and routine practice: current status and recommendations from the EIBALL* subcommittee of the European Society of Radiology (ESR). *Insights Imaging.* 2019;10(1):87. doi: 10.1186/s13244-019-0764-0
7. Jones EF, Buatti JM, Shu HK, et al. Clinical trial design and development work group within the quantitative imaging network. *Tomography.* 2020;6(2):60–64. doi: 10.18383/j.tom.2019.00022
8. European Society of Radiology (ESR). ESR statement on the validation of imaging biomarkers. *Insights Imaging.* 2020;11(1):76. doi: 10.1186/s13244-020-00872-9
9. Grimm LJ, Zhang J, Mazurowski MA. Computational approach to radiogenomics of breast cancer: Luminal A and luminal B molecular subtypes are associated with imaging features on routine breast MRI extracted using computer vision algorithms. *J Magn Reson Imaging.* 2015;42(4):902–907. doi: 10.1002/jmri.24879
10. Nie K, Shi L, Chen Q, et al. Rectal cancer: Assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. *Clin Cancer Res.* 2016;22(21):5256–5264. doi: 10.1158/1078-0432.CCR-15-2997
11. Kim H, Park CM, Lee M, et al. Impact of reconstruction algorithms on ct radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability. *PLoS One.* 2016;11(10):e0164924. doi: 10.1371/journal.pone.0164924
12. Ohri N, Duan F, Snyder BS, et al. Pretreatment 18F-FDG PET textural features in locally advanced non-small cell lung cancer: Secondary analysis of ACRIN 6668/RTOG 0235. *J Nucl Med.* 2016;57(6):842–848. doi: 10.2967/jnumed.115.166934
13. Zhang B, Tian J, Dong D, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clin Cancer Res.* 2017;23(15):4259–4269. doi: 10.1158/1078-0432.CCR-16-2910
14. Nakatsugawa M, Cheng Z, Goatman KA, et al. Radiomic analysis of salivary glands and its role for predicting xerostomia in irradiated head and neck cancer patients. *Int J Radiat Oncol Biol Phys.* 2016;96(2 suppl):S217. doi: 10.1016/j.ijrobp.2016.06.539
15. Shafiee MJ, Chung AG, Khalvati F, et al. Discovery radiomics via evolutionary deep radiomic sequencer discovery for pathologically proven lung cancer detection. *J Med Imaging.* 2016;4(4):041305. doi: 10.1117/1.JMI.4.4.041305
16. Echegaray S, Nair V, Kadoch M, et al. A rapid segmentation-insensitive «Digital Biopsy» method for radiomic feature extraction: method and pilot study using ct images of non-small cell lung cancer. *Tomography.* 2016;2(4):283–294. doi: 10.18383/j.tom.2016.00163
17. Li H, Galperin-Aizenberg M, Pryma D, et al. Unsupervised machine learning of radiomic features for predicting treatment response and overall survival of early stage non-small cell lung cancer patients treated with stereotactic body radiation therapy. *Radiother Oncol.* 2018;129(2):218–226. doi: 10.1016/j.radonc.2018.06.025
18. Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Transactions on Medical Imaging.* 2016;35(5):1299–1312. doi: 10.1109/TMI.2016.2535302
19. Elguindi S, Zelefsky MJ, Jiang J, et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. *Phys Imaging Radiat Oncol.* 2019;12:80–86. doi: 10.1016/j.phro.2019.11.006
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* 2016;278(2):563–577. doi: 10.1148/radiol.2015151169
21. Buckler AJ, Bresolin L, Dunnick NR, et al. Quantitative imaging test approval and biomarker qualification: Interrelated but distinct activities. *Radiology.* 2011;259(3):875–884. doi: 10.1148/radiol.10100800
22. Alobaidli S, McQuaid S, South C, et al. The role of texture analysis in imaging as an outcome predictor and potential tool in radiotherapy treatment planning. *Br J Radiol.* 2014;87(1042):20140369. doi: 10.1259/bjr.20140369
23. Li H, Giger ML, Lan L, et al. Comparative analysis of image-based phenotypes of mammographic density and parenchymal patterns in distinguishing between BRCA1/2 cases, unilateral cancer cases, and controls. *J Med Imaging.* 2014;1(3):031009. doi: 10.1117/1.JMI.1.3.031009
24. Goh V, Ganeshan B, Nathan P, et al. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. *Radiology.* 2011;261(1):165–171. doi: 10.1148/radiol.11110264
25. Yip C, Davnall F, Kozarski R, et al. Assessment of changes in tumor heterogeneity following neoadjuvant chemotherapy in primary esophageal cancer. *Dis Esophagus.* 2015;28(2):172–179. doi: 10.1111/dote.12170
26. Park JE, Kim HS. Radiomics as a quantitative imaging biomarker: practical considerations and the current standpoint in neuro-oncologic studies. *Nucl Med Mol Imaging.* 2018;52(2):99–108. doi: 10.1007/s13139-017-0512-7
27. Sergunova KA, Akhmad ES, Semenov DS, et al. Medical physicist's participation in quality assurance and safety in magnetic resonance imaging. *Medical physics.* 2020;(3):78–85. (In Russ).
28. Clements JB, Baird CT, de Boer SF, et al. AAPM medical physics practice guideline 10.a.: Scope of practice for clinical medical physics. *J Appl Clin Med Phys.* 2018;19(6):11–25. doi: 10.1002/acm2.12469
29. Shukla-Dave A, Obuchowski NA, Chenevert TL, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging.* 2019;49(7):e101–e121. doi: 10.1002/jmri.26518
30. Russek SE, Boss M, Jackson EF, et al. Characterization of NIST/ISMRM MRI System Phantom. *Proc Intl Soc Mag Reson Med.* 2012;20:2456.
31. Kuo MD, Jamshidi N. Behind the numbers: Decoding molecular phenotypes with radiogenomics –guiding principles and technical considerations. *Radiology.* 2014;270(2):320–325. doi: 10.1148/radiol.13132195
32. Narang S, Lehrer M, Yang D, et al. Radiomics in glioblastoma: current status, challenges and potential opportunities. *Transl Cancer Res.* 2016;5(4):383–397. doi: 10.21037/tcr.2016.06.31
33. O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol.* 2017;14(3):169–186. doi: 10.1038/nrclinonc.2016.162
34. Buizza G, Toma-Dasu I, Lazzeroni M, et al. Early tumor response prediction for lung cancer patients using novel longitudinal

- pattern features from sequential PET/CT image scans. *Phys Med*. 2018;54:21–29. doi: 10.1016/j.ejmp.2018.09.003
35. Raunig DL, McShane LM, Pennello G, et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. *Stat Methods Med Res*. 2015;24(1):27–67. doi: 10.1177/0962280214537344
36. Obuchowski NA, Reeves AP, Huang EP, et al. Quantitative imaging biomarkers: a review of statistical methods for computer algorithm comparisons. *Stat Methods Med Res*. 2015;24(1):68–106. doi: 10.1177/0962280214537390
37. Elguindi S, Zelefsky MJ, Jiang J, et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. *Phys Imaging Radiat Oncol*. 2019;12:80–86. doi: 10.1016/j.phro.2019.11.006
38. Morozov SP, Vladimirov AV, Klyashorny VG, et al. Clinical trials of software based on intelligent technologies (radiation diagnostics). Methodological recommendations. Moscow; 2019. 33 p. (In Russ).
39. Sullivan DC, Obuchowski NA, Kessler LG, et al. Metrology standards for quantitative imaging biomarkers. *Radiology*. 2015;277(3):813–825. doi: 10.1148/radiol.2015142202
40. Shur J, Blackledge M, D'Arcy J, et al. MRI texture feature repeatability and image acquisition factor robustness, a phantom study and in silico study. *Eur Radiol Exp*. 2021;5(1):2. doi: 10.1186/s41747-020-00199-6
41. Bane O, Hectors SJ, Wagner M, et al. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. *Magn Reson Med*. 2018;79(5):2564–2575. doi: 10.1002/mrm.26903
42. He Y, Liu Y, Dyer BA, et al. 3D-printed breast phantom for multi-purpose and multi-modality imaging. *Quant Imaging Med Surg*. 2019;9(1):63–74. doi: 10.21037/qims.2019.01.05
43. Scheuermann JS, Reddin JS, Opanowski A, et al. Qualification of national cancer institute-designated cancer centers for quantitative PET/CT imaging in clinical trials. *J Nucl Med*. 2017;58(7):1065–1071. doi: 10.2967/jnumed.116.186759
44. Obuchowski NA, Barnhart HX, Buckler AJ, et al. Statistical issues in the comparison of quantitative imaging biomarker algorithms using pulmonary nodule volume as an example. *Stat Methods Med Res*. 2015;24(1):107–140. doi: 10.1177/0962280214537392
45. Kessler LG, Barnhart HX, Buckler AJ, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res*. 2015;24(1):9–26. doi: 10.1177/0962280214537333
46. Napel S, Mu W, Jardim-Perassi BV, et al. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. *Cancer*. 2018;124(24):4633–4649. doi: 10.1002/cnrc.31630
47. Rozhkova NI, Bozhenko VK, Burdina II, et al. Radiogenomics of breast cancer as new vector of interdisciplinary integration of radiation and molecular biological technologies (literature review). *Medical alphabet*. 2020;(20):21–29. (In Russ). doi: 10.33667/2078-5631-2020-20-21-29
48. Antropova N, Huynh BQ, Giger ML. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. *Med Phys*. 2017;44(10):5162–5171. doi: 10.1002/mp.12453
49. Antunovic L, Gallivanone F, Sollini M, et al. [¹⁸F]FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur J Nucl Med Mol Imaging*. 2017;44(12):1945–1954. doi: 10.1007/s00259-017-3770-9
50. Ha S, Park S, Bang JI, et al. Metabolic radiomics for pretreatment ¹⁸F-FDG PET/CT to characterize locally advanced breast cancer: histopathologic characteristics, response to neoadjuvant chemotherapy, and prognosis. *Sci Rep*. 2017;7(1):1556. doi: 10.1038/s41598-017-01524-7
51. Guo W, Li H, Zhu Y, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. *J Med Imaging (Bellingham)*. 2015;2(4):041007. doi: 10.1117/1.JMI.2.4.041007
52. Saha A, Harowicz MR, Grimm LJ, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer*. 2018;119(4):508–516. doi: 10.1038/s41416-018-0185-8
53. Dong Y, Feng Q, Yang W, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur Radiol*. 2018;28(2):582–591. doi: 10.1007/s00330-017-5005-7
54. Chan HM, van der Velden BH, Loo CE, Gilhuijs KG. Eigentumors for prediction of treatment failure in patients with early-stage breast cancer using dynamic contrast-enhanced MRI: a feasibility study. *Phys Med Biol*. 2017;62(16):6467–6485. doi: 10.1088/1361-6560/aa7dc5
55. Braman NM, Etesami M, Prasanna P, et al. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res*. 2017;19(1):57. doi: 10.1186/s13058-017-0846-1
56. Chamming's F, Ueno Y, Ferré R, et al. Features from computerized texture analysis of breast cancers at pretreatment MR imaging are associated with response to neoadjuvant chemotherapy. *Radiology*. 2018;286(2):412–420. doi: 10.1148/radiol.2017170143
57. Partridge SC, Zhang Z, Newitt DC, et al. Diffusion-weighted MRI findings predict pathologic response in neoadjuvant treatment of breast cancer: the ACRIN 6698 multicenter trial. *Radiology*. 2018;289(3):618–627. doi: 10.1148/radiol.2018180273
58. Park H, Lim Y, Ko ES, et al. Radiomics signature on magnetic resonance imaging: association with disease-free survival in patients with invasive breast cancer. *Clin Cancer Res*. 2018;24(19):4705–4714. doi: 10.1158/1078-0432.CCR-17-3783
59. Nasrullah N, Sang J, Alam MS, et al. Automated lung nodule detection and classification using deep learning combined with multiple strategies. *Sensors (Basel)*. 2019;19(17):3722. doi: 10.3390/s19173722
60. Carter BW, Godoy MC, Erasmus JJ. Predicting malignant nodules from screening CTs. *J Thorac Oncol*. 2016;11(12):2045–2047. doi: 10.1016/j.jtho.2016.09.117
61. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. doi: 10.1038/ncomms5006
62. Zhou H, Dong D, Chen B, et al. Diagnosis of distant metastasis of lung cancer: based on clinical and radiomic features. *Transl Oncol*. 2018;11(1):31–36. doi: 10.1016/j.tranon.2017.10.010
63. Liu Y, Kim J, Balagurunathan Y, et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. *Clin Lung Cancer*. 2016;17(5):441–448.e6. doi: 10.1016/j.clcc.2016.02.001

64. Fave X, Zhang L, Yang J, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci Rep*. 2017;7(1):588. doi: 10.1038/s41598-017-00665-z
65. Coroller TP, Agrawal V, Narayan V, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. *Radiother Oncol*. 2016;119(3):480–486. doi: 10.1016/j.radonc.2016.04.004
66. Kickingereder P, Neuberger U, Bonekamp D, et al. Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. *Neuro Oncol*. 2018;20(6):848–857. doi: 10.1093/neuonc/nox188
67. Pérez-Beteta J, Molina-García D, Ortiz-Alhambra JA, et al. Tumor surface regularity at MR imaging predicts survival and response to surgery in patients with glioblastoma. *Radiology*. 2018;288(1):218–225. doi: 10.1148/radiol.2018171051
68. Zhou Z, Chen L, Sher D, et al. Predicting lymph node metastasis in head and neck cancer by combining many-objective radiomics and 3-dimensional convolutional neural network through evidential reasoning. *Annu Int Conf IEEE Eng Med Biol Soc*. 2018;2018:1–4. doi: 10.1109/EMBC.2018.8513070
69. Wang G, He L, Yuan C, et al. Pretreatment MR imaging radiomics signatures for response prediction to induction chemotherapy in patients with nasopharyngeal carcinoma. *Eur J Radiol*. 2018;98:100–106. doi: 10.1016/j.ejrad.2017.11.007
70. Chen X, Oshima K, Schott D, et al. Assessment of treatment response during chemoradiation therapy for pancreatic cancer based on quantitative radiomic analysis of daily CTs: An exploratory study. *PLoS One*. 2017;12:e0178961. doi: 10.1371/journal.pone.0178961
71. Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol*. 2016;34(18):2157–2164. doi: 10.1200/JCO.2015.65.9128
72. Lin YC, Lin G, Hong JH, et al. Diffusion radiomics analysis of intratumoral heterogeneity in a murine prostate cancer model following radiotherapy: Pixelwise correlation with histology. *J Magn Reson Imaging*. 2017;46(2):483–489. doi: 10.1002/jmri.25583
73. Chaddad A, Kucharczyk MJ, Niazi T. Multimodal radiomic features for the predicting gleason score of prostate cancer. *Cancers (Basel)*. 2018;10(8):249. doi: 10.3390/cancers10080249
74. Ognerubov NA, Shatov IA, Shatov AV. Radiogenomics and radiomics in the diagnostics of malignant tumours: a literary review. *Tambov University Reports. Series: Natural and Technical Sciences*. 2017;22(6):1453–1460. (In Russ). doi: 10.20310/1810-0198-2017-22-6-1453-1460

ОБ АВТОРАХ

*** Хоружая Анна Николаевна;**

адрес: Россия, 109029, Москва, Средняя Калитниковская ул., д. 28, стр. 1; ORCID: <https://orcid.com/0000-0003-4857-5404>; eLibrary SPIN: 7948-6427; e-mail: a.khoruzhaya@npcmr.ru

Ахмад Екатерина Сергеевна;

ORCID: <https://orcid.com/0000-0002-8235-9361>;
eLibrary SPIN: 5891-4384

Семенов Дмитрий Сергеевич;

ORCID: <https://orcid.com/0000-0002-4293-2514>;
eLibrary SPIN: 2278-7290

AUTHORS' INFO

*** Anna N. Khoruzhaya, MD;**

address: 28-1 Srednyaya Kalitnikovskaya str., 109029 Moscow, Russia; ORCID: <https://orcid.com/0000-0003-4857-5404>; eLibrary SPIN: 7948-6427; e-mail: a.khoruzhaya@npcmr.ru

Ekaterina S. Akhmad;

ORCID: <https://orcid.com/0000-0002-8235-9361>;
eLibrary SPIN: 5891-4384

Dmitry S. Semenov;

ORCID: <https://orcid.com/0000-0002-4293-2514>;
eLibrary SPIN: 2278-7290

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD70479>

Роль маммографии в радиомике рака молочной железы

В.Г. Говорухина¹, С.С. Семенов^{2, 3}, П.Б. Гележе², В.В. Диденко^{2, 4},
С.П. Морозов², А.Е. Андрейченко²

¹ Московский государственный университет имени М.В. Ломоносова, Москва, Российская Федерация

² Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Российская Федерация

³ Московский клинический научно-практический центр имени А.С. Логинова, Москва, Российская Федерация

⁴ Городская клиническая онкологическая больница № 1, Москва, Российская Федерация

АННОТАЦИЯ

Маммография — в настоящее время единственный способ скрининга рака молочной железы (РМЖ). Хотя цифровая маммография служит основным и наиболее широкодоступным методом для выявления РМЖ, её эффективность в обнаружении и оценке внутриопухолевой гетерогенности опухоли ограничена. Пункционная биопсия не может отразить гистологической картины опухоли в целом из-за небольшого размера образца ткани или опухоли. По этой причине выбор подходящего лечения и определение прогноза становится затруднительным. В этом случае такой неинвазивный подход, как медицинская визуализация, даёт более полное представление об опухоли, перспективен при «виртуальной биопсии», а также в контроле прогрессирования заболевания и ответа на терапию.

Радиомика с помощью текстурного анализа позволяет взглянуть на снимок как на группу числовых характеристик, выйти за пределы привычного качественного зрительного восприятия интенсивностей и перейти к более глубокому анализу цифровых, пиксельных данных с целью повышения точности дифференциальной диагностики. Метод радиогеномики, являясь естественным продолжением радиомики, фокусируется на определении экспрессии генов исходя из лучевого фенотипа опухоли. В обзоре рассматриваются возможности применения маммографии в радиомике и радиогеномике РМЖ.

В статье представлен обзор литературы баз данных PubMed, Medline, Springer, eLibrary, а также найденных с помощью Google Scholar актуальных российских научных статей. Полученная релевантная информация объединена, структурирована и проанализирована с целью изучения роли маммографии в радиомике РМЖ.

Ключевые слова: рак молочной железы; маммография; радиомика; радиогеномика; искусственный интеллект.

Как цитировать

Говорухина В.Г., Семенов С.С., Гележе П.Б., Диденко В.В., Морозов С.П., Андрейченко А.Е. Роль маммографии в радиомике рака молочной железы // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 185–199. DOI: <https://doi.org/10.17816/DD70479>

DOI: <https://doi.org/10.17816/DD70479>

The role of mammography in breast cancer radiomics

Veronika G. Govorukhina¹, Serafim S. Semenov^{2, 3}, Pavel B. Gelezhe², Vera V. Didenko^{2, 4}, Sergey P. Morozov², Anna E. Andreychenko²

¹ Lomonosov Moscow State University, Moscow, Russian Federation

² Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of Moscow Health Care, Moscow, Russian Federation

³ Moscow Clinical Scientific Center named after A.S. Loginov, Moscow, Russian Federation

⁴ City Clinical Oncology Hospital № 1, Moscow, Russian Federation

ABSTRACT

Mammography is still the only screening method for breast cancer. Although digital mammography is the most common and widely used method for detecting breast cancer, it is ineffective at detecting and assessing intratumoral heterogeneity. Due to the small size of the tissue sample or tumor, biopsies often fail to represent the entire tumor. For this reason, selecting a treatment and determining a patient's prognosis becomes difficult. In this case, medical imaging is a noninvasive approach that can provide a more comprehensive view of the entire tumor, act as a "virtual biopsy," and be useful for monitoring disease progression and response to therapy.

Radiomics with texture analysis allows you to look at an image as a group of numerical data, moving beyond the usual visual perception and into a deeper analysis of digital, pixel data to improve the accuracy of differential diagnosis. Radiogenomics is a natural extension of radiomics that focuses on determining gene expression based on radiologic tumor phenotype. The purpose of this review is to evaluate the role of mammography in breast cancer radiomics and radiogenomics.

The article presents a literature review of relevant Russian scientific articles found in databases such as PubMed, Medline, Springer, eLibrary, and Google Scholar. The information obtained was then pooled, structured, and analyzed to examine the role of mammography in breast cancer screening radiomics.

Keywords: breast cancer; mammography; radiomics; radiogenomics; artificial intelligence.

To cite this article

Govorukhina VG, Semenov SS, Gelezhe PB, Didenko VV, Morozov SP, Andreychenko AE. The role of mammography in breast cancer radiomics. *Digital Diagnostics*. 2021;2(2):185–199. DOI: <https://doi.org/10.17816/DD70479>

Received: 11.05.2021

Accepted: 04.07.2021

Published: 12.07.2021

DOI: <https://doi.org/10.17816/DD70479>

乳房造影检查在乳腺癌放射学中的作用

Veronika G. Govorukhina¹, Serafim S. Semenov^{2, 3}, Pavel B. Gelezhe², Vera V. Didenko^{2, 4}, Sergey P. Morozov², Anna E. Andreychenko²

¹ Lomonosov Moscow State University, Moscow, Russian Federation

² Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of Moscow Health Care, Moscow, Russian Federation

³ Moscow Clinical Scientific Center named after A.S. Loginov, Moscow, Russian Federation

⁴ City Clinical Oncology Hospital № 1, Moscow, Russian Federation

简评

乳房造影检查是目前筛查乳腺癌的唯一方法。尽管数字乳房x线照相术是检测乳腺癌的主要和最广泛可用的方法，但其在检测和评估肿瘤的肿瘤内异质性方面的有效性有限。由于组织样本或肿瘤体积小，穿刺活检不能反映整个肿瘤的组织学图片。由于这个原因，选择适当的治疗方法和确定预后变得复杂。在这种情况下，医学成像这样的非侵入性方法给出了肿瘤的更完整的画面，是有希望的»虚拟活检«，以及用于监测疾病的进展和对治疗的反应。

使用纹理分析的放射学允许您将图像视为一组数值特征，超越通常的定性视觉感知强度，并继续深入分析数字，像素数据，以提高差分诊断的准确性。放射基因组学方法是放射组学的自然延伸，侧重于根据肿瘤的放射表型确定基因的表达。该综述探讨了在乳腺癌的放射组学和放射基因组学中使用乳房造影照相术的可能性。

本文概述了PubMed, Medline, Springer, eLibrary数据库的文献，以及使用Google学术搜索找到的相关俄罗斯科学文章。将获得的相关信息组合，结构与分析，以研究乳房造影照相术在乳腺癌放射组学中的作用。

关键词：乳腺癌;乳房x光检查;放射学;放射基因组学;人工智能。

引用本文：

Govorukhina VG, Semenov SS, Gelezhe PB, Didenko VV, Morozov SP, Andreychenko AE. 乳房造影检查在乳腺癌放射学中的作用. *Digital Diagnostics*. 2021;2(2):185–199. DOI: <https://doi.org/10.17816/DD70479>

收到: 11.05.2021

接受: 04.07.2021

发布日期: 12.07.2021

РАК МОЛОЧНОЙ ЖЕЛЕЗЫ: АКТУАЛЬНОСТЬ, ХАРАКТЕРИСТИКА

Рак молочной железы (РМЖ) — актуальная проблема современной онкологии, поскольку по частоте развития занимает первое место среди всех злокачественных новообразований у женщин [1]. За 2018 год в России заболеваемость РМЖ составила 89,8 случаев на 100 000 женского населения [1], за 2019 год выявлено 73 366 случаев РМЖ, из них доля пациентов с III и IV стадией составила 27,7% [2].

РМЖ — гетерогенное заболевание, что подразумевает различие в типах морфологии опухоли и экспрессионных подтипах в зависимости от рецепторного статуса РМЖ [3, 4]. Рецепторный статус РМЖ включает экспрессию рецептора эстрогена (estrogen receptors, ER), рецептора прогестерона (progesterone receptors, PR) и рецептора эпидермального фактора роста, тип 2 (human epidermal growth factor receptor 2; HER2). Иммунохимическое окрашивание на маркер пролиферации Ki67 и рецептор эпидермального фактора роста (epidermal growth factor receptor, EGFR) также используются при определении молекулярного подтипа РМЖ [4].

Выделяют 5 молекулярных подтипов РМЖ:

- 1) люминальный A [ER+, PR+ высокие ($\geq 20\%$), HER2-, Ki-67 низкий ($\leq 20\%$): эстрогензависимые малоагрессивные опухоли; нет избытка экспрессии рецепторов белка HER2; отличается высокой экспрессией гена ER;
- 2) люминальный B [ER+, PR+ низкий ($\leq 20\%$), HER2-, Ki-67 высокий]: эстрогензависимые опухоли; нет избытка экспрессии рецепторов белка HER2;
- 3) люминальный B [ER+, HER2+, уровень Ki-67 любой, PR любые]: эстрогензависимые агрессивные опухоли; выражена амплификация онкогена HER2; выражена экспрессия гена ER;
- 4) HER2 положительный [ER- и PR-, Ki-67 любой, HER2+]: эстрогеннезависимые агрессивные опухоли; выражена амплификация онкогена HER2;
- 5) тройной негативный (базальноподобный): эстрогеннезависимые агрессивные опухоли с наихудшими показателями выживаемости (ER-, PR-, HER2-) [3–5].

Известно, что биология опухоли влияет на подбор терапии, а также на прогноз исхода: пациентки с ER+ и PR+ имеют более длительную безрецидивную выживаемость, в то время как тройной негативный РМЖ (ТНРМЖ) (ER-, PR-, HER2-) имеет наиболее агрессивное течение и наихудшие показатели выживаемости [3, 6]. Использование биологических маркеров для выявления подтипов РМЖ увеличивает выживаемость пациенток благодаря более точной диагностике заболевания. Например, пациентки с экспрессией ER и PR в опухоли должны получать эндокринную терапию; при экспрессии HER2 — анти-HER2-терапию [7].

Внутриопухолевая гетерогенность — неоднородность морфологического строения и варибельность экспрессии различного рода маркеров отдельными группами клеток в пределах одной опухоли [8, 9]. Морфологическую внутриопухолевую гетерогенность можно описать как различие в разных областях опухоли, т.е. пространственную гетерогенность, или как прогрессирование опухоли во времени — гетерогенность во времени [8]. Из-за такой неоднородности новообразований, а также из-за небольшого размера образца пункционной ткани биопсия не может отражать гистологическую картину опухоли в целом. По этой причине выбор подходящего лечения и определение прогноза становится затруднительным. Проведение биопсии также бывает затруднено при небольшом размере опухоли. В этом случае такой неинвазивный подход, как медицинская визуализация, даёт более полное представление об опухоли, перспективен при «виртуальной биопсии», а также в контроле прогрессирования заболевания и ответа на терапию [7, 10–12].

РАННЯЯ ДИАГНОСТИКА РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ, ПРОГНОЗИРОВАНИЕ ИСХОДА ТЕРАПИИ

Эффективный метод снижения смертности пациенток — выявление рака на ранних стадиях [13]. Маммография представляет собой пока единственный способ скрининга и диагностики РМЖ [10]. Хотя цифровая маммография служит основным методом для раннего выявления РМЖ, её эффективность в обнаружении находок ограничена: наблюдается более низкая чувствительность маммографии у пациенток с высокой плотностью молочных желёз (типы ACR-C и D по классификации американской коллегии радиологов (American College of Radiology (ACR)) [14], поскольку существует вероятность перекрытия на изображении патологического образования фиброгланулярными структурами [15, 16]. Несмотря на сниженную чувствительность у одной из групп пациенток, цифровая маммография на данный момент обладает наилучшим сочетанием чувствительности и специфичности при диагностике РМЖ, однако эти два показателя разнятся в пределах 75–90 и 80–90% соответственно в зависимости от страны [15]. В недавнем исследовании O. Demircioglu и соавт. [17] было показано, что интерпретация изображений низкого качества врачами-рентгенологами с небольшим опытом приводит к гипердиагностике и ненужным болезненным инвазивным процедурам примерно в половине клинических случаев [6, 15, 17].

Последние достижения в области технологий искусственного интеллекта (ИИ), применяемых для анализа изображений, перспективны для обнаружения опухолей и снижения нагрузки на врачей, оценки лечения

и контроля прогрессирования заболевания [6]. Однако при РМЖ раннее определение прогноза заболевания и прогнозирование ответа на терапию — центральные задачи клинической практики и исследований. С этой точки зрения открыты другие возможности ИИ с применением текстурного анализа для определения подтипа рака и прогнозирования ответа на лечение [6, 18, 19].

Анализ маммографических исследований включает в себя интерпретацию изображений врачом-рентгенологом с оценкой структуры опухоли, отношения её к окружающим тканям, особенностей расположения и строения микрокальцинатов. Для создания действительно персонализированной терапии также требуется количественная (объективная) оценка образования [6].

Внутриопухолевая гетерогенность имеет важное значение для точной диагностики, клинических прогнозов (ответ на лечение, выживаемость, прогрессирование заболевания и др.) и лечения онкологических заболеваний [20, 21]. Раннее обнаружение устойчивости опухоли к терапии имеет решающее значение для улучшения исходов, что позволяет вовремя изменить схему лечения [6].

Таким образом, существует необходимость в повышении эффективности выявления, прогнозирования исхода и ответа на лечение РМЖ. Особый набор методов, объединяемых в радиомик и радиогеномик, на данный момент набирает силу в качестве инструмента для максимизации информации, которая может быть извлечена практически из любой модальности цифровых медицинских изображений [15].

РАДИОМИКА И РАДИОГЕНОМИКА

Лучевые диагностические изображения содержат информацию, отражающую патофизиологические процессы, и эта взаимосвязь может быть выявлена с помощью количественного анализа изображений [22]. Иными словами, характеристики опухоли на клеточном и генетическом уровнях отражаются в фенотипических паттернах опухоли, которые могут проявляться и быть обнаружены на изображениях [23].

Радиомика — процесс, состоящий из этапов подготовки и последующего количественного анализа многомерных данных, полученных из цифровых медицинских изображений (суффикс «мик» присутствует в названиях направлений молекулярной биологии, специализирующихся на больших объёмах данных [24]). Радиомика описывается как анализ изображений с использованием определённых алгоритмов, направленных на извлечение числовых характеристик изображений, с целью создания моделей классификации для усовершенствования поддержки принятия медицинских решений, а также для определения прогноза заболевания [25, 26] и лечения [27], что особенно ценно для персонализированной терапии. В радиомике одна область интереса изображения используется

для получения множества (иногда десятков и сотен) числовых характеристик, каждое из которых может нести определённый теоретико-информационный аспект (часто называемый «радиомическим признаком»), который недоступен при обычном просмотре изображений [15]. Радиомика использует алгоритмы автоматического извлечения признаков текстуры цифровых медицинских изображений для преобразования данных медицинской визуализации в набор данных порядковой статистики [28]. Иными словами, радиомика с помощью текстурного анализа позволяет взглянуть на снимок как на группу числовых характеристик, выйти за пределы привычного зрительного восприятия и перейти к анализу многомерных данных.

Радиогеномика — технология, устанавливающая связь между генотипом пациента и фенотипом визуализации. Необходимо отметить, что термин «радиогеномика» может также употребляться по отношению к генетической изменчивости и её связи с ответом на лучевую терапию [29, 30], но чаще используется при оценке взаимосвязи между характеристиками изображения опухоли или какого-то заболевания с его образцами экспрессии генов и мутациями генов [25, 26].

Метод радиогеномики даёт возможность определить экспрессию генов в опухоли исходя из её лучевого фенотипа. Это важно, поскольку опухоли гетерогенны, и радиомические данные извлекаются из области интереса (опухоли) целиком, а не из отдельного образца [22]. Радиогеномика также позволяет оценить ответ на лечение, основываясь не только на традиционном измерении размера опухоли в динамике [25]. Сочетание радиомики и радиогеномики приведёт к обнаружению генных нарушений по изображениям [6]. Радиомика и радиогеномика повышают точность клинического диагноза и имеют прогностическое значение за счёт выявления взаимосвязей между различными типами клинических данных [22].

ЭТАПЫ РАДИОМИКИ

Рассматривая радиомик как процесс, можно выделить несколько основных этапов: получение изображений; выделение области интереса; извлечение радиомических признаков из области интереса (текстурный анализ изображений); анализ текстурных признаков и построение различных моделей прогнозирования и классификации с использованием полученных радиомических данных с вариантом включения дополнительной информации (например, клинических, демографических или геномных данных; наличия коморбидных состояний) [22, 23, 31]. Этапы радиомики продемонстрированы на рис. 1, а более подробная их характеристика описана ниже.

1. Определение клинической проблемы и **получение цифровых медицинских изображений**, исключая исследования низкого качества.



Рис. 1. На схеме продемонстрированы типичные этапы радиомики. После получения медицинских изображений (1) вручную или автоматическим путём проводится их сегментация (2). Радиомические признаки первого и более высоких порядков извлекаются из сегментированных областей интереса с использованием специального программного обеспечения или модулей языков программирования (3). Далее выполняются анализ и отбор наиболее значимых из полученных текстурных признаков. На последнем этапе на основе проанализированных радиомических данных проводится построение различных клинко-диагностических моделей классификации или прогнозирования (4).

2. Сегментация изображений до основных исследуемых зон интереса [32], например до злокачественного новообразования для оценки его внутриопухолевой гетерогенности. Многие опухоли имеют нечёткие границы, что делает воспроизводимость их сегментации сложной задачей [33]. Предпочтительно использование полу-или полностью автоматического выделения области интереса при помощи специального программного обеспечения, однако в некоторых случаях требуется уточнение экспертов и выделение вручную [23, 34]. Процесс выбора области интереса не стандартизован, и область интереса может содержать как всю опухоль, так и некоторые её части [35, 36]. Определение области интереса вручную трудоёмко и вариативно по причине различия интерпретации изображений различными врачами-рентгенологами [33], что в итоге влияет на точность построенных радиомических моделей, однако современные технологии глубокого обучения с использованием больших данных способны минимизировать это влияние [37].

3. Извлечение множества радиомических признаков из сегментированной области интереса с помощью математических операций с цифровыми значениями интенсивностей и взаиморасположения пикселей или вокселей на изображениях. Извлечённые количественные признаки подразделяются на категории: морфологические (объём и форма), гистограммные признаки (описание интенсивности уровней серого тона) первого, второго и более высоких порядков [26, 34].

Морфологические признаки отражают форму области интереса. Для планарных изображений актуальны 2D-признаки формы, например отношение периметра к поверхности и округлость — мера приближения формы области интереса к форме круга. Так, например, отношение поверхности к объёму у опухоли звёздчатой формы будет больше, чем у округлой [31].

Гистограммные признаки первого порядка [38] отражают распределение интенсивностей уровня серого для пикселей области интереса. Наиболее часто встречающиеся признаки этой категории — среднее и медиана — указывают на ширину диапазона интенсивностей; энтропия — мера неравномерности распределения интенсивностей (более высокие значения соответствуют более гетерогенной области) [39]. Однако статистические признаки первого порядка не отражают пространственного расположения пикселей.

Гистограммные признаки второго порядка [38], иначе называемые текстурными признаками, отражают пространственное отношение между двумя соседними пикселями с одинаковым или же различным значением яркости. Наиболее часто применяемый метод для извлечения текстурных признаков основан на использовании матрицы совместной встречаемости уровней серого тона, иначе называемой матрицей совпадения уровня серого (gray level co-occurrence matrix, GLCM) — это матрица, число строк и столбцов которой отражают значения уровня яркости серого; в ячейках матрицы отражается число раз, когда соответствующие значения серого находятся в определённой взаимосвязи (угол и расстояние между анализируемыми пикселями). Например, признаки, получаемые при использовании такой матрицы, включают энтропию второго порядка, указывающую на гетерогенность; энергию, описывающую гомогенность изображения; контрастность, определяющую локальное изменение интенсивностей [10]. В радиомике анализ текстуры предоставляет информацию о мере внутриопухолевой гетерогенности [22, 40].

На рис. 2 приведено сравнение гистограммных признаков первого и второго порядков, продемонстрировано формирование матрицы смежности уровня серого тона: на рис. 2 (а) показаны два исходных изображения,

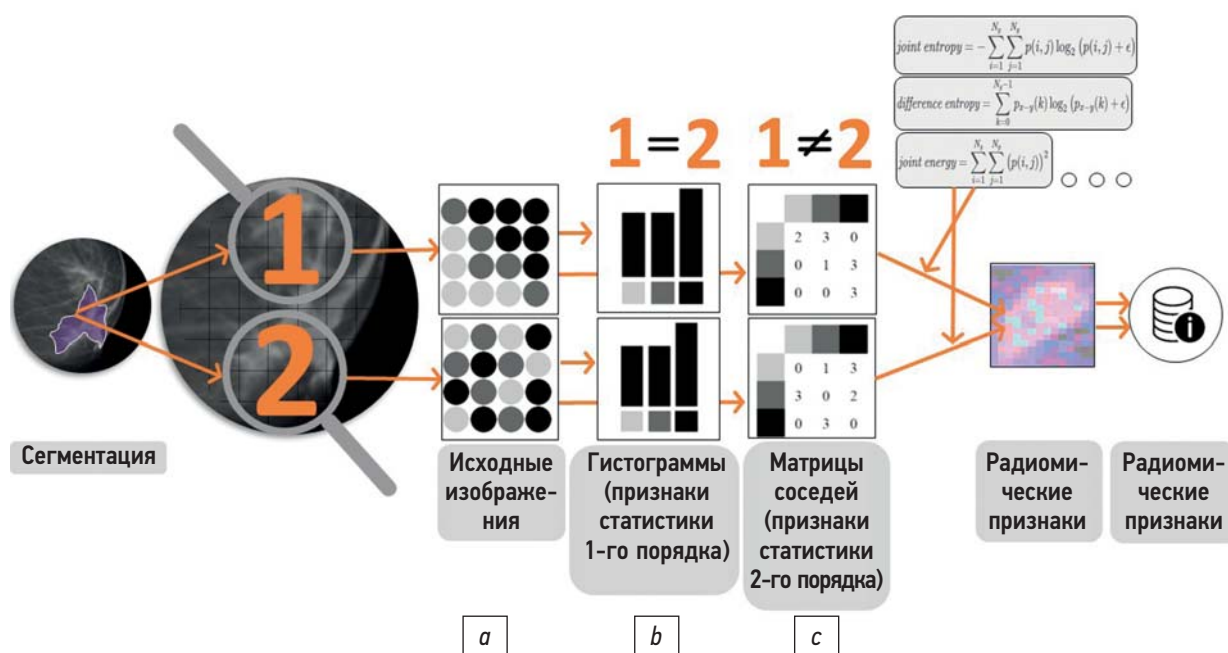


Рис. 2. Сравнение гистограммных признаков первого и второго порядка. Две различные исходные области интереса сегментированного изображения (a) содержат одинаковое количество пикселей светло-серого, тёмно-серого и чёрного оттенков. Гистограммы яркости на основе количества пикселей определённых оттенков (гистограммные признаки первого порядка) одинаковы (b). Эти признаки не отражают взаиморасположения пикселей. Матрицы смежности (гистограммные признаки второго порядка) отражают гетерогенность изображений (c). В дальнейшем математическими алгоритмами из полученных гистограмм интенсивностей и матриц смежностей уровня серого рассчитывается множество радиомических признаков для анализа и моделирования.

на 2 (b) — гистограммы первого порядка; на рис. 2 (c) матрицы смежности уровней серого тона получены для исходных изображений. В заголовках строк и столбцов этих матриц указаны номера оттенка серого цвета. В каждую ячейку таблицы помещено количество горизонтальных пар пикселей, в которых пиксели с оттенком, указанным в заголовке строки и столбца этой ячейки, расположены относительно друг друга под углом 0° [41]. В дальнейшем математическими алгоритмами из полученных гистограмм интенсивностей и матриц смежностей уровня серого рассчитывается множество радиомических признаков для анализа и моделирования.

4. Анализ и моделирование: полученные радиомические признаки в зависимости от поставленного вопроса можно анализировать различными способами — от статистических моделей до методов машинного обучения.

Учитывая большое количество данных, извлекаемых из изображений, первым этапом проводится отбор или сокращение признаков. Следует исключить невоспроизводимые признаки, поскольку они, скорее всего, приведут к ложным результатам построенных моделей [42, 43].

Второй этап — многомерный анализ данных [31] и построение моделей трёх основных групп — предиктивных, объяснительных и описательных [15]. Описательные модели используются для получения общей картины каждого признака, резюмируя основные его

характеристики. Так, часто используемые для биомедицинских данных объяснительные методы сосредоточены на способности модели устанавливать связь между признаком и исходом, например, взаимосвязь текстурных характеристик матрицы совпадений уровня серого и морфологического типа РМЖ внутри области интереса. Формирование предиктивных моделей происходит с помощью методов машинного обучения: такие модели анализируют вероятность определённых исходов на основе полученных входных данных [15], например, радиомическая модель предсказания отсутствия ответа на неоадьювантную химиотерапию РМЖ. Качество и воспроизводимость результатов работы полученных моделей должны быть оценены перед применением в клинических условиях [31].

ОПЫТ, ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ МАММОГРАФИИ В РАДИОМИКЕ И РАДИОГЕНОМИКЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ

Распознавание злокачественного новообразования

Сложный и самый важный шаг в маммографии — это разделение находок на маммограммах

на доброкачественные и злокачественные [44]. В недавнем исследовании N. Мао и соавт. [45] было продемонстрировано, что использование количественных признаков совместно с ИИ может обеспечить большую диагностическую эффективность при использовании маммографии по сравнению с эффективностью выполнения диагностики опытными врачами-рентгенологами [15].

Процесс причисления микрокальцинатов к доброкачественному либо злокачественному процессу по изображениям всё ещё остаётся сложной задачей для врачей-рентгенологов [46]. На данный момент есть возможность проведения текстурного анализа изображений совместно с методами ИИ при обнаружении подозрительных кальцинатов с потенциальным уменьшением количества ненужных биопсий [47, 48].

Особенности паренхимы молочной железы могут отражать биологические факторы риска развития РМЖ. В работе H. Li и соавт. [49] было показано, что использование текстурных признаков, извлечённых из маммограмм поражённой и контралатеральной (с нормальной паренхимой) желёз, позволяет повысить точность цифровой маммографии в диагностике РМЖ. Исследования показывают, что радиомика с высокой чувствительностью и специфичностью позволяет классифицировать злокачественные новообразования молочной железы от доброкачественных [50].

Определение подтипов РМЖ

Исследования в области радиогеномики последних лет подтвердили корреляцию МР-признаков визуализации РМЖ с молекулярными подтипами — люминальным А, люминальным В, HER2 и ТНРМЖ [51]. Хотя из маммографических изображений можно получить меньше информации, чем при магнитно-резонансной томографии (МРТ), на данный момент есть несколько исследований, в которых продемонстрированы возможности маммографии в радиомике и радиогеномике РМЖ. В исследовании W. Ма и соавт. [10] была показана возможность прогноза молекулярного подтипа РМЖ путём извлечения радиомических характеристик из маммографических изображений. Наиболее значимыми оказались следующие признаки: округлость, вогнутость, среднее значение уровня серого и корреляция. Результаты показали, что люминальные и ТНРМЖ обладают особыми текстурными признаками, в отличие от других подтипов, что позволяет количественно отличать их с помощью радиомики.

У некоторых пациенток с РМЖ использование неоадъювантной химиотерапии не даёт эффективного терапевтического ответа, что ведёт к задержке хирургического вмешательства, плохому прогнозу и увеличению затрат на лечение. Было показано, что использование радиомики совместно с независимыми клиническими факторами риска (например Ki-67 индекса, HER2

статуса) улучшает модель предсказания отсутствия ответа на неоадъювантную химиотерапию [52].

Раннее обнаружение с помощью медицинской визуализации более агрессивного подтипа РМЖ, а именно ТНРМЖ, позволит врачам назначать лечение до окончательного подтверждения биопсией [53]. В исследовании H.X. Zhang и соавт. [53] ТНРМЖ по сравнению с другими подтипами имел большую округлость и вогнутость; площадь под ROC-кривой (Receiver Operating Curve; классическая ROC-кривая: график зависимости чувствительности от специфичности [54]) использовалась для оценки точности этих двух признаков в дифференциации ТНРМЖ от других подтипов РМЖ и составила более 0,70 [53, 55]. В этом исследовании коэффициент асимметрии (гистограммный признак, отражающий асимметрию распределения значений относительно среднего) всех подтипов был менее 0 (отрицательная или левосторонняя асимметрия). Было показано, что значение коэффициента асимметрии ТНРМЖ ниже коэффициентов остальных исследуемых подтипов. Следовательно, вышеуказанные радиомические признаки можно рассматривать как потенциальные маркеры отличия ТНРМЖ от других подтипов РМЖ в будущем [53].

Прогнозирование развития РМЖ и возможность персонализированного скрининга

Технологии на основе радиомики помогут в развитии персонализированного скрининга, т.е. составления индивидуального плана скрининга для пациенток путём создания инструментов индивидуальной оценки риска и включения их в средства поддержки принятия решений при маммографическом скрининге, а также индивидуальных интервалов скрининга [56–58]. Известно, что более высокая плотность молочных желёз связана с большим риском развития РМЖ [59]. Термин «плотность» означает степень ослабления рентгеновского излучения при прохождении через железу и отражает распределение фиброгладулярной ткани. Однако определение только лишь плотности не отражает всей сложности структуры желёз. Текстурные признаки, получаемые из изображений, были предложены в качестве маркеров изменений в паренхиме, которые указывают на связь с развитием РМЖ [57, 59]. В исследовании D. Kontos и соавт. [59] (2019) были выделены радиомические фенотипы, отражающие сложность паренхимы на маммограммах (помимо плотности) и имеющие независимую связь с РМЖ. В отличие от обычного определения плотности, текстурные признаки отразили едва различимую и более локализованную сложность паренхиматозного рисунка. Плотность молочных желёз была различна среди фенотипов низкой и средней сложности паренхимы, но схожа для других фенотипов. Интерес представляют полученные данные о фенотипе с наименьшей сложностью (сложность паренхимы) у женщин

с высокой плотностью молочных желёз из-за большей их гомогенности, тогда как фенотип с низкой и средней степенью сложности паренхимы включал в себя незначительное количество изображений с высокой плотностью [59].

Предоперационное определение метастазов в подмышечных лимфатических узлах

Наиболее частая локализация метастазов РМЖ — подмышечные лимфатические узлы. Статус подмышечных лимфатических узлов — важный фактор для оценки общей и безрецидивной выживаемости пациенток с РМЖ [60]. Точное предоперационное определение статуса подмышечных лимфоузлов может предоставить врачам информацию для решения вопроса о необходимости проведения лимфодиссекции и назначении адъювантной терапии. В настоящее время статус определяется при помощи биопсии сторожевого лимфатического узла, что может привести к осложнениям, в частности повреждению кровеносных сосудов, нервов и развитию лимфедемы; а диагностика с помощью визуализационных методов имеет низкую чувствительность [60]. В работе J. Yang и соавт. [60] была разработана модель с включением извлечённых из маммограмм радиомических признаков, которая при сочетании с дополнительной клинико-патологической информацией способна служить неинвазивным методом для предоперационного определения метастазов в подмышечных лимфатических узлах.

ОГРАНИЧЕНИЯ ПРИМЕНЕНИЯ РАДИОМИКИ

Хотя радиомика и радиогеномика обладают большим потенциалом для развития персонализированной медицины, для этих технологий необходима валидация на независимом наборе данных для подтверждения их диагностической и прогностической ценности. Необходимо время для того, чтобы эти технологии приобрели существенное практическое значение в онкологических исследованиях, и ещё больше времени потребуется до начала применения их в клинической практике. Такие ограничения связаны с тем, что имеющиеся большие объёмы данных в настоящий момент не содержат полной характеристики пациентов [6]. Сложность воспроизводимости результатов радиомики связана с недостатками на каждом её этапе: на разном оборудовании и протоколах визуализации получаются разные текстурные признаки [61, 62]; золотой стандарт сегментации опухоли вручную требует много времени и зависит от оператора [63]; полу- и автоматическая сегментация, снижающие вариабельность [64, 65], не стандартизированы; между текстурными признаками существует очевидная повторяемость, что приводит к необходимости уменьшения размеров данных [66, 67];

нет чёткого объяснения, какова связь между единичей радиомики (базовым блоком текстуры) и тканями человека. В обстоятельствах, когда основная теория не ясна, а технические методы не стандартизированы, любые полученные «значимые» результаты исследований должны быть пересмотрены [68].

ПЕРСПЕКТИВА ИСПОЛЬЗОВАНИЯ МАГНИТНО-РЕЗОНАНСНОЙ ТОМОГРАФИИ В РАДИОМИКЕ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ

Накоплены убедительные данные о том, что МРТ молочных желёз превосходит в диагностической точности традиционные методы диагностики, такие как маммография [69]. Помимо выявления большего количества протокового рака *in situ*, МРТ молочных желёз часто меняет стадию онкологического процесса, что способствует оптимизации лечебного процесса.

Установлено, что радиомические признаки, извлечённые из МР-изображений молочных желёз, отражают гетерогенность опухоли, её васкуляризацию [70], позволяют дифференцировать протоковый рак от доброкачественного очага [71]. Существующие радиомические модели пока отстают от маммологов экспертного уровня по показателю площади под кривой (AUC) для дифференциальной диагностики доброкачественных и злокачественных очагов [72]. Однако показаны многообещающие результаты в выявлении подозрительных (BI-RADS 4 и 5) очагов с помощью радиомики диффузионно-взвешенных изображений [73].

Радиомика, по-видимому, может помочь в принятии клинических решений, потенциально избегая инвазивных вмешательств в подмышечной области. Два различных исследования показали, что радиомическая модель способна предсказывать метастазы сторожевых лимфатических узлов [74, 75], что имеет колоссальное клиническое значение.

Другое применение радиомики связано с индексом пролиферации Ki-67, который используется в качестве прогностического маркера при РМЖ [72]. Последние исследования изучали возможность прогнозирования экспрессии маркера пролиферации Ki-67 с помощью радиомики серий динамического контрастного усиления [76–79].

ЗАКЛЮЧЕНИЕ

Одна из ключевых концепций радиомики состоит в том, что лучевые диагностические изображения содержат данные, которые могут дать больше информации об области интереса, чем считалось ранее. Маммография — наиболее эффективный метод для раннего обнаружения РМЖ. Маммографические изображения

применимы для радиомического анализа, что может быть использовано для распознавания злокачественного новообразования, определения подтипов РМЖ, прогнозирования развития заболевания и ответа на лечение.

Технологии на основе радиомики, а именно в области маммографии, в будущем могут быть включены в средства поддержки принятия врачебных решений для определения стратегии индивидуального скрининга, наблюдения и, возможно, превентивной терапии. Однако необходимо учитывать, что радиомика находится на ранней стадии развития, и предстоит большой объём исследований перед началом клинического применения.

Маммография с применением радиомики позволяет получать важную диагностическую и прогностическую информацию о РМЖ, что обещает привести к снижению потребности в инвазивных и зачастую трудных для выполнения процедурах. Хотя методы на основе радиомики не заменят биопсию в ближайшем будущем, ожидая подкрепления новыми исследованиями, интеграция радиомики и радиогеномики в клиническую практику станет одной из важных и перспективных задач для достижения цели снижения смертности от РМЖ.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении поисково-аналитической работы и подготовке статьи.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. В.Г. Говорухина, С.С. Семенов — поиск публикаций по теме, анализ литературы, написание текста; П.Б. Гележе, В.В. Диденко — поиск публикаций по теме, написание текста, экспертная оценка обзора литературы, С.П. Морозов, А.Е. Андрейченко — определение основной направленности обзора, экспертная оценка обзора литературы, систематизация и финальное редактирование обзора. Авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Funding source. The authors declare that there is no external funding for the exploration and analysis work.

Competing interests. The authors declare no obvious and potential conflicts of interest related to the publication of this article.

Authors' contribution. V.G. Govorukhina, S.S. Semenov — search for publications on the topic, analysis of literature, writing a text; P.B. Gelezhe, V.V. Didenko — search for publications on the topic, text writing, expert assessment of literature, S.P. Morozov, A.E. Andreychenko — defining the main focus of the review, expert assessment of the literature review, systematization and editing of the review. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

СПИСОК ЛИТЕРАТУРЫ

- Каприн, А.Д., Старинский, В.В., Петрова Г.В. Злокачественные новообразования в России в 2018 году (заболеваемость и смертность). Москва: МНИОИ им. П.А. Герцена, 2019. 250 с.
- Каприн А.Д., Старинский В.В. Состояние онкологической помощи населению России в 2019 году. Москва: МНИОИ им. П.А. Герцена, 2020. 239 с.
- Гришина К.А., Музаффарова Т.А., Хайленко В.А., Карпунин А.В. Молекулярно-генетические маркеры рака молочной железы // Опухоли женской репродуктивной системы. 2016. Vol. 12, N 3. P. 36–42. doi: 10.17650/1994-4098-2016-12-3-36-42
- Wu M., Ma J. Association between imaging characteristics and different molecular subtypes of breast cancer // Acad Radiol. 2017. Vol. 24, N 4. P. 426–434. doi: 10.1016/j.acra.2016.11.012
- Стенина М.Б., Жукова Л.Г., Королева И.А., и соавт. Практические рекомендации по лекарственному лечению рака молочной железы // Malig tumours. 2021. Vol. 10. N 3.
- Pesapane F., Suter M.B., Rotili A., et al. Will traditional biopsy be substituted by radiomics and liquid biopsy for breast cancer diagnosis and characterisation? // Med Oncol. 2020. Vol. 37, N 4. P. 29. doi: 10.1007/s12032-020-01353-1
- Januškevičienė I., Petrikaitė V. Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations // Life Sci. 2019. Vol. 239. P. 117009. doi: 10.1016/j.lfs.2019.117009
- Turashvili G., Brogi E. Tumor heterogeneity in breast cancer // Front Med. 2017. Vol. 4. P. 227. doi: 10.3389/fmed.2017.00227
- Завьялова М.В., Вторушин С.В., Цыганов М.М. Внутриопухолевая гетерогенность: природа и биологическое значение. Обзор // Биохимия. 2013. Т. 78, № 11. С. 1531–1549.
- Ma W., Zhao Y., Ji Y., et al. Breast cancer molecular subtype prediction by mammographic radiomic features // Acad Radiol. 2019. Vol. 26, N 2. P. 196–201. doi: 10.1016/j.acra.2018.01.023
- Aerts H.J., Velazquez E.R., Leijenaar R.T., et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach // Nat Commun. 2014. Vol. 5. P. 4006. doi: 10.1038/ncomms5006
- Lin F., Wang Z., Zhang K., et al. Contrast-Enhanced spectral mammography-based radiomics nomogram for identifying benign and malignant breast lesions of Sub-1 cm // Front Oncol. 2020. Vol. 10. P. 573630. doi: 10.3389/fonc.2020.573630
- Li X., Qin G., He Q., et al. Digital breast tomosynthesis versus digital mammography: integration of image modalities enhances deep learning-based breast mass classification // Eur Radiol. 2020. Vol. 30, N 2. P. 778–788. doi: 10.1007/s00330-019-06457-5
- Nazari S.S., Mukherjee P. An overview of mammographic density and its association with breast cancer // Breast Cancer. 2018. Vol. 25, N 3. P. 259–267. doi: 10.1007/s12282-018-0857-5

15. Conti A., Duggento A., Indovina I., et al. Radiomics in breast cancer classification and prediction // *Semin Cancer Biol.* 2021. Vol. 72. P. 238–250.
16. Hogg P., Kelly J., Mercer C. Digital mammography: A holistic approach // Springer. 2015. 309 p.
17. Demircioglu O., Uluer M., Aribal E. How many of the biopsy decisions taken at inexperienced breast radiology units were correct? // *J Breast Heal.* 2017. Vol. 13, N 1. P. 23–26. doi: 10.5152/tjbh.2016.2962
18. Valdora F., Houssami N., Rossi F., et al. Rapid review: radiomics and breast cancer // *Breast Cancer Res Treat.* 2018. Vol. 169, N 2. P. 217–229. doi: 10.1007/s10549-018-4675-4
19. Mayerhoefer M.E., Materka A., Langset G., et al. Introduction to radiomics // *J Nucl Med.* 2020. Vol. 61, N 4. P. 488–495. doi: 10.2967/jnumed.118.222893
20. Lambin P., Zindler J., Vanneste B.G., et al. Decision support systems for personalized and participative radiation oncology // *Adv Drug Deliv Rev.* 2017. Vol. 109. P. 131–153. doi: 10.1016/j.addr.2016.01.006
21. Wen Y.L., Leech M. Review of the role of radiomics in tumour risk classification and prognosis of cancer // *Anticancer Research.* 2020. Vol. 40, N 7. P. 3605–3618. doi: 10.21873/anticancer.14350
22. Gillies R.J., Kinahan P.E., Hricak H. Radiomics: Images are more than pictures, they are data // *Radiology.* 2016. Vol. 278, N 2. P. 563–577. doi: 10.1148/radiol.2015151169
23. Lambin P., Leijenaar R.T., Deist T.M., et al. Radiomics: The bridge between medical imaging and personalized medicine // *Nat Rev Clin Oncol.* 2017. Vol. 14, N 12. P. 749–762. doi: 10.1038/nrclinonc.2017.141
24. Vailati-Riboni M., Palombo V., Looor J.J. What are omics sciences? // *Periparturient Diseases of Dairy Cows: A Systems Biology Approach.* 2017. P. 1–7. doi: 10.1007/978-3-319-43033-1_1
25. Porcu M., Solinas C., Mannelli L., et al. Radiomics and “radi-omics” in cancer immunotherapy: a guide for clinicians // *Crit Rev Oncol Hematol.* 2020. Vol. 154. P. 103068. doi: 10.1016/j.critrevonc.2020.103068
26. Ognerubov N.A., Shatov I.A., Shatov A.V. Radiogenomics and radiomics in the diagnostics of malignant tumours: a literary review // *Tambov Univ Reports Ser Nat Tech Sci.* 2017. Vol. 22, N 6–2. P. 1453–1460. doi: 10.20310/1810-0198-2017-22-6-1453-1460
27. Nasief H., Hall W., Zhenget C., al. Improving treatment response prediction for chemoradiation therapy of pancreatic cancer using a combination of delta-radiomics and the clinical biomarker CA19-9 // *Front Oncol.* 2020. Vol. 9. P. 1464. doi: 10.3389/fonc.2019.01464
28. Lambin P., Rios-Velazquez E., Leijenaar R., et al. Radiomics: Extracting more information from medical images using advanced feature analysis // *Eur J Cancer.* 2012. Vol. 48, N 4. P. 441–446. doi: 10.1016/j.ejca.2011.11.036
29. Kerns S.L., Ostrer H., Rosenstein B.S. Radiogenomics: Using genetics to identify cancer patients at risk for development of adverse effects following radiotherapy // *Cancer Discovery.* 2014. Vol. 4, N 2. P. 155–165. doi: 10.1158/2159-8290
30. Neri E., Del Re M., Paia F., et al. Radiomics and liquid biopsy in oncology: the holons of systems medicine // *Insights Imaging.* 2018. Vol. 9, N 6. P. 915–924. doi: 10.1007/s13244-018-0657-7
31. Rizzo S., Botta F., Raimondi S., et al. Radiomics: the facts and the challenges of image analysis // *Eur Radiol Exp.* 2018. Vol. 2, N 1. P. 36. doi: 10.1186/s41747-018-0068-z
32. Kumar V., Gu Y., Basu S., et al. Radiomics: The process and the challenges // *Magn Reson Imaging.* 2012. Vol. 30, N 9. P. 1234–1248. doi: 10.1016/j.mri.2012.06.010
33. Sala E., Mema E., Himoto Y., et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging // *Clinical Radiology.* 2017. Vol. 72, N 1. P. 3–10. doi: 10.1016/j.crad.2016.09.013
34. Lee S.H., Park H., Ko E.S. Radiomics in breast imaging from techniques to clinical applications: A review // *Korean J Radiol.* 2020. Vol. 21, N 7. P. 779–792. doi: 10.3348/kjr.2019.0855
35. Tagliafico A.S., Piana M., Schenone D., et al. Overview of radiomics in breast cancer diagnosis and prognostication // *Breast.* 2020. Vol. 49. P. 74–80. doi: 10.1016/j.breast.2019.10.018
36. Li H., Giger M.L., Huo Z., et al. Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: effect of ROI size and location // *Med Phys.* 2004. Vol. 31, N 3. P. 549–555. doi: 10.1118/1.1644514
37. Holbrook M.D., Blocker S.J., Mowery Y.M., et al. Mri-based deep learning segmentation and radiomics of sarcoma in mice // *Tomography.* 2020. Vol. 6, N 1. P. 23–33. doi: 10.18383/j.tom.2019.00021
38. Прэтт У. Цифровая обработка изображений. Москва : Мир, 1982. 480 с.
39. Santos J.M., Oliveira B.C., de Araujo-Filho J., et al. State-of-the-art in radiomics of hepatocellular carcinoma: a review of basic principles, applications, and limitations // *Abdom Radiol.* 2020. Vol. 45, N 2. P. 342–353. doi: 10.1007/s00261-019-02299-3
40. Avanzo M., Stancanella J., El Naqa I. Beyond imaging: the promise of radiomics // *Phys Medica.* 2017. Vol. 38. P. 122–139. doi: 10.1016/j.ejmp.2017.05.071
41. Verma M., Raman B., Murala S. Local extrema co-occurrence pattern for color and texture image retrieval // *Neurocomputing.* 2015. Vol. 165. P. 255–269. doi: 10.1016/j.neucom.2015.03.015
42. Tunali I., Hall L.O., Napel S., et al. Stability and reproducibility of computed tomography radiomic features extracted from peritumoral regions of lung cancer lesions // *Med Phys.* 2019. Vol. 46, N 11. P. 5075–5085. doi: 10.1002/mp.13808
43. Nasief H., Zheng C., Schott D., et al. A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer // *NPJ Precis Oncol.* 2019. Vol. 3, N 1. P. 25. doi: 10.1038/s41698-019-0096-z
44. Cui Y., Li Y., Xing D., et al. Improving the prediction of benign or malignant breast masses using a combination of image biomarkers and clinical parameters // *Front Oncol.* 2021. Vol. 11. P. 629321. doi: 10.3389/fonc.2021.629321
45. Mao N., Yin P., Wang Q., et al. Added value of radiomics on mammography for breast cancer diagnosis: a feasibility study // *J Am Coll Radiol.* 2019. Vol. 16, N 4, Pt A. P. 485–491. doi: 10.1016/j.jacr.2018.09.041
46. Fanizzi A., Basile T.M., Losurdo L., et al. A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis // *BMC Bioinformatics.* 2020. Vol. 21, Suppl 2. P. 91. doi: 10.1186/s12859-020-3358-4
47. Stelzer P.D., Steding O., Raudner M.W., et al. Combined texture analysis and machine learning in suspicious calcifications detected by mammography: Potential to avoid unnecessary stereotactical biopsies // *Eur J Radiol.* 2020. Vol. 132. P. 109309. doi: 10.1016/j.ejrad.2020.109309
48. Karahaliou A., Skiadopoulos S., Boniatis I., et al. Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis // *Br J Radiol.* 2007. Vol. 80, N 956. P. 648–656. doi: 10.1259/bjr/30415751

49. Li H., Mendel K.R., Lan L., et al. Digital mammography in breast cancer: Additive value of radiomics of breast parenchyma // *Radiology*. 2019. Vol. 291, N 1. P. 15–20. doi: 10.1148/radiol.2019181113
50. Parekh V.S., Jacobs M.A. MPRAD: A multiparametric radiomics framework // *arXiv*. 2018.
51. Rozhkova N.I., Bozhenko V.K., Burdina I.I., et al. Radiogenomics of breast cancer as new vector of interdisciplinary integration of radiation and molecular biological technologies (literature review) // *Med Alph*. 2020. N 20. P. 21–29. doi: 10.33667/2078-5631-2020-20-21-29
52. Wang Z., Lin F., Ma H., et al. Contrast-Enhanced spectral mammography-based radiomics nomogram for the prediction of neo-adjuvant chemotherapy-insensitive breast cancers // *Front Oncol*. 2021. Vol. 11. P. 605230. doi: 10.3389/fonc.2021.605230
53. Zhang H.X., Sun Z.Q., Cheng Y.G., et al. A pilot study of radiomics technology based on X-ray mammography in patients with triple-negative breast cancer // *J Xray Sci Technol*. 2019. Vol. 27, N 3. P. 485–492. doi: 10.3233/XST-180488
54. Morozov S.P., Vladimirovsky A.V., Klyashtorny V.G., et al. Clinical acceptance of software based on artificial intelligence technologies (Radiology). Moscow: Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies; 2019. 45 p.
55. Mandrekar J.N. Receiver operating characteristic curve in diagnostic test assessment // *J Thorac Oncol*. 2010. Vol. 5, N 9. P. 1315–1316. doi: 10.1097/JTO.0b013e3181ec173d
56. Rahbar H., McDonald E.S., Lee J.M., et al. How can advanced imaging be used to mitigate potential breast cancer overdiagnosis? // *Academic Radiology*. 2016. Vol. 23, N 6. P. 768–773. doi: 10.1016/j.acra.2016.02.008
57. Pinker K. Beyond breast density: Radiomic phenotypes enhance assessment of breast cancer risk // *Radiology*. 2019. Vol. 290, N 1. P. 50–51. doi: 10.1148/radiol.2018182296
58. Sun W., Tseng T.L., Qian W., et al. Using multiscale texture and density features for near-term breast cancer risk analysis // *Med Phys*. 2015. Vol. 42, N 6. P. 2853–2862. doi: 10.1118/1.4919772
59. Kontos D., Winham S.J., Oustimov A., et al. Radiomic phenotypes of mammographic parenchymal complexity: Toward augmenting breast density in breast cancer risk assessment // *Radiology*. 2019. Vol. 290, N 1. P. 41–49. doi: 10.1148/radiol.2018180179
60. Yang J., Wang T., Yang L., et al. Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method // *Sci Rep*. 2019. Vol. 9, N 1. P. 4429. doi: 10.1038/s41598-019-40831-z
61. Zhao B., Tan Y., Tsai W.Y., et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging // *Sci Rep*. 2016. Vol. 6. P. 23428. doi: 10.1038/srep23428
62. Lu L., Ehmke R.C., Schwartz L.H., Zhao B. Assessing agreement between radiomic features computed for multiple CT imaging settings // *PLoS One*. 2016. Vol. 11, N 12. P. e0166550. doi: 10.1371/journal.pone.0166550
63. Velazquez E.R., Parmar C., Jermoumi M., et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer // *Sci Rep*. 2013. Vol. 3. P. 3529. doi: 10.1038/srep03529
64. Qiu Q., Duan J., Gong G., et al. Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma // *Transl Cancer Res*. 2017. Vol. 6, N 5. doi: 10.21037/tcr.2017.09.47
65. Qiu Q., Duan J., Duan Z., et al. Reproducibility and non-redundancy of radiomic features extracted from arterial phase CT scans in hepatocellular carcinoma patients: Impact of tumor segmentation variability // *Quant Imaging Med Surg*. 2019. Vol. 9, N 3. P. 453–464. doi: 10.21037/qims.2019.03.02
66. Hunter L.A., Krafft S., Stingo F., et al. High quality machine-robust image features: Identification in nonsmall cell lung cancer computed tomography images // *Med Phys*. 2013. Vol. 40, N 12. P. 121916. doi: 10.1118/1.4829514
67. O'Connor J.P., Aboagye E.O., Adams J.E., et al. Imaging biomarker roadmap for cancer studies // *Nat Rev Clin Oncol*. 2017. Vol. 14, N 3. P. 169–186. doi: 10.1038/nrclinonc.2016.162
68. Chalkidou A., O'Doherty M.J., Marsden P.K. False discovery rates in PET and CT studies with texture features: A systematic review // *PLoS One*. 2015. Vol. 10, N 5. P. e0124165. doi: 10.1371/journal.pone.0124165
69. Mann R.M., Kuhl C.K., Moy L. Contrast-enhanced MRI for breast cancer screening // *J Magn Reson Imaging*. 2019. Vol. 50, N 2. P. 377–390. doi: 10.1002/jmri.26654
70. Parekh V.S., Jacobs M.A. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI // *NPJ Breast Cancer*. 2017. Vol. 3. P. 43. doi: 10.1038/s41523-017-0045-3
71. Whitney H.M., Taylor N.S., Drukker K., et al. Additive benefit of radiomics over size alone in the distinction between benign lesions and luminal a cancers on a large clinical breast MRI dataset // *Acad Radiol*. 2019. Vol. 26, N 2. P. 202–209. doi: 10.1016/j.acra.2018.04.019
72. Crivelli P., Ledda R.E., Parascandolo N., et al. A new challenge for radiologists: radiomics in breast cancer // *BioMed Research International*. 2018. Vol. 2018. P. 6120703. doi: 10.1155/2018/6120703
73. Bickelhaupt S., Paech D., Kickingereder P., et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography // *J Magn Reson Imaging*. 2017. Vol. 46, N 2. P. 604–616. doi: 10.1002/jmri.25606
74. Han L., Zhu Y., Liu Z., et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer // *Eur Radiol*. 2019. Vol. 29, N 7. P. 3820–3829. doi: 10.1007/s00330-018-5981-2
75. Dong Y., Feng Q., Yang W., et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI // *Eur Radiol*. 2018. Vol. 28, N 2. P. 582–591. doi: 10.1007/s00330-017-5005-7
76. Ma W., Ji Y., Qi L., et al. Breast cancer Ki67 expression prediction by DCE-MRI radiomics features // *Clin Radiol*. 2018. Vol. 73, N 10. P. 909.e1–909.e5. doi: 10.1016/j.crad.2018.05.027
77. Jagadish K., Sheela G.M., Naidu B.P., et al. Big data analytics and radiomics to discover diagnostics and therapeutics for gastric cancer // *Recent Advancements in Biomarkers and Early Detection of Gastrointestinal Cancers*. 2020. P. 213–219. doi: 10.1007/978-981-15-4431-6_12
78. Scheckenbach K. Radiomics: Big data instead of biopsies in the future? // *Laryngorhinootologie*. 2018. Vol. 97, S 01. P. S114–S141. doi: 10.1055/s-0043-121964
79. European Society of Radiology (ESR). Medical imaging in personalised medicine: a white paper of the research committee of the European Society of Radiology (ESR) // *Insights Imaging*. 2015. Vol. 6, N 2. P. 141–155. doi: 10.1007/s13244-015-0394-0

REFERENCES

1. Kaprin AD, Starinsky VV, Petrova GV. Malignant neoplasms in Russia in 2018 (morbidity and mortality). Moscow : P.A. Herzen Moscow State Research Institute; 2019. 250 p. (In Russ).
2. Kaprin AD, Starinsky VV. The state of oncological care for the population of Russia in 2019. Moscow : P.A. Herzen Moscow State Research Institute; 2020. 239 p. (In Russ).
3. Grishina KA, Muzaffarova TA, Khailenko VA, Karpukhin AV. Molecular genetic markers of breast cancer. *Tumors of the female reproductive system*. 2016;12(3):36–42. (In Russ). doi: 10.17650/1994-4098-2016-12-3-36-42
4. Wu M, Ma J. Association between imaging characteristics and different molecular subtypes of breast cancer. *Acad Radiol*. 2017;24(4):426–434. doi: 10.1016/j.acra.2016.11.012
5. Stenina MB, Zhukova LG, Koroleva IA, et al. Practical recommendations for the drug treatment of breast cancer. *Malign tumours*. 2021;10(3). (In Russ).
6. Pesapane F, Suter MB, Rotili A, et al. Will traditional biopsy be substituted by radiomics and liquid biopsy for breast cancer diagnosis and characterisation? *Med Oncol*. 2020;37(4):29. doi: 10.1007/s12032-020-01353-1
7. Januškevičienė I, Petrikaitė V. Heterogeneity of breast cancer: the importance of interaction between different tumor cell populations. *Life Sci*. 2019;(239):117009. doi: 10.1016/j.lfs.2019.117009
8. Turashvili G, Brogi E. Tumor heterogeneity in breast cancer. *Front Med*. 2017;4:227. doi: 10.3389/fmed.2017.00227
9. Zavyalova MV, Vtorushin SV, Tsyganov MM. Intra-tumor heterogeneity: nature and biological significance. Review. *Biochemistry*. 2013;78(11):1531–1549. (In Russ).
10. Ma W, Zhao Y, Ji Y, et al. Breast cancer molecular subtype prediction by mammographic radiomic features. *Acad Radiol*. 2019;26(2):196–201. doi: 10.1016/j.acra.2018.01.023
11. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. doi: 10.1038/ncomms5006
12. Lin F, Wang Z, Zhang K, et al. Contrast-Enhanced spectral mammography-based radiomics nomogram for identifying benign and malignant breast lesions of Sub-1 cm. *Front Oncol*. 2020;10:573630. doi: 10.3389/fonc.2020.573630
13. Li X, Qin G, He Q, et al. Digital breast tomosynthesis versus digital mammography: integration of image modalities enhances deep learning-based breast mass classification. *Eur Radiol*. 2020;30(2):778–788. doi: 10.1007/s00330-019-06457-5
14. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. *Breast Cancer*. 2018;25(3):259–267. doi: 10.1007/s12282-018-0857-5
15. Conti A, Duggento A, Indovina I, et al. Radiomics in breast cancer classification and prediction. *Semin Cancer Biol*. 2021;72:238–250. doi: 10.1016/j.semcancer.2020.04.002
16. Hogg P, Kelly J, Mercer C. Digital mammography: A holistic approach. Springer; 2015. 309 p.
17. Demircioglu O, Uluer M, Aribal E. How many of the biopsy decisions taken at inexperienced breast radiology units were correct? *J Breast Heal*. 2017;13(1):23–26. doi: 10.5152/tjbh.2016.2962
18. Valdora F, Houssami N, Rossi F, et al. Rapid review: radiomics and breast cancer. *Breast Cancer Res Treat*. 2018;169(2):217–229. doi: 10.1007/s10549-018-4675-4
19. Mayerhoefer ME, Materka A, Langset G, et al. Introduction to radiomics. *J Nucl Med*. 2020;61(4):488–495. doi: 10.2967/jnumed.118.222893
20. Lambin P, Zindler J, Vanneste BG, et al. Decision support systems for personalized and participative radiation oncology. *Adv Drug Deliv Rev*. 2017;109:131–153. doi: 10.1016/j.addr.2016.01.006
21. Wen YL, Leech M. Review of the role of radiomics in tumour risk classification and prognosis of cancer. *Anticancer Research*. 2020;40(7):3605–3618. doi: 10.21873/anticancer.14350
22. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology*. 2016;278(2):563–577. doi: 10.1148/radiol.2015151169
23. Lambin P, Leijenaar RT, Deist TM, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749–762. doi: 10.1038/nrclinonc.2017.141
24. Vailati-Riboni M, Palombo V, Looor JJ. What are omics sciences? Periparturient diseases of dairy cows: a systems biology approach. 2017. P. 1–7. doi: 10.1007/978-3-319-43033-1_1
25. Porcu M, Solinas C, Mannelli L, et al. Radiomics and “radi-omics” in cancer immunotherapy: a guide for clinicians. *Crit Rev Oncol Hematol*. 2020;154:103068. doi: 10.1016/j.critrevonc.2020.103068
26. Ognerubov NA, Shatov IA, Shatov AV. Radiogenomics and radiomics in the diagnostics of malignant tumours: a literary review. *Tambov Univ Reports Ser Nat Tech Sci*. 2017;22(6-2):1453–1460. doi: 10.20310/1810-0198-2017-22-6-1453-1460
27. Nasief H, Hall W, Zhenget C, et al. Improving treatment response prediction for chemoradiation therapy of pancreatic cancer using a combination of delta-radiomics and the clinical biomarker CA19-9. *Front Oncol*. 2020;9:1464. doi: 10.3389/fonc.2019.01464
28. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441–446. doi: 10.1016/j.ejca.2011.11.036
29. Kerns SL, Ostrer H, Rosenstein BS. Radiogenomics: Using genetics to identify cancer patients at risk for development of adverse effects following radiotherapy. *Cancer Discovery*. 2014;4(2):155–165. doi: 10.1158/2159-8290
30. Neri E, Del Re M, Paiar F, et al. Radiomics and liquid biopsy in oncology: the holons of systems medicine. *Insights Imaging*. 2018;9(6):915–924. doi: 10.1007/s13244-018-0657-7
31. Rizzo S, Botta F, Raimondi S, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018;2(1):36. doi: 10.1186/s41747-018-0068-z
32. Kumar V, Gu Y, Basu S, et al. Radiomics: The process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234–1248. doi: 10.1016/j.mri.2012.06.010
33. Sala E, Mema E, Himoto Y, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clinical Radiology*. 2017;72(1):3–10. doi: 10.1016/j.crad.2016.09.013
34. Lee SH, Park H, Ko ES. Radiomics in breast imaging from techniques to clinical applications: A review. *Korean J Radiol*. 2020;21(7):779–792. doi: 10.3348/kjr.2019.0855
35. Tagliafico AS, Piana M, Schenone D, et al. Overview of radiomics in breast cancer diagnosis and prognostication. *Breast*. 2020;(49):74–80. doi: 10.1016/j.breast.2019.10.018

36. Li H, Giger ML, Huo Z, et al. Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: Effect of ROI size and location. *Med Phys*. 2004;31(3):549–555. doi: 10.1118/1.1644514
37. Holbrook MD, Blocker SJ, Mowery YM, et al. Mri-based deep learning segmentation and radiomics of sarcoma in mice. *Tomography*. 2020;6(1):23–33. doi: 10.18383/j.tom.2019.00021
38. Pratt W. Digital Image Processing. Moscow : World; 1982. 480 p.
39. Santos JM, Oliveira BC, de Araujo-Filho J, et al. State-of-the-art in radiomics of hepatocellular carcinoma: a review of basic principles, applications, and limitations. *Abdom Radiol*. 2020;45(2):342–353. doi: 10.1007/s00261-019-02299-3
40. Avanzo M, Stancanella J, El Naqa I. Beyond imaging: The promise of radiomics. *Phys Medica*. 2017;(38):122–139. doi: 10.1016/j.ejmp.2017.05.071
41. Verma M, Raman B, Murala S. Local extrema co-occurrence pattern for color and texture image retrieval. *Neurocomputing*. 2015;165:255–269. doi: 10.1016/j.neucom.2015.03.015
42. Tunali I, Hall LO, Napel S, et al. Stability and reproducibility of computed tomography radiomic features extracted from peritumoral regions of lung cancer lesions. *Med Phys*. 2019;46(11):5075–5085. doi: 10.1002/mp.13808
43. Nasief H, Zheng C, Schott D, et al. A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer. *NPJ Precis Oncol*. 2019;3(1):25. doi: 10.1038/s41698-019-0096-z
44. Cui Y, Li Y, Xing D, et al. Improving the prediction of benign or malignant breast masses using a combination of image biomarkers and clinical parameters. *Front Oncol*. 2021;11:629321. doi: 10.3389/fonc.2021.629321
45. Mao N, Yin P, Wang Q, et al. Added value of radiomics on mammography for breast cancer diagnosis: a feasibility study. *J Am Coll Radiol*. 2019;16(4 Pt A):485–491. doi: 10.1016/j.jacr.2018.09.041
46. Fanizzi A, TM, Losurdo L, et al. A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis. *BMC Bioinformatics*. 2020;21(Suppl 2):91. doi: 10.1186/s12859-020-3358-4
47. Stelzer PD, Steding O, Raudner MW, et al. Combined texture analysis and machine learning in suspicious calcifications detected by mammography: Potential to avoid unnecessary stereotactical biopsies. *Eur J Radiol*. 2020;132:109309. doi: 10.1016/j.ejrad.2020.109309
48. Karahaliou A, Skiadopoulos S, Boniatis I, et al. Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis. *Br J Radiol*. 2007;80(956):648–656. doi: 10.1259/bjr/30415751
49. Li H, Mendel KR, Lan L, et al. Digital mammography in breast cancer: Additive value of radiomics of breast parenchyma. *Radiology*. 2019;291(1):15–20. doi: 10.1148/radiol.2019181113
50. Parekh VS, Jacobs MA. MPRAD: A multiparametric radiomics framework. *arXiv*. 2018.
51. Rozhkova NI, Bozhenko VK, Burdina II, et al. Radiogenomics of breast cancer as new vector of interdisciplinary integration of radiation and molecular biological technologies (literature review). *Med Alph*. 2020;(20):21–29. doi: 10.33667/2078-5631-2020-20-21-29
52. Wang Z, Lin F, Ma H, et al. Contrast-Enhanced spectral mammography-based radiomics nomogram for the prediction of neo-adjuvant chemotherapy-insensitive breast cancers. *Front Oncol*. 2021;11:605230. doi: 10.3389/fonc.2021.605230
53. Zhang HX, Sun ZQ, Cheng YG, et al. A pilot study of radiomics technology based on X-ray mammography in patients with triple-negative breast cancer. *J Xray Sci Technol*. 2019;27(3):485–492. doi: 10.3233/XST-180488
54. Morozov SP, Vladzimirsky AV, Klyashtornyy VG, et al. Clinical acceptance of software based on artificial intelligence technologies (Radiology). Moscow: Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies; 2019. 45 p.
55. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315–1316. doi: 10.1097/JTO.0b013e3181ec173d
56. Rahbar H, McDonald ES, Lee JM, et al. How can advanced imaging be used to mitigate potential breast cancer overdiagnosis? *Academic Radiology*. 2016;23(6):768–773. doi: 10.1016/j.acra.2016.02.008
57. Pinker K. Beyond breast density: Radiomic phenotypes enhance assessment of breast cancer risk. *Radiology*. 2019;290(1):50–51. doi: 10.1148/radiol.2018182296
58. Sun W, Tseng TL, Qian W, et al. Using multiscale texture and density features for near-term breast cancer risk analysis. *Med Phys*. 2015;42(6):2853–2862. doi: 10.1118/1.4919772
59. Kontos D, Winham SJ, Oustimov A, et al. Radiomic phenotypes of mammographic parenchymal complexity: Toward augmenting breast density in breast cancer risk assessment. *Radiology*. 2019;290(1):41–49. doi: 10.1148/radiol.2018180179
60. Yang J, Wang T, Yang L, et al. Preoperative prediction of axillary lymph node metastasis in breast cancer using mammography-based radiomics method. *Sci Rep*. 2019;9(1):4429. doi: 10.1038/s41598-019-40831-z
61. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep*. 2016;6:23428. doi: 10.1038/srep23428
62. Lu L, Ehmke RC, Schwartz LH, Zhao B. et al. Assessing agreement between radiomic features computed for multiple CT imaging settings. *PLoS One*. 2016;11(12):e0166550. doi: 10.1371/journal.pone.0166550
63. Velazquez ER, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. *Sci Rep*. 2013;3:3529. doi: 10.1038/srep03529
64. Qiu Q, Duan J, Gong G, et al. Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma. *Transl Cancer Res*. 2017;6(5). doi: 10.21037/tcr.2017.09.47
65. Qiu Q, Duan J, Duan Z, et al. Reproducibility and non-redundancy of radiomic features extracted from arterial phase CT scans in hepatocellular carcinoma patients: Impact of tumor segmentation variability. *Quant Imaging Med Surg*. 2019;9(3):453–464. doi: 10.21037/qims.2019.03.02
66. Hunter LA, Krafft S, Stingo F, et al. High quality machine-robust image features: Identification in nonsmall cell lung cancer computed tomography images. *Med Phys*. 2013;40(12):121916. doi: 10.1118/1.4829514
67. O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*. 2017;14(3):169–186. doi: 10.1038/nrclinonc.2016.162
68. Chalkidou A, O'Doherty M, Marsden PK. False discovery rates in PET and CT studies with texture features: a systematic review. *PLoS One*. 2015;10(5):e0124165. doi: 10.1371/journal.pone.0124165

69. Mann RM, Kuhl CK, Moy L. Contrast-enhanced MRI for breast cancer screening. *J Magn Reson Imaging*. 2019;50(2):377–390. doi: 10.1002/jmri.26654
70. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer*. 2017;3:43. doi: 10.1038/s41523-017-0045-3
71. Whitney HM, Taylor NS, Drukker K, et al. Additive benefit of radiomics over size alone in the distinction between benign lesions and luminal a cancers on a large clinical breast MRI dataset. *Acad Radiol*. 2019;26(2):202–209. doi: 10.1016/j.acra.2018.04.019
72. Crivelli P, Ledda RE, Parascandolo N, et al. A new challenge for radiologists: radiomics in breast cancer. *BioMed Research International*. 2018;2018:6120703. doi: 10.1155/2018/6120703
73. Bickelhaupt S, Paech D, Kickingeder P, et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography. *J Magn Reson Imaging*. 2017;46(2):604–616. doi: 10.1002/jmri.25606
74. Han L, Zhu Y, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol*. 2019;29(7):3820–3829. doi: 10.1007/s00330-018-5981-2
75. Dong Y, Feng Q, Yang W, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur Radiol*. 2018;28(2):582–591. doi: 10.1007/s00330-017-5005-7
76. Ma W, Ji Y, Qi L, et al. Breast cancer Ki67 expression prediction by DCE-MRI radiomics features. *Clin Radiol*. 2018;73(10):909.e1–909.e5. doi: 10.1016/j.crad.2018.05.027
77. Jagadish K, Sheela GM, Naidu BP, et al. Big data analytics and radiomics to discover diagnostics and therapeutics for gastric cancer. In: *Recent Advancements in Biomarkers and Early Detection of Gastrointestinal Cancers*. 2020. P. 213–219. doi: 10.1007/978-981-15-4431-6_12
78. Scheckenbach K. Radiomics: Big data instead of biopsies in the future? *Laryngorhinotologie*. 2018;97(S 01):S114–S141. doi: 10.1055/s-0043-121964
79. European Society of Radiology (ESR). Medical imaging in personalised medicine: a white paper of the research committee of the European Society of Radiology (ESR). *Insights Imaging*. 2015;6(2):141–155. doi: 10.1007/s13244-015-0394-0

ОБ АВТОРАХ

* Семенов Серафим Сергеевич;

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: <https://orcid.org/0000-0003-2585-0864>;
eLibrary SPIN: 4790-0416; e-mail: s.semenov@npcmr.ru

Говорухина Вероника Георгиевна;

e-mail: govorukhinaver@gmail.com;
ORCID: <https://orcid.org/0000-0002-1611-9618>;
eLibrary SPIN: 7038-8580

Гележе Павел Борисович, к.м.н.;

e-mail: gelezhe.pavel@gmail.com;
ORCID: <https://orcid.org/0000-0003-1072-2202>;
eLibrary SPIN: 4841-3234

Диденко Вера Владимировна;

e-mail: didenko@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-9068-1273>;
eLibrary SPIN: 5033-8376

Морозов Сергей Павлович, д.м.н., профессор;

e-mail: morozov@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720

Андрейченко Анна Евгеньевна, к.физ.-мат.н.;

e-mail: a.andreychenko@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-6359-0763>;
eLibrary SPIN: 6625-4186

AUTHORS' INFO

* Serafim S. Semenov, MD;

address: 24-1 Petrovka street, 127051 Moscow, Russia;
ORCID: <https://orcid.org/0000-0003-2585-0864>;
eLibrary SPIN: 4790-0416; e-mail: s.semenov@npcmr.ru

Veronika G. Govorukhina, MD;

e-mail: govorukhinaver@gmail.com;
ORCID: <https://orcid.org/0000-0002-1611-9618>;
eLibrary SPIN: 7038-8580

Pavel B. Gelezhe, MD, Cand. Sci. (Med.);

e-mail: gelezhe.pavel@gmail.com;
ORCID: <https://orcid.org/0000-0003-1072-2202>;
eLibrary SPIN: 4841-3234

Vera V. Didenko, MD;

e-mail: didenko@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-9068-1273>;
eLibrary SPIN: 5033-8376

Sergey P. Morozov, MD, Dr. Sci. (Med.), Professor;

e-mail: morozov@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-6545-6170>;
eLibrary SPIN: 8542-1720

Anna E. Andreychenko, Cand. Sci. (Phys.-Math.);

e-mail: a.andreychenko@npcmr.ru;
ORCID: <https://orcid.org/0000-0001-6359-0763>;
eLibrary SPIN: 6625-4186

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD60532>

Кавернозные мальформации головного мозга и современные взгляды на их лечение

Е.Н. Гиря¹, В.Е. Синицын², А.С. Токарев^{1, 3}¹ Научно-исследовательский институт скорой помощи имени Н.В. Склифосовского, Москва, Российская Федерация² Московский государственный университет имени М.В. Ломоносова, Москва, Российская Федерация³ Департамент здравоохранения города Москвы, Москва, Российская Федерация

АННОТАЦИЯ

Кавернозные мальформации головного мозга благодаря развитию современных методов нейровизуализации являются в последние годы всё чаще обнаруживаемой патологией. Несмотря на доброкачественный характер течения в большинстве случаев, данные образования могут приводить к развитию судорожного синдрома и серьёзным неврологическим нарушениям. Как правило, причинами клинических симптомов являются кровоизлияния в структуру каверном и окружающую паренхиму головного мозга. Выбор тактики ведения пациентов с кавернозными мальформациями головного мозга зависит от типа мальформации, её размеров, локализации, наличия повторных кровоизлияний и клинической картины.

Данный обзор литературы посвящён современным методам лечения кавернозных мальформаций головного мозга, в частности хирургическим подходам. В случаях глубинного расположения очагов в функционально значимых зонах головного мозга, для которых характерен максимальный риск осложнений при хирургическом вмешательстве, альтернативными выступают методы лучевой терапии, такие как стереотаксическая радиохирургия, протонная терапия. Рассматриваются возможности, эффективность и безопасность стереотаксического радиохирургического лечения, использование протонной терапии в лечении кавернозных мальформаций. Выявлены преимущества лучевых методов лечения кавернозных мальформаций.

Ключевые слова: кавернозные мальформации; лучевая диагностика; МРТ; обзор; аппарат Гамма-нож; протонная терапия; радиохирургическое лечение; стереотаксическая лазерная абляция.

Как цитировать

Гиря Е.Н., Синицын В.Е., Токарев А.С. Кавернозные мальформации головного мозга и современные взгляды на их лечение // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 200–210. DOI: <https://doi.org/10.17816/DD60532>

DOI: <https://doi.org/10.17816/DD60532>

Cavernous malformations of the brain and modern views on their treatment

Elena N. Girya¹, Valentin E. Sinitsyn², Aleksei S. Tokarev^{1, 3}

¹ Sklifosovsky Research Institute for Emergency Medicine, Moscow, Russian Federation

² Lomonosov Moscow State University, Moscow, Russian Federation

³ Moscow Health Department, Moscow, Russian Federation

ABSTRACT

Cavernous malformations of the brain have become an increasingly common pathology in recent years, thanks to the advancement of modern methods of neuroimaging. Despite the benign nature of the course in most cases, these formations can cause convulsions and serious neurological disorders. Typically, clinical manifestations are caused by hemorrhages in the structure of the cavernous and surrounding parenchyma of the brain. The management strategy chosen for patients with cerebral cavernous malformations is determined by the type of malformation, its size, localization, the presence of repeated hemorrhages, and the clinical picture.

This literature review focuses on modern methods of treating cerebral cavernous malformations. The main methods of treatment for cavernous malformations of the brain, particularly surgical treatment, have been analyzed. If surgical intervention is not possible, alternative methods of treatment include radiation therapy, such as stereotactic radiosurgery, and proton therapy, in cases of deep location of foci in functionally significant areas of the brain, which are characterized by the highest risk of complications. The possibilities, efficacy, and safety of stereotactic radiosurgical treatment are discussed, as well as the use of proton therapy in the treatment of cavernous malformations. Furthermore, radiation therapy has been shown to be beneficial for cavernous malformations.

Keywords: cavernous malformations; radiation diagnostics; MRI; review; Gamma knife; proton therapy; radiosurgical treatment; stereotactic laser ablation.

To cite this article

Girya EN, Sinitsyn VE, Tokarev AS. Cavernous malformations of the brain and modern views on their treatment. *Digital Diagnostics*. 2021;2(2):200–210. DOI: <https://doi.org/10.17816/DD60532>

Received: 12.02.2021

Accepted: 05.04.2021

Published: 01.07.2021

DOI: <https://doi.org/10.17816/DD60532>

大脑海绵状畸形及其治疗的现代观点

Elena N. Girya¹, Valentin E. Sinitsyn², Aleksei S. Tokarev^{1, 3}

¹ Sklifosovsky Research Institute for Emergency Medicine, Moscow, Russian Federation

² Lomonosov Moscow State University, Moscow, Russian Federation

³ Moscow Health Department, Moscow, Russian Federation

简评

由于现代神经影像学方法的发展，近年来大脑海绵状畸形已成为越来越可检测的病理。尽管在大多数情况下病程的性质是良性的，但这些形成可导致惊厥综合征和严重神经系统疾病的发展。基本上临床症状的原因是洞穴结构和大脑周围实质的出血。大脑海绵状畸形患者的管理策略的选择取决于畸形的类型，其大小，定位，反复出血的存在和临床情况。

这篇文献综述致力于海绵体畸形的现代治疗方法。我们分析治疗脑海绵状畸形的�主要方法，特别是手术治疗。无法手术干预的时候，在大脑功能显着区域的病灶深度定位的情况下，其特征在于并发症的最大风险，放射治疗的替代方法是如立体定向放射外。同时审查立体定向放射外科治疗的可能性，有效性和安全性，使用质子治疗治疗海绵体畸形。揭示了治疗海绵体畸形的辐射方法的优点。

关键词：海绵体畸形；放射诊断；MRI；综述；伽玛刀装置；质子治疗；放射外科治疗；立体定向激光消融。

引用本文：

Girya EN, Sinitsyn VE, Tokarev AS. 大脑海绵状畸形及其治疗的现代观点. *Digital Diagnostics*. 2021;2(2):200–210.

DOI: <https://doi.org/10.17816/DD60532>

收到: 12.02.2021

接受: 05.04.2021

发布日期: 01.07.2021

ВВЕДЕНИЕ

Кавернозные мальформации (КМ) представляют собой сосудистые образования головного и спинного мозга с низким уровнем кровотока, состоящие из каверн с эндотелиальной выстилкой [1–4]. КМ обнаруживаются как в супра-, так и в инфратенториальных областях головного мозга, реже — в спинном мозге [5–8].

Данные образования являются вторыми по распространённости сосудистыми мальформациями центральной нервной системы после аномалий венозного развития [9–11].

Распространённость КМ одинакова у мужчин и женщин. Постановка диагноза, как правило, происходит между вторым и четвертым десятилетием жизни, хотя КМ также могут выявляться у детей. В большинстве случаев КМ могут клинически не проявляться, однако с течением времени могут вызвать серьёзную очаговую и общемозговую неврологическую симптоматику, часто обусловленную разрывом КМ с кровоизлиянием в структуру образований и окружающее вещество мозга [12].

Несмотря на то, что, по данным ряда исследований, к настоящему времени уровни рисков развития кровоизлияний и судорог у данной категории пациентов установлены, чёткое выявление модифицируемых факторов риска представляет собой значительную сложность.

Ведение больных с КМ включает динамическое наблюдение за ними либо выполнение хирургического вмешательства [13, 14].

ХИРУРГИЧЕСКОЕ ЛЕЧЕНИЕ КАВЕРНОЗНЫХ МАЛЬФОРМАЦИЙ ГОЛОВНОГО МОЗГА

Микрохирургическая резекция по-прежнему остаётся золотым стандартом лечения КМ, способным навсегда избавить пациента от сопутствующих проявлений этой патологии и рисков развития неврологического дефицита, связанного с кровоизлияниями. Оценка риска выполнения хирургического вмешательства зависит от размеров и локализации образования, близости к поверхности мозга и опыта хирурга [15]. Целью хирургического лечения является полное удаление КМ и окружающих потенциальных эпилептогенных зон [16]. Однако, если эти образования расположены рядом с жизненно важными структурами (расстояние составляет менее 1 см), полное удаление может приводить к послеоперационным неврологическим нарушениям. В случаях локализации КМ в таких областях головного мозга, как таламус, базальные ганглии или ствол головного мозга, хирургическое вмешательство, как правило, выполняется только при частых рецидивирующих кровоизлияниях или при существенном ухудшении состояния пациента.

Ряд авторов отмечает, что относительно низкая частота осложнений хирургического лечения тем не менее превышает риск развития кровотечений у тех пациентов, у которых ранее они не были диагностированы. Таким образом, хирургическое удаление бессимптомных очагов, особенно при их глубокой локализации или расположении в области ствола мозга, является неоправданным.

Глубокое расположение очагов (в базальных ганглиях или таламусе) требует выполнения технически сложной операции, при которой могут быть затронуты критически важные структуры головного мозга, в том числе ядра и тракты белого вещества; имеется риск повреждений перфорирующих артерий. Частота послеоперационных осложнений при выполнении данного вида вмешательств даже у опытных специалистов составляет 5–18%, частота летальных исходов приближается к 2% [17].

В целом, несмотря на прогресс и совершенствование хирургической техники, остаётся большое количество пациентов, которым хирургические методы не показаны или лечение которых оказывается неполным, и КМ продолжают функционировать. В лечении этого контингента больных всё большее место отводится стереотаксическому облучению (радиохирургия и стереотаксическая лучевая терапия).

ВОЗМОЖНОСТИ, ЭФФЕКТИВНОСТЬ И БЕЗОПАСНОСТЬ РАДИОХИРУРГИЧЕСКОГО ЛЕЧЕНИЯ КАВЕРНОЗНЫХ МАЛЬФОРМАЦИЙ ГОЛОВНОГО МОЗГА

В последние годы появляется всё больше сообщений об использовании лучевых методов лечения артериовенозных мальформаций и дуральных артериовенозных фистул [18–20]. В отдельных работах продемонстрирована возможность применения этого метода и к лечению КМ. Основным показанием к использованию такого рода лечения являются КМ до 3 см в диаметре, расположенные в глубоких отделах мозга — зонах, для которых характерен максимальный риск осложнений. Одним из основных методов лучевой терапии, применяемых в настоящее время для лечения пациентов с КМ, является стереотаксическое радиохирургическое лечение. В ряде неконтролируемых исследований было показано, что риск рецидива кровотечения после выполнения радиохирургических вмешательств снижается у больных через 2 года.

Ретроспективное исследование С.С. Lee и соавт. было посвящено изучению эффективности и безопасности радиохирургического лечения с использованием аппарата Гамма-нож в лечении пациентов с КМ головного мозга [21]. Авторы проанализировали результаты лечения 261 пациента с 331 симптоматической КМ; средний возраст пациентов составил 39,9 года, средний объём

КМ — 3,1 мл. Средняя доза облучения в течение всего периода лечения составила 11,9 Гр. Пациентов наблюдали в течение 69 мес. Ряду пациентов диагноз КМ был поставлен после первоначального кровоизлияния; всего до лечения было диагностировано 136 кровоизлияний.

В целом исследователями был сделан вывод, что радиохирургическое лечение снизило риск кровоизлияний у пациентов с КМ, поэтому использованный метод является эффективным альтернативным вариантом лечения пациентов при сложности осуществления хирургического доступа к КМ или при наличии у пациента тяжёлых сопутствующих заболеваний.

A.U. Kefeli и соавт. была предпринята попытка оценить результаты лечения КМ ствола головного мозга с помощью Гамма-ножа [22]. В исследование было включено 82 пациента, у которых до лечения отмечалось от 1 до 3 геморагических событий, подтверждённых рентгенологически. После проведённого лечения средний целевой объём составил 0,3 мл, предельная доза облучения — 12 Гр. Пациентов наблюдали в среднем в течение 25,5 мес до операции и в течение 50,3 мес после операции. Годовая частота кровоизлияний до лечения составляла 8,6%. После лечения в течение всего периода наблюдения только у 3 пациентов произошло повторное кровоизлияние. Таким образом, частота повторных кровоизлияний в течение года составила 0,87%, т.е. риск подобных осложнений был значительно снижен при использовании радиохирургического метода.

Величина риска кровоизлияния при КМ до настоящего времени чётко не определена. Наблюдение за естественным течением КМ свидетельствует, что годовой риск кровоизлияния колеблется от 2,3 до 4,1%, тогда как при хирургическом лечении его величина составляет от 2,7 до 6,8% в год [23, 24]. Однако риск рецидивирующего кровоизлияния при КМ увеличивается после первоначального кровоизлияния в разы, достигая 40% [25].

R. Wen и соавт. был выполнен метаанализ с целью оценки клинической эффективности радиохирургического лечения КМ с использованием аппарата Гамма-нож, по результатам которого не наблюдалось существенных различий по частоте кровоизлияний между первыми двумя годами послеоперационного периода и последующими двумя годами (ОР 2,81; 95% ДИ 0,20–13,42) [26].

В исследованиях последних лет установлено ежегодное снижение частоты кровотечений с 39,5 до 7,2% в течение первых 2 лет после проведения лечения КМ с использованием аппарата Гамма-нож, а в последующие годы — с 3,6 до 1% [22, 27, 28].

D. Kondziolka и соавт. при изучении частоты кровоизлияний в течение всего срока наблюдения КМ обнаружили, что годовая частота кровотечений до выполнения радиохирургического вмешательства составляет 5,9%, через 2 года — 1,1% [29]. R. Aboukais и соавт. продемонстрировали снижение этого показателя с 3,16 до 2,46% [30]. По данным R. Lopez-Serrano

и соавт., ежегодная частота кровотечений составляла 3,06% до радиохирургического лечения и порядка 1,4% после его проведения [31].

Некоторые авторы считают, что эффективность использования Гамма-ножа в полной мере проявляется через 2–3 года после радиохирургического лечения, что обусловлено уменьшением объёма КМ со временем вследствие процессов склероза и тромбооблитерации сосудов после облучения [31, 32].

Обсуждается вопрос, связано ли снижение частоты кровоизлияний с выполнением радиохирургических вмешательств или является следствием естественного течения КМ [21].

Предполагают, что в основе механизмов радиохирургического лечения сосудистых мальформаций лежат такие процессы, как пролиферация эндотелиальных клеток и гиалинизация, приводящая к закрытию просвета сосуда. R. Gewirtz и соавт. и I. Nyáry и соавт. выполнили гистологическое исследование тканей КМ у пациентов, которым проводилось радиохирургическое лечение, в результате чего были выявлены признаки фибриноидного некроза, разрушения эндотелиальных клеток и выраженного фиброза в строме соединительной ткани [33, 34].

K. Park и соавт. были проанализированы отдалённые результаты радиохирургического лечения симптоматических КМ ствола головного мозга с использованием Гамма-ножа у 45 пациентов (14 мужчин, 31 женщина) [27]. Пациентов наблюдали более 5 лет; средняя длительность 9,31 (от 5,1 до 19,4) года. У всех больных в анамнезе было одно или несколько симптоматических кровотечений до проведения радиохирургического лечения. Эти кровоизлияния сопровождались проявлениями неврологического дефицита, в том числе нарушениями функции черепно-мозговых нервов, гемипарезами, гемисенсорным дефицитом, спастичностью, хореей. Средний целевой объём КМ составил 1,82 см³, медиана предельной дозы облучения — 13 Гр. На основании полученных результатов авторами было сделано заключение, что радиохирургическое лечение с использованием Гамма-ножа является безопасным и клинически эффективным методом лечения КМ, позволяющим снизить частоту рецидивов кровоизлияния.

До 2019 г. было проведено три крупных исследования по применению Гамма-ножа (>100 случаев, не менее 4 лет наблюдения) в лечении повторных геморагических или симптоматических КМ [35–37]. В общей сложности в эти исследования было включено 530 пациентов. Результаты исследований Y. Kida показали, что ежегодная частота кровоизлияний после радиохирургического лечения КМ с использованием Гамма-ножа снизилась с 9,5% (в течение 1 года) до 4,7% (в течение 2 лет) [37]. В других исследованиях ежегодная частота кровоизлияний после применения этого метода снизилась с 15% (в течение 2 лет) до 2,4% (после 2 лет) [35].

В качестве факторов, влияющих на частоту кровоизлияний у пациентов, которым проводится радиохирургическое лечение, одни исследователи рассматривают пол пациента, выраженность неврологических проявлений до вмешательства, размер КМ, степень отёка окружающих тканей, дозу облучения [36]. В то же время В. Kim и соавт. не выявили статистически значимых различий по частоте кровоизлияний в зависимости от объёма КМ, дозы облучения, пола и возраста пациента на момент проведения лечения с использованием Гамма-ножа [38].

Общим осложнением для большинства пациентов с КМ являются эпилептические приступы, при этом предполагают наличие корреляции между развитием кровоизлияний и судорогами. Нередко при кровоизлияниях пациенты с КМ испытывают сопутствующие головные боли или головокружения [37]. Экспериментальные исследования показали, что подобным эпилептогенным фактором может служить отложение метаболитов кровяного сгустка, особенно железа. Исследованиями с помощью магнитно-резонансного томографа (МРТ) подтверждена связь развития судорог с кровоизлияниями по времени у данной категории пациентов. Другим фактором риска развития судорог при данной патологии считается локализация КМ, в первую очередь супратенториальная, архикортикальная и мезиотемпоральная. В исследовании К. Menzler и соавт. в сопоставлении с данными МРТ было показано, что из 81 пациента с КМ с вовлечением коры головного мозга у 49 наблюдались судороги, тогда как ни у одного из 17 пациентов с исключительно подкорковой локализацией КМ судорог не отмечено [39].

Рассматривая осложнения радиохирургического лечения КМ, следует отметить в первую очередь опасность развития радиационного поражения головного мозга с появлением неврологических нарушений, включающих головную боль, головокружение, паралич лицевого нерва, парестезию лица, диплопию, дизартрию и слабость в конечностях [30]. Другим серьёзным побочным эффектом является радиационный некроз, который может способствовать развитию опухолей [40].

Следует отметить, что некоторые исследователи выражают беспокойство по поводу способности радиационного облучения индуцировать образование новых КМ, особенно у детей, а также в случаях семейного заболевания [41].

Оптимальная предельная доза облучения при проведении радиохирургического лечения КМ ствола головного мозга чётко не определена, однако С. Lee и соавт. и В. S. Kim и соавт. полагают, что предельная величина 11 Гр является достаточной для уменьшения риска радиационных осложнений этого вида лечения [21, 38]. Использование дозы такого уровня является эффективным, при этом было зафиксировано снижение риска кровоизлияний до 2,4% через 2 года после применения Гамма-ножа, а также улучшение неврологического статуса;

частота радиационно-индуцированных осложнений составила 2,32%.

В целом безопасная в отношении радиотоксичности терапевтическая доза облучения при радиохирургическом лечении КМ ствола головного мозга составляет 11–13 Гр [42].

В соответствии с современными рекомендациями к выполнению радиохирургических вмешательств, этот подход следует рассматривать в качестве метода лечения одиночных КМ с кровоизлиянием в анамнезе в тех областях головного мозга, где хирургический риск повреждения тканей является неприемлемо высоким [43]. Экспертное мнение предполагает, что применение этих методов не рекомендуется в случаях, когда КМ доступны для хирургического лечения, а также в отсутствии симптомов и при семейной форме патологии.

Потенциально перспективным методом лечения КМ с эпилептоидными проявлениями рассматривается также стереотаксическая лазерная абляция этих образований [44].

Таким образом, радиохирургическое лечение КМ головного мозга является относительно безопасным подходом, при использовании которого не отмечается разрывов сосудов и повреждений мозговой ткани. Использование этого метода подразумевает однократное подведение всей дозы облучения, с одной стороны, необходимой для получения необходимого результата, с другой — достаточно безопасной для окружающего мозгового вещества. Такой подход характеризуется наиболее высокой эффективностью в лечении КМ. В то же время размеры (объём) КМ в ряде случаев не позволяют безопасно использовать желаемые дозы облучения, при этом уменьшение дозы приводит к снижению эффективности воздействия [45].

По мнению С. С. Lee и соавт., в прошлом эффективность радиохирургического лечения КМ была ограничена недостаточными возможностями методов нейровизуализации, высокими дозами облучения (>15 Гр) и неполным или чрезмерным охватом целевой области [21]. Достижения в области нейровизуализации (применение МРТ), оптимизация доз облучения и планирование вмешательства с помощью соответствующего программного обеспечения значительно снизили риск осложнений при применении данного вида лечения.

ПРОТОННАЯ ТЕРАПИЯ В ЛЕЧЕНИИ КАВЕРНОЗНЫХ МАЛЬФОРМАЦИЙ

Протонная терапия — ещё более усовершенствованный метод лучевой терапии при невозможности хирургического удаления или отказе больного от операции. Протонная терапия КМ, как и стереотаксическое радиохирургическое лечение, решает задачу достижения облитерации в структуре образования и тем самым снижения риска последующих кровоизлияний. Преимуществом

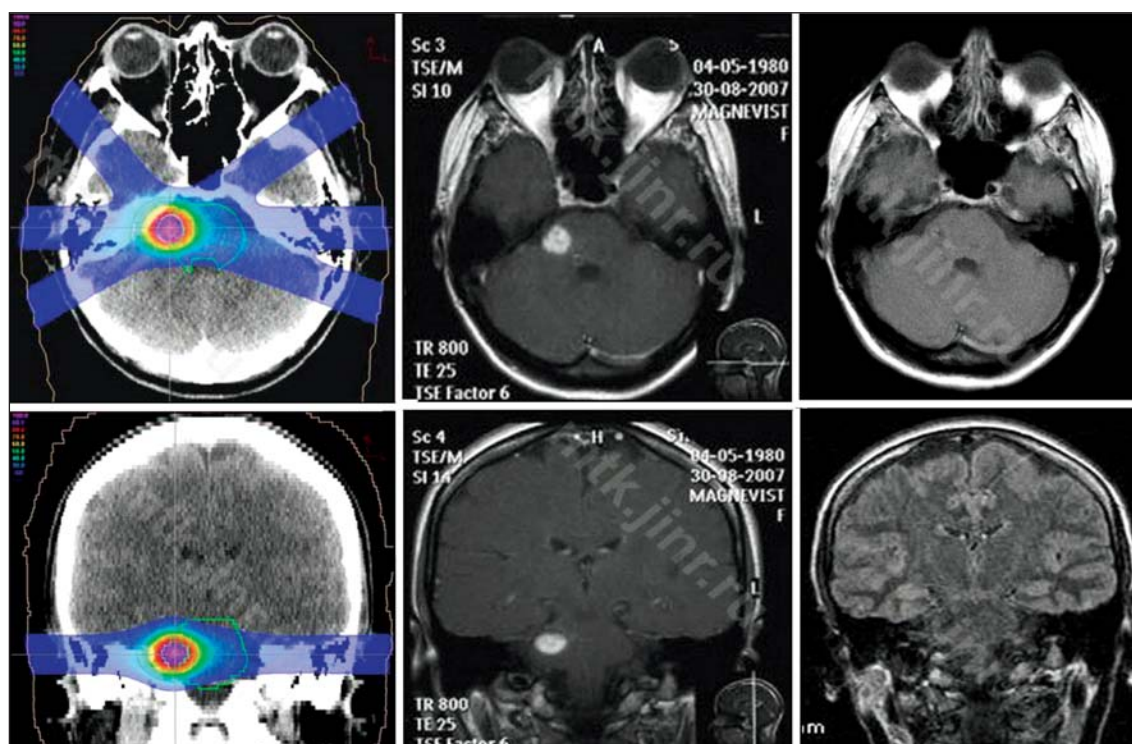


Рис. План протонной радиохирургии каверномы околостволовой локализации: МРТ с контрастом до лечения и спустя 3 мес (полная резорбция каверномы).

протонной терапии является возможность осуществления достаточно точного облучения опухоли (точность около 0,5 мм) с минимальным повреждением здоровых тканей и снижением риска побочных эффектов [46].

Эффект от лечения наблюдается в период от 5 до 90 мес. Полная облитерация новообразования достигается в 70% случаев. План протонной радиохирургии каверномы околостволовой локализации представлен на рисунке [47].

ЗАКЛЮЧЕНИЕ

Кавернозные мальформации представляют собой сосудистые новообразования головного мозга, механизм развития которых основан на процессах сосудистой пролиферации, дисморфизма и геморрагической ангиопатии. Основными причинами клинических симптомов являются повторные кровоизлияния в структуру кавернозных ангиом с последующим отложением железа в окружающих тканях головного мозга, что может служить причиной возникновения очагов эпилептогенеза, особенно при локализации каверном в мезиотемпоральной и архикортикальной областях головного мозга. Совершенствование методов диагностики и лечения данной патологии является мультидисциплинарной проблемой.

Выбор метода лечения зависит от типа мальформации, её размеров, локализации, наличия предыдущих кровоизлияний. В связи с тем, что у значительной части пациентов с КМ риск осложнений хирургического вмешательства является высоким, в отношении этой

категории больных, как и пациентов с семейной формой КМ, крайне актуальной является разработка альтернативных методов хирургического лечения. К таким направлениям относятся используемые в настоящее время методы стереотаксической лучевой терапии.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении поисково-аналитической работы и подготовке статьи.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. Е.Н. Гиря — поиск публикаций по теме, анализ литературы, написание текста; А.С. Токарев — определение основной направленности обзора, экспертная оценка обзора литературы, обработка полученных результатов; В.Е. Синицын — экспертная оценка обзора литературы, обработка полученных результатов, систематизация и финальное редактирование обзора. Авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Funding source. The study had no sponsorship.

Competing interests. The authors declare no obvious and potential conflicts of interest related to the publication of this article.

Authors' contribution. E.N. Giryа — search for publications on the topic, analysis of literature, text writing; A.S. Tokarev — determination of the main focus of the review, expert assessment of

the literature review, processing of the results obtained; V.E. Siniysyn — expert assessment of the literature review, processing of the obtained results, systematization and final editing of the review. All authors made a substantial contribution to the conception

of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

СПИСОК ЛИТЕРАТУРЫ

1. Муха А.М., Дашьян В.Г., Крылов В.В. Кавернозные мальформации головного мозга // Неврологический журнал. 2013. Т. 18, № 5. С. 46–51.
2. Попов В.Е., Лившиц М.И., Башлачев М.Г., Наливкин А.Е. Кавернозные мальформации у детей: обзор литературы // Альманах клинической медицины. 2018. Т. 46, № 2. С. 146–159. doi: 10.18786/2072-0505-2018-46-2-146-159
3. Caton M.T., Shenoy V.S. Cerebral Cavernous Malformations. Treasure Island (FL): StatPearls Publishing; 2019. doi: 10.20959/wjpr202011-18275
4. Flemming K.D., Brown R.D. Epidemiology and natural history of intracranial vascular malformations. In: H.R. Winn, ed. Youmans & Winn neurological surgery, 7th ed. Amsterdam: Elsevier; 2017. P. 3446–3463e7.
5. Готко А.В., Kivelev J.V., Сон А.С. Кавернозные мальформации головного и спинного мозга // Український нейрохірургічний журнал. 2013. № 3. С. 10–15.
6. Родич А., Смянович А., Сидорович Р., и др. Современные подходы к хирургическому лечению кавернозных ангиом головного мозга // Наука и инновации. 2018. Т. 10, № 188. С. 70–73.
7. Gross B.A., Du R. Natural history of cerebral arteriovenous malformations: a meta-analysis // J Neurosurg. 2013. Vol. 118, N 2. P. 437–443. doi: 10.3171/2012.10.JNS121280
8. Kearns K.N., Chen C.J., Tvrdik P., et al. Outcomes of surgery for brainstem cavernous malformations: a systematic review // Stroke. 2019. Vol. 50, N 10. P. 2964–2966. doi: 10.1161/STROKEAHA.119.026120
9. Сазонов И.А., Белоусова О.Б. Кавернозная мальформация, вызвавшая развитие обширной острой субдуральной гематомы. Случай из практики и обзор литературы // Вопросы нейрохирургии им. Н.Н. Бурденко. 2019. Т. 83, № 3. С. 73–76. doi: 10.17116/neiro20198303173
10. Mouchtouris N., Chalouhi N., Chitale A., et al. Management of cerebral cavernous malformations: from diagnosis to treatment // Scientific World Journal. 2015. Vol. 2015, N 808314. doi: 10.1155/2015/808314
11. Negoto T., Terachi S., Baba Y., et al. Symptomatic brainstem cavernoma of elderly patients: timing and strategy of surgical treatment. Two case reports and review of the literature // World Neurosurgery 2018. Vol. 111. P. 227–234. doi: 10.1016/j.wneu.2017.12.111
12. Batra S., Lin D., Recinos P.F., et al. Cavernous malformations: natural history, diagnosis and treatment // Nature Reviews Neurology. 2009. Vol. 5, N 12. P. 659–670. doi: 10.1038/nrneurol.2009.177
13. Al-Shahi Salman R., White P.M., Counsell C.E., et al. Outcome after conservative management or intervention for unruptured brain arteriovenous malformations // JAMA. 2014. Vol. 311, N 16. P. 1661–1669. doi: 10.1001/jama.2014.3200
14. Hakim A.A., Gralla J., Rozeik C., et al. Anomalies and normal variants of the cerebral arterial supply: A comprehensive pictorial review with a proposed workflow for classification and significance // J Neuroimaging. 2018. Vol. 28, N 1. P. 14–35. doi: 10.1111/jon.12475
15. Полковников А.Ю. Артериовенозные мальформации полушарий большого мозга с ядром малого и среднего размеров, особенности клинических проявлений и методы хирургического лечения // Актуальні питання фармацевтичної і медичної науки та практики. 2013. Т. 2, № 12. С. 033–038.
16. Chang E.F., Wang D.D., Barkovich A.J., et al. Predictors of seizure freedom after surgery for malformations of cortical development // Ann Neurol. 2011. Vol. 70, N 1. P. 151–162. doi: 10.1002/ana.22399
17. Pasqualin A., Meneghelli P., Giammarusti A., Turazzi S. Results of surgery for cavernomas in critical supratentorial areas // Acta Neurochir Suppl. 2014. Vol. 119. P. 117–123. doi: 10.1007/978-3-319-02411-0_20
18. Маряшев С.А. Стереотаксическое облучение артериовенозных мальформаций головного мозга: Автореф. дис. ... д-ра мед. наук. Москва, 2016. 26 с.
19. Shimizu K., Kosaka N., Yamamoto T., et al. Arterial spin labeling perfusion-weighted MRI for long-term follow-up of a cerebral arteriovenous malformation after stereotactic radiosurgery // Acta Radiol Short Rep. 2014. Vol. 3, N 1. doi: 10.1177/2047981613510160
20. Seo Y., Kim D.G., Dho Y.S., et al. A Retrospective analysis of the outcomes of dural arteriovenous fistulas treated with gamma knife radiosurgery: a single-institution experience // Stereotactic Functional Neurosurgery. 2018. Vol. 96, N 1. P. 46–53. doi: 10.1159/000486685
21. Lee C.C., Wang W.H., Yang H.C., et al. Gamma Knife radiosurgery for cerebral cavernous malformation // Sci Rep. 2019. Vol. 9, N 1. P. 19743. doi: 10.1038/s41598-019-56119-1
22. Kefeli A.U., Sengoz M., Peker S. Gamma knife radiosurgery for hemorrhagic brainstem cavernomas // Turk Neurosurg. 2018. Vol. 29, N 1. P. 14–19. doi: 10.5137/1019-5149.JTN.21690-17.1
23. Cantu C., Murillo-Bonilla L., Arauz A., et al. Predictive factors for intracerebral hemorrhage in patients with cavernous angiomas // Neurol Res. 2005. Vol. 27, N 3. P. 314–318. doi: 10.1179/016164105X39914
24. Wang C.C., Liu A., Zhang J., et al. Surgical management of brainstem cavernous malformations: report of 137 cases // Surg Neurol. 2003. Vol. 59, N 6. P. 444–454. doi: 10.1016/s0090-3019(03)00187-3
25. Ablak A.A., Turner J.D., Mitha A.P., et al. Surgical approaches to brainstem cavernous malformations // Neurosurgical Focus. 2010. Vol. 29, N 3. P. E8. doi: 10.3171/2010.6.FOCUS10128
26. Wen R., Shi Y., Gao Y., et al. The efficacy of gamma knife radiosurgery for cavernous malformations: a meta-analysis and review // World Neurosurgery. 2019. Vol. 123. P. 371–377. doi: 10.1016/j.wneu.2018.12.046
27. Park S.H., Hwang S.K. Gamma knife radiosurgery for symptomatic brainstem intra-axial cavernous malformations // World Neurosurgery. 2013. Vol. 80, N 6. P. 261–266. doi: 10.1016/j.wneu.2012.09.013
28. Frischer J.M., Gatterbauer B., Holzer S., et al. Microsurgery and radiosurgery for brainstem cavernomas: effective and complemen-

- tary treatment options // *World Neurosurgery*. 2014. Vol. 81, N 3-4. P. 520–528. doi: 10.1016/j.wneu.2014.01.004
29. Kondziolka D., Lunsford L.D., Flickinger J.C., Kestle J.R. Reduction of hemorrhage risk after stereotactic radiosurgery for cavernous malformations // *J Neurosurg*. 1995. Vol. 83, N 5. P. 825–831. doi: 10.3171/jns.1995.83.5.0825
30. Aboukais R., Estrade L., Devos P., et al. Gamma knife radiosurgery of brainstem cavernous malformations // *Stereotact Funct Neurosurg*. 2016. Vol. 94, N 6. P. 397–403. doi: 10.1159/000452844
31. Lopez-Serrano R., Martinez N.E., Kusak M.E., et al. Significant hemorrhage rate reduction after gamma knife radiosurgery in symptomatic cavernous malformations: long-term outcome in 95 case series and literature review // *Stereotact Funct Neurosurg*. 2017. Vol. 95, N 6. P. 369–378. doi: 10.1159/000480664
32. Shin S.S., Murdoch G., Hamilton R.L., et al. Pathological response of cavernous malformations following radiosurgery // *J Neurosurg*. 2015. Vol. 123, N 4. P. 938–944. doi: 10.3171/2014.10.jns14499
33. Gewirtz R.J., Steinberg G.K., Crowley R., Levy R.P. Pathological changes in surgically resected angiographically occult vascular malformations after radiation // *Neurosurgery*. 1998. Vol. 42, N 4. P. 738–741. doi: 10.1097/00006123-199804000-00031
34. Nyary I., Major O., Hanzely Z., Szeifert G.T. Histopathological findings in a surgically resected thalamic cavernous hemangioma 1 year after 40-Gy irradiation // *J Neurosurg*. 2005. Vol. 102, Suppl. P. 56–58. doi: 10.3171/sup.2005.102.s_supplement.0056
35. Monaco E.A., Khan A.A., Niranjan A., et al. Stereotactic radiosurgery for the treatment of symptomatic brainstem cavernous malformations // *Neurosurgical Focus*. 2010. Vol. 29, N 3. E11. doi: 10.3171/2010.7.FOCUS10151
36. Tian K.B., Zheng J.J., Ma J.P., et al. Clinical course of untreated thalamic cavernous malformations: hemorrhage risk and neurological outcomes // *J Neurosurg*. 2017. Vol. 127, N 3. P. 480–491. doi: 10.3171/2016.8.JNS16934
37. Kida Y. Radiosurgery for cavernous malformations in basal ganglia, thalamus and brainstem // *Prog Neurol Surg*. 2009. Vol. 22. P. 31–37. doi: 10.1159/00016338
38. Kim B.S., Yeon J.Y., Kim J.S. et al. Gamma Knife radiosurgery of the symptomatic brain stem cavernous angioma with low marginal dose // *Clin Neurol Neurosurg*. 2014. Vol. 126, P.110–114. doi: 10.1016/j.clineuro.2014.08.028
39. Menzler K. Epileptogenicity of cavernomas depends on (archi-) cortical localization // *Neurosurgery*. 2010. Vol. 67. P. 918–924. doi: 10.1227/NEU.0b013e3181eb5032
40. Liao C., Visocchi M., Zhang W., et al. Management of cerebral radiation necrosis: a retrospective study of 12 patients // *Acta Neurochir Suppl*. 2017. Vol. 124. P. 195–201. doi: 10.1007/978-3-319-39546-3_30
41. Akers A., Salman R.A., Awad I., et al. Synopsis of guidelines for the clinical management of cerebral cavernous malformations: consensus recommendations based on systematic literature review by the angioma alliance scientific advisory board clinical experts panel // *Neurosurgery*. 2017. Vol. 80, N 5. P. 665–680. doi: 10.1093/neuros/nyx091
42. Garcia R.M., Ivan M.E., Lawton M.T. Brainstem cavernous malformations: surgical results in 104 patients and a proposed grading system to predict neurological outcomes // *Neurosurgery*. 2015. Vol. 76, N 3. P. 265–277. doi: 10.1227/NEU.0000000000000602
43. Крылов В.В., Дашьян В.Г., Шетова И.М., и др. Нейрохирургическая помощь больным с сосудистыми заболеваниями головного мозга в Российской Федерации // *Нейрохирургия*. 2017. № 4. С. 11–20.
44. Willie J.T., Malcolm J.G., Stern M.A., et al. Safety and effectiveness of stereotactic laser ablation for epileptogenic cerebral cavernous malformations // *Epilepsia*. 2019. Vol. 60, N 2. P. 220–232. doi: 10.1111/epi.14634
45. Amponsah K., Ellis T.L., Chan M.D., et al. Retrospective analysis of imaging techniques for treatment planning and monitoring of obliteration for gamma knife treatment of cerebral arteriovenous malformation // *Neurosurgery*. 2012. Vol. 71, N 4. P. 893–899. doi: 10.1227/neu.0b013e3182672a83
46. Chen C.C., Chapman P., Petit J., et al. Proton radiosurgery in neurosurgery // *Neurosurg Focus*. 2007. Vol. 23, N 6. E5. doi: 10.3171/FOC-07/12/E5
47. Медико-технический комплекс: Объединенный институт ядерных исследований. Кавернозные ангиомы (каверномы) [Интернет]. Режим доступа: http://mtk.jinr.ru/index.php?option=com_content&task=view&id=33&Itemid=53. Дата обращения: 14.03.2021.

REFERENCES

1. Mukha AM, Dashyan VG, Krylov VV. Cavernous malformations of the brain. *Neurological Journal*. 2013;18(5):46–51. (In Russ).
2. Popov VE, Livshits MI, Bashlachev MG, Nalivkin AE. Cavernous malformations in children: a literature review. *Almanac of Clinical Medicine*. 2018;46(2):146–159. (In Russ). doi: 10.18786/2072-0505-2018-46-2-146-159
3. Caton MT, Shenoy VS. Cerebral Cavernous Malformations. Treasure Island (FL): StatPearls Publishing; 2019. doi: 10.20959/wjpr202011-18275
4. Flemming KD, Brown RD Jr. Epidemiology and natural history of intracranial vascular malformations. In: H.R. Winn, ed. Youmans and Winn neurological surgery, 7th ed. Amsterdam: Elsevier; 2017.
5. Gotko AV, Kivelev JV, Sleep AS. Cavernous malformations of the brain and spinal cord. *Ukrainian neurosurgical Journal*. 2013;(3):10–15. (In Russ).
6. Rodich A, Smeyanovich A, Sidorovich R, et al. Modern approaches to the surgical treatment of cavernous angiomas of the brain. *Science and Innovation*. 2018;10(188):70–73. (In Russ).
7. Gross BA, Du R. Natural history of cerebral arteriovenous malformations: a meta-analysis. *J. Neurosurg*. 2013;118(2):437–443. doi: 10.3171/2012.10.JNS121280
8. Kearns KN, Chen CJ, Tvrdik P, et al. Outcomes of surgery for brainstem cavernous malformations: a systematic review. *Stroke*. 2019;50(10):2964–2966. doi: 10.1161/STROKEAHA.119.026120
9. Sazonov IA, Belousova OB. Cavernous malformation, which caused the development of extensive acute subdural hematoma. Case study and literature review. *Questions of neurosurgery named after N.N. Burdenko*. 2019;3(3):73–76. (In Russ). doi: 10.17116/neiro20198303173
10. Mouchtouris N, Chalouhi N, Chitale A, et al. Management of cerebral cavernous malformations: from diagnosis to

- treatment. *Scientific World Journal*. 2015;2015:808314. doi: 10.1155/2015/808314
11. Negoto T, Terachi S, Baba Y, et al. Symptomatic brainstem cavernoma of elderly patients: timing and strategy of surgical treatment. Two case reports and review of the literature. *World Neurosurg*. 2018;111:227–234. doi: 10.1016/j.wneu.2017.12.111
 12. Batra S, Lin D, Recinos PF, et al. Cavernous malformations: natural history, diagnosis and treatment. *Nature Reviews Neurology*. 2009;5(12):659–670. doi: 10.1038/nrneurol.2009.177
 13. Al-Shahi Salman R, White PM, Counsell CE, et al. Outcome after conservative management or intervention for unruptured brain arteriovenous malformations. *JAMA*. 2014;311(16):1661–1669. doi: 10.1001/jama.2014.3200
 14. Hakim AA, Gralla J, Rozeik C, et al. Anomalies and normal variants of the cerebral arterial supply: A comprehensive pictorial review with a proposed workflow for classification and significance. *J Neuroimaging*. 2018;28(1):14–35. doi: 10.1111/jon.12475
 15. Polkovnikov AY. Arteriovenous malformations of the cerebral hemispheres with a small and medium-sized nucleus, features of clinical manifestations and methods of surgical treatment. *Current Issues of Pharmaceutical and Medical Science and Practice*. 2013;2(12):033–038. (In Russ).
 16. Chang EF, Wang DD, Barkovich A, et al. Predictors of seizure freedom after surgery for malformations of cortical development. *Ann Neurol*. 2011;70(1):151–162. doi: 10.1002/ana.22399
 17. Pasqualin A, Meneghelli P, Giammarusti A, Turazzi S. Results of surgery for cavernomas in critical supratentorial areas. *Acta Neurochir Suppl*. 2014;119:117–123. doi: 10.1007/978-3-319-02411-0_20
 18. Maryashev SA. Stereotactic irradiation of arteriovenous malformations of the brain [dissertation abstract]. Moscow; 2016. 26 p. (In Russ).
 19. Shimizu K, Kosaka N, Yamamoto T, et al. Arterial spin labeling perfusion-weighted MRI for long-term follow-up of a cerebral arteriovenous malformation after stereotactic radiosurgery. *Acta Radiol Short Rep*. 2014;3(1). doi: 10.1177/2047981613510160
 20. Seo Y, Kim DG, Dho YS, et al. A retrospective analysis of the outcomes of dural arteriovenous fistulas treated with gamma knife radiosurgery: a single-institution experience. *Stereotactic Functional Neurosurgery*. 2018;96(1):46–53. doi: 10.1159/000486685
 21. Lee CC, Wang WH, Yang HC, et al. Gamma Knife radiosurgery for cerebral cavernous malformation. *Sci Rep*. 2019;9(1):19743. doi: 10.1038/s41598-019-56119-1
 22. Kefeli AU, Sengoz M, Peker S. Gamma knife radiosurgery for hemorrhagic brainstem cavernomas. *Turk Neurosurg*. 2018;29(1):14–19. doi: 10.5137/1019-5149.JTN.21690-17.1
 23. Cantu C, Murillo-Bonilla L, Arauz A, et al. Predictive factors for intracerebral hemorrhage in patients with cavernous angiomas. *Neurol Res*. 2005;27(3):314–318. doi: 10.1179/016164105X39914
 24. Wang CC, Liu A, Zhang J, et al. Surgical management of brainstem cavernous malformations: report of 137 cases. *Surg Neurol*. 2003;59(6):444–454. doi: 10.1016/s0090-3019(03)00187-3
 25. Abila AA, Turner JD, Mitha AP, et al. Surgical approaches to brainstem cavernous malformations. *Neurosurgical Focus*. 2010;29(3):E8. doi: 10.3171/2010.6.FOCUS10128
 26. Wen R, Shi Y, Gao Y, et al. The efficacy of gamma knife radiosurgery for cavernous malformations: a meta-analysis and review. *World Neurosurgery*. 2019;123:371–377. doi: 10.1016/j.wneu.2018.12.046
 27. Park SH, Hwang SK. Gamma knife radiosurgery for symptomatic brainstem intra-axial cavernous malformations. *World Neurosurgery*. 2013;80(6):261–266. doi: 10.1016/j.wneu.2012.09.013
 28. Frischer JM, Gatterbauer B, Holzer S, et al. Microsurgery and radiosurgery for brainstem cavernomas: effective and complementary treatment options. *World Neurosurgery*. 2014;81(3-4):520–528. doi: 10.1016/j.wneu.2014.01.004
 29. Kondziolka D, Lunsford LD, Flickinger JC, Kestle JR. Reduction of hemorrhage risk after stereotactic radiosurgery for cavernous malformations. *J Neurosurg*. 1995;83(5):825–831. doi: 10.3171/jns.1995.83.5.0825
 30. Aboukai R, Estrade L, Devos P, et al. Gamma knife radiosurgery of brainstem cavernous malformations. *Stereotact Funct Neurosurg*. 2016;94(6):397–403. doi: 10.1159/000452844
 31. Lopez-Serrano R, Martinez NE, Kusak ME, et al. Significant hemorrhage rate reduction after gamma knife radiosurgery in symptomatic cavernous malformations: long-term outcome in 95 case series and literature review. *Stereotact Funct Neurosurg*. 2017;95(6):369–378. doi: 10.1159/000480664
 32. Shin SS, Murdoch G, Hamilton RL, et al. Pathological response of cavernous malformations following radiosurgery. *J Neurosurg*. 2015;123(4):938–944. doi: 10.3171/2014.10.jns14499
 33. Gewirtz RJ, Steinberg GK, Crowley R, Levy RP. Pathological changes in surgically resected angiographically occult vascular malformations after radiation. *Neurosurgery*. 1998;42(4):738–741. doi: 10.1097/00006123-199804000-00031
 34. Nyary I, Major O, Hanzely Z, Szeifert GT. Histopathological findings in a surgically resected thalamic cavernous hemangioma 1 year after 40-gy irradiation. *J Neurosurg*. 2005;102(Suppl):56–58. doi: 10.3171/sup.2005.102.s_supplement.0056
 35. Monaco EA, Khan AA, Niranjana A, et al. Stereotactic radiosurgery for the treatment of symptomatic brainstem cavernous malformations. *Neurosurgical Focus*. 2010;29(3):E11. doi: 10.3171/2010.7.FOCUS10151
 36. Tian KB, Zheng JJ, Ma JP, et al. Clinical course of untreated thalamic cavernous malformations: hemorrhage risk and neurological outcomes. *J Neurosurg*. 2017;127(3):480–491. doi: 10.3171/2016.8.JNS16934
 37. Kida Y. Radiosurgery for cavernous malformations in basal ganglia, thalamus and brainstem. *Prog Neurol Surg*. 2009;22:31–37. doi: 10.1159/000163380
 38. Kim BS, Yeon JY, Kim JS, et al. Gamma Knife radiosurgery of the symptomatic brain stem cavernous angioma with low marginal dose. *Clin Neurol Neurosurg*. 2014;126:110–114. doi: 10.1016/j.clineuro.2014.08.028
 39. Menzler K. Epileptogenicity of cavernomas depends on (archi-) cortical localization. *Neurosurgery*. 2010;67:918–924. doi: 10.1227/NEU.0b013e3181eb5032
 40. Liao C, Visocchi M, Zhang W, et al. Management of cerebral radiation necrosis: a retrospective study of 12 patients. *Acta Neurochir Suppl*. 2017;124:195–201. doi: 10.1007/978-3-319-39546-3_30
 41. Akers A, Salman RA, Awad I, et al. Synopsis of guidelines for the clinical management of cerebral cavernous malformations: consensus recommendations based on systematic literature review by the angioma alliance scientific advisory board clinical experts panel. *Neurosurgery*. 2017;80(5):665–680. doi: 10.1093/neuros/nyx091

42. Garcia RM, Ivan ME, Lawton MT. Brainstem cavernous malformations: surgical results in 104 patients and a proposed grading system to predict neurological outcomes. *Neurosurgery*. 2015;76(3):265–277. doi: 10.1227/NEU.0000000000000602
43. Krylov VV, Dashyan VG, Shetova IM, et al. Neurosurgical care for patients with cerebral vascular diseases in the Russian Federation. *Neurosurgery*. 2017;(4):11–20. (In Russ).
44. Willie JT, Malcolm JG, Stern MA, et al. Safety and effectiveness of stereotactic laser ablation for epileptogenic cerebral cavernous malformations. *Epilepsia*. 2019;60(2):220–232. doi: 10.1111/epi.14634
45. Amponsah K, Ellis TL, Chan MD, et al. Retrospective analysis of imaging techniques for treatment planning and monitoring of obliteration for gamma knife treatment of cerebral arteriovenous malformation. *Neurosurgery*. 2012;71(4):893–899. doi: 10.1227/neu.0b013e3182672a83
46. Chen CC, Chapman P, Petit J, et al. Proton radiosurgery in neurosurgery. *Neurosurg Focus*. 2007;23(6):E5. doi: 10.3171/FOC-07/12/E5
47. Medical and Technical Complex: Joint Institute for Nuclear Research. Cavernous angiomas (cavernomas) [Internet]. Available from: http://mtk.jinr.ru/index.php?option=com_content&task=view&id=33&Itemid=53

ОБ АВТОРАХ

*** Гиря Елена Николаевна,**

адрес: Россия, 129010, Москва, Большая Сухаревская пл., д. 3;
тел.: +7 (495) 608-34-50;
ORCID: <https://orcid.org/0000-0001-5875-1489>;
eLibrary SPIN: 4793-7748; e-mail: mishka_77@list.ru

Синицын Валентин Евгеньевич, д.м.н., профессор;

ORCID: <https://orcid.org/0000-0002-5649-2193>;
eLibrary SPIN: 8449-6590; e-mail: vsini@mail.ru

Токарев Алексей Сергеевич, к.м.н.;

ORCID: <https://orcid.com/0000-0002-8415-5602>;
eLibrary SPIN: 1608-0630; e-mail: alex_am_00@mail.ru

AUTHORS' INFO

*** Elena N. Giryа, MD;**

address: 3 Bol'shaya Sukharevskaya ploshcad, 129010, Moscow, Russia; tel.: +7 (495) 608-34-50;
ORCID: <https://orcid.org/0000-0001-5875-1489>;
eLibrary SPIN: 4793-7748; e-mail: mishka_77@list.ru

Valentin E. Sinitsyn, MD, Dr. Sci. (Med.), Professor;

ORCID: <https://orcid.org/0000-0002-5649-2193>;
eLibrary SPIN: 8449-6590; e-mail: vsini@mail.ru

Alexey S. Tokarev, MD, Cand. Sci. (Med.);

ORCID: <https://orcid.com/0000-0002-8415-5602>;
eLibrary SPIN: 1608-0630; e-mail: alex_am_00@mail.ru

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD71630>

Проблема утомления в лучевой диагностике: многообещающая область для научных исследований

S. Taylor-Phillips, C. Stinton

Warwick Medical School, The University of Warwick, Ковентри, Великобритания

АННОТАЦИЯ

Утомление врачей-рентгенологов может быть причиной большого количества врачебных ошибок. В данном обзоре описываются последние исследования, посвящённые изучению проблемы утомления в области лучевой диагностики. Мы приводим данные о методах измерений и влиянии утомления на точность результатов, получаемых в ходе лабораторных исследований и в условиях клинической практики. В статье раскрываются широкие возможности для будущих исследований в этой области, включая дополнительные меры, способствующие уменьшению количества ошибок, связанных с утомлением, и дальнейшее понимание того, каким образом показатели утомления соотносятся с возможными врачебными ошибками. Изучается возможность ответа на эти вопросы с помощью крупных популяционных наблюдательных исследований и комплексных рандомизированных контролируемых испытаний, проводимых в условиях реальной клинической практики.

Данная публикация является перепечатанной версией статьи [Taylor-Phillips S, Stinton C. Fatigue in radiology: a fertile area for future research. *Br J Radiol.* 2019;92:20190043. doi: 10.1259/bjr.20190043] с переводом оригинала на русский язык.

Ключевые слова: утомление; врачебные ошибки; лучевая диагностика; врачи-рентгенологи.

Как цитировать

Taylor-Phillips S., Stinton C. Проблема утомления в радиологии: многообещающая область для научных исследований // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 211–222. DOI: <https://doi.org/10.17816/DD71630>

DOI: <https://doi.org/10.17816/DD71630>

Fatigue in radiology: a fertile area for future research

Sian Taylor-Phillips, Chris Stinton

Warwick Medical School, The University of Warwick, Coventry, UK

ABSTRACT

Fatigue in radiologists may be responsible for a large number of medical errors. This review describes the latest research on fatigue in radiology. This includes measurement methods, and recent evidence on how fatigue affects accuracy in laboratory test conditions and in clinical practice. The extensive opportunities for future research in the area are explored, including testing interventions to reduce fatigue-related error, and further understanding of which fatigue measures correlate with errors. Finally we explore the possibility of answering these questions using large population-based observational studies and pragmatic integrated randomised controlled trials.

This publication is the reprint with Russian translation from original: Taylor-Phillips S, Stinton C. Fatigue in radiology: a fertile area for future research. *Br J Radiol.* 2019;92:20190043. doi: 10.1259/bjr.20190043.

Keywords: fatigue; medical errors; radiology; radiologists.

To cite this article

Taylor-Phillips S, Stinton C. Fatigue in radiology: a fertile area for future research. *Digital Diagnostics.* 2021;2(2):211–222.

DOI: <https://doi.org/10.17816/DD71630>

Received: 17.06.2021

Accepted: 17.06.2021

Published: 07.07.2021

DOI: <https://doi.org/10.17816/DD71630>

辐射诊断中的疲劳问题：一个很有前途的研究领域

Sian Taylor-Phillips, Chris Stinton

Warwick Medical School, The University of Warwick, Coventry, UK

简评

放射科医生的疲劳会导致大量的医疗错误。本文综述了辐射诊断领域中疲劳的最新研究进展。我们提供有关测量方法的数据以及疲劳对实验室和临床结果准确性的影响。文章揭示了该领域未来研究的广阔空间，包括有助于减少疲劳误差的附加措施，并进一步了解疲劳指标与可能出现的医疗失误的关系。研究了通过大型种群观察研究和综合随机对照试验回答这些问题的可行性，在真正的临床实践中进行。

本出版物是该文章的重印版本[Taylor-Phillips S, Stinton C. Fatigue in radiology: a fertile area for future research. *Br J Radiol* 2019;92:20190043. doi: 10.1259/bjr.20190043] 翻译成俄语。

关键词： 疲劳； 医疗差错； 放射诊断； 放射科医生。

引用本文：

Taylor-Phillips S, Stinton C. Fatigue in radiology: a fertile area for future research. *Digital Diagnostics*. 2021;2(2):211–222.

DOI: <https://doi.org/10.17816/DD71630>

收到: 17.06.2021

接受: 17.06.2021

发布日期: 07.07.2021

ВВЕДЕНИЕ

Утомление медицинских работников отрицательно сказывается на их самочувствии и моральном состоянии и может приводить к врачебным ошибкам. Данный обзор посвящён изучению проблемы утомления, которое может стать причиной врачебных ошибок в области лучевой диагностики. Ежегодно в больницах скорой медицинской помощи Национальной службы здравоохранения Великобритании (National Health Service, NHS) может происходить более 11 000 предотвратимых смертей, из которых на диагностические ошибки приходится более 8000 [1]. Эти приблизительные данные основаны на предыдущих оценках, согласно которым тяжёлая форма инвалидности или смерть в результате медицинских вмешательств наступает у 60 000–255 000 пациентов, находящихся в лечебных учреждениях NHS [2]. Какова доля случаев, которые могут быть связаны с утомлением, доподлинно неизвестно. Тем не менее мы точно знаем, что в конце рабочей смены врач-рентгенолог может допускать большее количество ошибок, чем в начале дня, и что более длительные смены в медицине связаны с более высокими показателями врачебных ошибок и травм [3–5]. Проблема усугубляется тем, что загруженность служб лучевой диагностики в последнее время резко возросла (по оценкам, до 10–12% в год) без соответствующего при этом увеличения численности персонала [6].

Несмотря на то, что утомление в лучевой диагностике может иметь крайне негативное значение, исследований в этой области крайне мало. В данной статье мы рассмотрим различные типы и классификации утомления, приведём текущую доказательную базу с использованием данных лабораторных тестов, наблюдательных исследований и испытаний, выполненных в условиях реальной клинической практики, а также обсудим направления будущих исследований.

ИЗМЕРЕНИЕ УТОМЛЕНИЯ

Утомление обычно описывается как чувство «слабости, вялости и желания отдохнуть», сопровождаемых пониженной работоспособностью [7, 8], и подразделяется, как правило, на два связанных между собой подтипа — физическое и умственное. Физическое утомление — это ухудшение способности мышц поддерживать мышечное сокращение, что приводит к трудностям с мышечной координацией и управлением двигательным аппаратом. Умственное утомление — это снижение способности выполнять умственные задачи [9]. Применительно к лучевой диагностике существует ещё один подтип — утомление глаз, связанное со снижением остроты зрения. Измерение утомления может быть объективным или субъективным, и каждый метод измерения может охватывать одну или несколько из этих категорий.

Для измерения утомления разработан ряд инструментов. Следует отметить, что некоторые показатели, используемые для оценки степени утомления, были разработаны для измерения усталости. «Утомление» и «усталость» (или сонливость) в литературе часто употребляются как взаимозаменяемые понятия [8, 10, 11]. В основе большинства оценочных инструментов лежит субъективная оценка, полученная с помощью шкал или контрольных перечней, составленных со слов пациентов. Так, A. Shahid и соавт. [8] описали краткосрочные (например, Каролинская и Стэнфордская шкалы сонливости [12, 13]) и общие (например, Эпвортская шкала сонливости [14], Опросник о качестве сна и пробуждения [15]) показатели сонливости. E. Krupinski и B.I. Reiner [10] сообщили о некоторых дополнительных способах оценки сонливости, включая Питтсбургский опросник для определения индекса качества сна [16] и Шведский опросник для определения профессиональной утомляемости (Swedish Occupational Fatigue Inventory, SOFI) [17].

Помимо субъективных, также были описаны объективные методы измерения утомления, например тест на поддержание бодрствования (оценка концентрации внимания в течение дня, в ходе которой участникам предлагают поддерживать бодрствование в затемнённом помещении), тест на определение психомоторной бдительности (оценка способности длительно удерживать произвольное внимание, где участникам предлагают реагировать на стимулы нажатием кнопки, после чего оценивается наличие/отсутствие реакции), а также тест на устойчивость внимания (ещё один метод измерения внимательности, когда участники выполняют нестимулирующие задачи, при этом реагируя на одни стимулы и игнорируя другие). В других работах описаны альтернативные подходы к измерению утомления, такие как мониторинг кровяного давления, кожно-гальванические реакции и частота сердечных сокращений [10, 11], и подчеркнута важность зрительного утомления при диагностической визуализации.

Интерпретация медицинских снимков является сложной и повторяющейся задачей визуального поиска, и ошибки (ложноотрицательные результаты) здесь встречаются относительно часто [18]. Помимо признаков, о которых сообщают сами участники, изучению зрительного утомления поможет оценка аккомодации (изменение кривизны хрусталика глаза для регулировки фокуса между объектами вдали и вблизи) и конвергенции (формирование/поддержание бинокулярного зрения путём движения глаз в противоположных направлениях). Способность к аккомодации и конвергенции снижается при утомлении [10, 11]. Другие исследователи предлагают оценивать в качестве объективных показателей зрительного утомления критическую частоту слияния (частота, при которой мерцающий свет воспринимается как стабильный) и частоту мигательных движений [19, 20]. Из всех этих показателей лишь немногие

учитывались при изучении утомления в области диагностической визуализации.

Многие исследователи пришли к разумному консенсусу относительно ряда объективных и субъективных инструментов, которые подходят для измерения утомления, тем не менее в одиночных работах прослеживаются попытки соотнести баллы, полученные по этим показателям, с клинически значимыми результатами в лучевой диагностике, такими как частота выявления патологии или частота возникновения ошибок. Не владея данными о взаимосвязи этих показателей с клиническими исходами, авторы исследований с меньшей вероятностью смогут добиться результатов, которые будут эффективными в клинической практике.

Хотя для улучшения общего состояния рентгенологов важно минимизировать утомление, понимание того, каким образом это состояние может привести к пропуску клинически значимых отклонений, может иметь гораздо большее значение. Именно поэтому в качестве альтернативного и объективного метода измерения утомления часто выступает оценка показателей точности тестов, измеряемых непрерывно во время сеанса чтения рентгенограмм или в течение всего рабочего дня врача-рентгенолога. Так, во многих исследованиях по психологии, где изучается способность молодых специалистов к устойчивому вниманию при визуальном поиске и выполнении задач на проверку бдительности, такие тесты длятся около часа [21]. Актуальность данная область исследований приобрела в связи с выявлением фактов снижения чувствительности у операторов радиолокационных станций к обнаружению вражеских самолётов или кораблей на экранах радаров в процессе выполнения задачи (декремент бдительности) [22]. Чтение рентгенологических данных считается аналогичной задачей с высокой интенсивностью, низким уровнем сигнала и схожими закономерностями [23], хотя различия в опыте и знаниях, самой задаче и её важности могут ограничивать обобщаемость полученных результатов.

В психологии также хорошо известен эффект пространённости, когда чувствительность рентгенологов к выявлению отклонений снижается при изучении серии снимков с меньшей распространённостью таких отклонений [24], что может быть обусловлено ожиданиями лучевых диагностов относительно возможной вероятности аномальности случая, до того как он будет рассмотрен. Это в свою очередь приводит к снижению чувствительности в процессе выполнения задачи при интерпретации рентгенологических данных с низкой распространённостью [25]. В работе E. Kompaniez-Dunigan и соавт. [26] по изучению зрительной адаптации при описании рентгеновских снимков говорится о необходимости дальнейшего изучения вопроса о её влиянии на точность интерпретируемых результатов в процессе выполнения задачи.

Имеются также результаты ряда исследований о циркадных ритмах (биологические процессы, которые ритмично изменяются на протяжении примерно 24 ч). Кроме того, в лабораторных условиях наблюдались связанные со временем изменения в работоспособности при выполнении ряда задач на проверку двигательных и зрительных навыков, чувствительности, вербальных задач и задач на запоминание [27].

В целом работоспособность при выполнении задач улучшается в течение дня и снижается ночью, хотя и с некоторыми вариациями [27]. Приведены также некоторые свидетельства снижения работоспособности после обеденного приёма пищи [28]. Помимо этого, выдвинута гипотеза, что циркадные ритмы влияют на точность рентгенологических исследований: некоторые авторы сообщали, что точность диагностики снижается при интерпретации снимков в конце рабочего дня, чего не наблюдается в начале рабочей смены [29, 30]. Другие исследования этого не подтверждают [31, 32].

ИССЛЕДОВАНИЯ, ПРОВЕДЁННЫЕ В ЛАБОРАТОРНЫХ УСЛОВИЯХ

Исследования по изучению влияния утомления на точность диагностической визуализации проводятся редко. N. Stec и соавт. [33] представили систематический обзор научных работ по данной теме, опубликованных до января 2017 г., выявив 27 соответствующих исследований, из которых менее половины были первичными ($n=10$) или «другими» типами.

E. Krupinski и соавт. провели серию экспериментальных исследований [29, 30, 34]. В первом исследовании оценивали точность выявления переломов костей в работе 40 интерпретаторов (20 врачей-рентгенологов и 20 ординаторов), которые изучали серию из 60 случаев по 2–4 снимка в каждом в начале обычного дня клинической практики (раннее состояние) и после его завершения (позднее состояние) [29]. Общее утомление оценивали с помощью опросника SOFI, а зрительное утомление (напряжение глаз) — с помощью автоматического рефрактометра WAM-5500 (Grand Seiko, Хиросима, Япония), который собирает данные о диаметре зрачка и показателях преломления. Авторы обнаружили, что субъективных сообщений об утомлении и напряжении глаз было значительно больше в «поздней» временной точке по сравнению с «ранней», и что диагностическая точность была значительно более низкой после завершения дня клинической практики, чем до его начала (площадь под кривой в «ранней» точке составляла 0,885, в «поздней» — 0,852). Однако непосредственную оценку связи между показателями утомления и выполнением задач не проводили.

Во втором исследовании 44 интерпретатора (22 практикующих врача-рентгенолога и 22 ординатора-рентгенолога) выявляли лёгочные узлы при компьютерной

томографии (КТ) грудной клетки [30]. Как и в предыдущем исследовании, работоспособность врачей проверяли до и после одного дня клинической практики. Общее утомление оценивали с помощью опросника SOFI, а зрительное — методом темновой конвергенции (конвергенция глаз в отсутствие стимулов). Полученные результаты были аналогичны результатам предыдущего исследования [29], т.е. диагностическая точность у ординаторов была значительно более низкой в конце дня по сравнению с его началом (79% против 75%). У практикующих рентгенологов такой разницы не наблюдалось. Общее и зрительное утомление не всегда было выше в группе «позднего времени». Степень утомления (т.е. баллы, согласно опроснику на определение утомления или степень зрительного напряжения) и выполнение задач напрямую не оценивались.

В третьем исследовании 20 рентгенологов изучали результаты КТ у пациентов со множественными травмами после 8-часового рабочего дня (так называемая группа с утомлением). Утомление измеряли в соответствии с исследованием [29]. Данные участников в состоянии утомления сопоставляли с данными 20 лиц без утомления, принимавших участие в предыдущих исследованиях той же исследовательской группы [29, 30]. При выявлении крупных переломов не было выявлено разницы между группой с утомлением и группой без утомления (площадь под ROC-кривой составляла 0,945 против 0,944 соответственно). Опять же, непосредственная оценка связи между уровнем утомления и выполнением задач не проводилась. Эти исследования дают некоторое представление о том, что квалификация и опыт рентгенолога могут влиять на степень испытываемого утомления и/или умеренно влиять на связь между утомлением и точностью визуализации, что заслуживает дальнейшего изучения.

T.N. Happa и соавт. [35] изучали влияние ночных смен на диагностическую точность и степень утомления на основе данных 12 рентгенологов (5 преподавателей и 7 ординаторов). Участники анализировали рентгенограммы костей в двух состояниях — один раз днём (состояние без утомления) и один раз утром после ночной смены (состояние с утомлением) — и заполняли опросник SOFI. Между этими двумя состояниями обнаружены существенные различия: более высокие показатели утомления по данным каждого из разделов опросника и более низкая диагностическая точность (0,806 против 0,926) после ночной смены.

Результаты всей серии экспериментальных исследований подтверждены выводом в обзоре N. Stec и соавт. [33], что зрительное утомление встречается у рентгенологов относительно часто (примерно 35% участников сообщали о напряжении глаз), и что зрительное и физическое утомление связаны с более низкой диагностической точностью (согласно оценке работоспособности на протяжении длительного периода времени).

ИССЛЕДОВАНИЯ, ПРОВЕДЁННЫЕ В УСЛОВИЯХ РЕАЛЬНОЙ КЛИНИЧЕСКОЙ ПРАКТИКИ

Хотя большинство данных о влиянии утомления на работоспособность при диагностической визуализации получено из экспериментальных исследований, выполнено также небольшое количество клинических испытаний, что особенно ценно для понимания важности утомления в «реальных условиях», поскольку предыдущие данные свидетельствуют о различиях между работоспособностью при выполнении экспериментальных задач и работоспособностью, наблюдаемой при визуализации в клинической практике [36, 37].

Ruutiainen и соавт. сравнили различия между предварительной и окончательной интерпретацией 8062 медицинских снимков, выполненной 10 ординаторами-рентгенологами, которые работали на протяжении долгого времени (более 10 ч подряд). Авторы обнаружили, что частота серьёзных несоответствий (которые потенциально могли повлиять на пациентов) была значительно более высокой в отчётах за последние 2 ч рабочей смены по сравнению с более ранним временем (2% по сравнению с 1%).

T.N. Happa и соавт. [5] выполнили ретроспективный анализ влияния продолжительности смены, графика работы (обычные дни, праздники, выходные, сверхурочные, резервное время работы) и рабочей нагрузки на точность интерпретации в большой выборке рентгенологических исследований ($n=2\,922\,377$), охватывающей широкий спектр медицинских специальностей. Исследователи обнаружили, что продолжительные смены и большие объёмы работы (что приводит к утомлению) стали причиной значительного числа серьёзных различий в интерпретации снимков. Кроме того, большая их часть приходилась именно на вторую половину рабочего дня.

Насколько нам известно, единственное из проведённых проспективных исследований по изучению степени утомления при диагностической визуализации — это масштабное рандомизированное контролируемое исследование CO-OPS (Changing case Order to Optimise patterns of performance in Screening), в которое были включены данные более миллиона женщин [38]. S. Taylor-Phillips и соавт. [38] изучили влияние снижения бдительности на работоспособность интерпретаторов, работающих в парах (рентгенологов, опытных рентгенологов и врачей-маммологов), которые исследовали маммограммы женщин в рамках Программы скрининга рака молочной железы (Великобритания) [38, 39]. Интерпретаторы последовательно оценивали в парах серии снимков, каждая из которых включала примерно по 35 маммограмм. Участников исследования разделили на 2 группы: в первой специалисты-аналитики в парах оценивали серии снимков в одинаковом порядке (контрольная группа),

во второй — один из пары интерпретаторов оценивал снимки в прямой последовательности, второй — в обратной (интервенционная группа). При снижении бдительности частота выявления рака уменьшалась бы со временем при выполнении задачи в контрольной группе (к концу анализа оба интерпретатора были бы в состоянии низкой бдительности), но не в интервенционной группе (интерпретаторы были бы в состоянии низкой бдительности на разных этапах исследования снимков). Тем не менее не наблюдалось никакой разницы в показателях выявления рака между двумя группами (ОР 1,01; 95% ДИ 0,97–1,06), что указывает на отсутствие снижения бдительности. Интересно, что количество повторных вызовов пациентов со временем уменьшалось, что фактически говорит об увеличении работоспособности специалиста-аналитика со временем выполнения задачи, с точки зрения положительной прогностической ценности.

C. Stinton и соавт. [39] в дальнейшем проанализировали ряд данных, использованных в рамках исследования CO-OPS. Авторы изучили показатели повторных вызовов и выявления рака интерпретаторами в течение дня, при этом данные были разделены на три равных промежутка времени, исходя из продолжительности изучения маммограмм: 9:00–16:59, 17:00–23:59 и 1:00–8:59. В течение дня наблюдались различия как в показателях повторных вызовов, так и в показателях выявления рака, при этом многоуровневая логистическая регрессия показала, что вероятность повторного вызова у женщин, чьи маммограммы были интерпретированы в период с 17:00 до 23:59, была в 1,07 раза выше (95% ДИ 1,03–1,11), чем у женщин, чьи маммограммы исследовали в течение других периодов времени ($p < 0,001$). Значимой связи между временным периодом и частотой выявления рака выявлено не было. Такая тенденция наблюдалась при анализе данных с учётом всех задействованных интерпретаторов и без учёта тех специалистов-аналитиков, которые не работали в вечернее время.

ВОЗМОЖНОСТИ ДЛЯ БУДУЩИХ ИССЛЕДОВАНИЙ

Существуют убедительные доказательства того, что утомление — это серьёзная проблема, которая влияет на количество врачебных ошибок в медицине в целом, и лучевая диагностика не является исключением. Требуются дальнейшие исследования, чтобы понять, с помощью каких мер можно снизить утомляемость рентгенологов и, в свою очередь, уменьшить количество последующих врачебных ошибок. Кроме того, существует множество субъективных и объективных показателей утомления, но для того, чтобы определить, какие из них связаны с врачебными ошибками, вызываемыми утомлением, необходимы опять же соответствующие исследования. Эти данные могут быть использованы

для выбора метода измерения утомления, что позволит выявлять состояние утомления, в результате которого допускаются врачебные ошибки в режиме реального времени. В других областях лучевой диагностики используемые технологические достижения, связанные с искусственным интеллектом, также могут влиять на некоторые элементы выполняемых рентгенологами задач, включая и те, которые приводят к утомлению. Следовательно, эти связи также требуют дальнейшего изучения. Наконец, появление «больших данных» даёт возможность ответить на поставленные вопросы в крупномасштабных наблюдательных и комплексных исследованиях в условиях реальной клинической практики с помощью методов, которые ранее не были широко доступны.

Исследований, изучающих эффективность мер по снижению утомления и ошибок, связанных с этим состоянием, в лучевой диагностике немного. В числе самых крупных из проведённых исследований следует отметить рандомизированное контролируемое исследование по изменению порядка изучения данных при скрининге рака молочной железы, подробно описанное выше [38]. Данное исследование показало, что используемые методы снижения утомляемости не были эффективными. Эффективность других потенциальных мер, предпринимаемых во время одного сеанса чтения, таких как планирование перерывов, прерывания, кофеин, дизайн рабочего места и окружающее освещение, ещё предстоит тщательно проверить в реальных условиях. При этом исследование степени утомления к концу длительной рабочей смены будет более предпочтительным, поскольку к этому времени его показатели будут значительно превышать те, которые наблюдаются на протяжении единичного сеанса чтения рентгенологических данных. Предпринимаемые меры в этом контексте могут быть более организационными: например, составление графика выполнения задач в течение смены таким образом, чтобы критически важные задачи выполнялись ближе к её началу; изменение продолжительности смены, времени начала и конца смены или численности персонала. Все подобные вмешательства должны быть направлены на то, чтобы помочь рентгенологам справиться с утомлением, а не на применение административного подхода, выполняемого в директивном порядке, поэтому дизайн исследований должен быть прагматичным и включать адаптацию к местным условиям.

Мы не знаем, какие показатели утомления связаны с увеличением числа врачебных ошибок. Требуются масштабные исследования с одновременным применением нескольких простых методов измерения утомления у одних и тех же рентгенологов, в одно и то же время и соотносением их с реальной частотой возникновения ошибок. Эти данные позволят определить, какие показатели могут служить косвенными признаками состояния утомления, приводящего к врачебным ошибкам. Затем эти же показатели можно будет использовать в качестве

косвенных результатов в научных исследованиях и полезных механизмов обратной связи в клинической практике.

Появление искусственного интеллекта, исследующего рентгеновские снимки, не является чем-то новым: это, скорее, автоматизированное обнаружение данных, которое применялось в течение многих лет с переменным успехом. Новым здесь является то, что разработки в области хранения данных, определения взаимосвязей и применения искусственного интеллекта привели к появлению нового поколения автоматизированных инструментов чтения изображений. Точность этих инструментов не является предметом данного обзора, но точность как таковая, по-видимому, зависит от доступа к большому количеству изображений с известной достоверностью данных, которые становятся всё более доступными. Важное значение приобретает понимание того, каким образом эти разработки связаны с понятием утомления в лучевой диагностике. Исследования утомления показали, что при некоторых обстоятельствах специфичность и положительная прогностическая значимость улучшаются в ходе сеанса чтения [40], что позволяет объяснить действие механизма, с помощью которого чтение серии снимков улучшает специфичность [41]. Таким образом, с учётом исследований человеческого фактора и утомления можно предположить, что если искусственный интеллект будет использоваться для отсеивания очевидных случаев без какой-либо патологии, то число сложных случаев, интерпретируемых рентгенологами, возрастёт, что может повысить утомляемость, но в то же время увеличит уровень распространённости и, следовательно, чувствительности [23, 24]. Если искусственный интеллект использовать для обнаружения подозрительных участков на рентгеновских снимках, сеанс чтения может прерываться ложноположительными подсказками, что негативно скажется на специфичности, как было в случае с автоматизированным обнаружением данных при скрининге молочной железы [42, 43].

Проблема измерения утомления у врачей-рентгенологов — это проблема статистической мощности, общей для исследователей, измеряющих все элементы точности и работоспособности специалистов этой области. Мы хотели бы знать, какое оборудование, рабочие места, часы работы, условия и тесты позволяют рентгенологам быть максимально точными. Мы хотели бы, чтобы результаты наших исследований были клинически значимыми, например измерение числа клинически значимых пропущенных случаев/отклонений от нормы. Однако для этого потребуется выполнение либо серии тестов с нереальным количеством и типами отклонений, либо проведение очень крупных исследований. Похоже, что будущее именно за крупными исследованиями, поскольку достоверно известно, что состояние врачей-рентгенологов, интерпретирующих расширенный набор

данных, не поддаётся обобщению в клинической практике [36, 37].

Эпоха «больших данных» представляет собой большой потенциал и риск для исследований. Автоматизированные системы ведения истории болезни пациентов и обработки изображений открывают фантастические возможности для исследований, которые находятся лишь на начальном этапе реализации. Программа скрининга рака молочной железы в Великобритании — яркий тому пример. Эта программа включает единую автоматизированную систему ведения медицинских карт пациентов, поэтому учёт данных осуществляется примерно одинаково по всей стране. Эта же программа подключена к Национальному реестру онкологических заболеваний, где содержатся данные долгосрочного наблюдения за пациентами, что позволило провести множество наблюдательных исследований. В последнее время в программу были внесены изменения для интеграции рандомизации. Внутренняя рандомизация и наблюдение за клинически значимыми исходами позволяет нам автоматизировать проведение рандомизированных контролируемых исследований, таких как CO-OPS с участием миллиона женщин [38] и исследование риска заболевания раком молочной железы, возрастающего с возрастом, с участием 6 млн женщин [44]. Компьютеризация данных позволяет проводить гораздо более крупные исследования при значительно меньших затратах. Интегрированные рандомизированные контролируемые испытания, проводимые в реальных условиях клинической практики, обеспечивают наименее предвзятую форму исследования утомления. В тех случаях, когда рандомизация невозможна или нецелесообразна, например при сравнении дизайнов кабинетов врачей-рентгенологов, более предпочтительными будут наблюдательные исследования.

При переходе к регулярному использованию «больших данных» в рентгенологических исследованиях и аудите существует реальный риск того, что некорректный анализ приведёт к неверным выводам. В настоящее время в лучевой диагностике имеется огромное количество как самих изображений, так и связанных с ними данных о пациентах. Наблюдательные исследования строятся на большом количестве вмешивающихся факторов, ограничивающих их проведение, поэтому необходимо проявлять осторожность при формулировании выводов на их основе. Повышенная доступность больших объёмов выборки может привести к увеличению числа систематических ошибок, связанных с предпочтительной публикацией положительных результатов исследования, поскольку результаты публикуются только тогда, когда они «интересны». Например, в Великобритании анализ наблюдений повышенной смертности пациентов, поступающих в больницы в выходные дни, был использован в качестве обоснования семидневной рабочей недели [45,

46]. Однако основным показателем являлась повышенная тяжесть заболевания у пациентов, поступающих в лечебные учреждения в выходные дни, при этом сами авторы заявили, что «предположение о том, что их (дополнительных смертей в выходные дни — прим. авт.) можно было избежать, было бы опрометчивым и неправильным». Это один из наиболее очевидных примеров того, насколько трудно установить причину и следствие при наличии только данных о наблюдениях. Обучение методам исследования должно стать более доступным для тех специалистов, которые проверяют и анализируют такие данные, чтобы при анализе учитывались вмешивающиеся факторы во избежание некорректных выводов. В исследованиях утомления особое внимание следует уделять индивидуальным различиям между врачами-рентгенологами, например путём их включения в анализ в качестве случайного эффекта.

Успех будущих исследований в области утомления будет в значительной степени зависеть от культуры в рабочей среде врачей-рентгенологов. Данные о работе отдельного рентгенолога могут стать очень мощным инструментом для постоянного улучшения работоспособности специалистов данной области, если они в точности передаются в благоприятных для обучения условиях. Аналогичным образом, данные исследований могут быть полностью анонимными как в отношении пациентов, так и в отношении врачей-рентгенологов, что позволяет проводить крупные исследования с меньшим предполагаемым риском. Если «большие данные» будут использоваться в качестве метода управления работоспособностью в конфликтных ситуациях, в обвинительной практике или даже в судебных исках по поводу неправомерных действий, они будут восприниматься, скорее, как средство наказания и тотального контроля. В этом случае врачи-рентгенологи вряд ли будут изъяслять

готовность открыто работать с «большими данными», что способствовало бы обучению и совершенствованию необходимых навыков.

ЗАКЛЮЧЕНИЕ

Существуют явные доказательства того, что утомление, особенно в конце длительной рабочей смены, способствует возникновению серьёзных врачебных ошибок и увеличивает риск того, что патологические изменения при диагностической визуализации не будут выявлены. Для того чтобы определить, какие меры помогут снизить количество ошибок, связанных с утомлением, и какие показатели утомления являются надёжными ранними индикаторами этого состояния, необходимо провести дополнительные исследования. Использование «больших данных» в сфере здравоохранения даёт возможность проведения крупных исследований для решения этих проблем, но при этом особое внимание следует уделить устранению вмешивающихся факторов.

ДОПОЛНИТЕЛЬНО

Д-р S. Taylor-Phillips и д-р C. Stinton получили поддержку от Национального института исследований в области здравоохранения (NIHR CLAHRC) (Уэст-Мидлендс, Великобритания). Д-р S. Taylor-Phillips является участником программы стипендий и грантов «NIHR Career Development Fellowship» (CDF-2016-09-018). Мнения, изложенные в статье, принадлежат авторам рукописи. Национальная служба здравоохранения, Национальный институт исследований в области здравоохранения и Департамент здравоохранения и социального обеспечения Великобритании не влияли на написание рукописи.

СПИСОК ЛИТЕРАТУРЫ

1. Hogan H., Healey F., Neale G., et al. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study // *BMJ Qual Saf.* 2012. Vol. 21, N 9. P. 737–745. doi: 10.1136/bmjqs-2011-001159
2. Chief Medical Officer. An organisation with a memory: Report of an expert group on learning from adverse events in the NHS. London: The British Institute of Radiology; 2000.
3. Lockley S.W., Barger L.K., Ayas N.T., et al. Effects of health care provider work hours and sleep deprivation on safety and performance // *Jt Comm J Qual Patient Saf.* 2007. Vol. 33, N 11, Suppl. P. 7–18. doi: 10.1016/S1553-7250(07)33109-7
4. Barger L.K., Cade B.E., Ayas N.T., et al. Extended work shifts and the risk of motor vehicle crashes among interns // *N Engl J Med.* 2005. Vol. 352, N 2. P. 125–134. doi: 10.1056/NEJMoa041401
5. Hanna T.N., Lamoureux C., Krupinski E.A., et al. Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations // *Radiology.* 2018. Vol. 287, N 1. P. 205–212. doi: 10.1148/radiol.2017170555
6. Clinical Radiology UK workforce census 2014 report. London: The British Institute of Radiology, 2015.
7. Pigeon W.R., Sateia M.J., Ferguson R.J. Distinguishing between excessive daytime sleepiness and fatigue: toward improved detection and treatment // *J Psychosom Res.* 2003. Vol. 54, N 1. P. 61–69. doi: 10.1016/s0022-3999(02)00542-1
8. Shahid A., Shen J., Shapiro C.M. Measurements of sleepiness and fatigue // *J Psychosom Res.* 2010. Vol. 69, N 1. P. 81–89. doi: 10.1016/j.jpsychores.2010.04.001
9. Xu R., Zhang C., He F., et al. How physical activities affect mental fatigue based on EEG energy, connectivity, and complexity // *Front Neurol.* 2018. Vol. 9. P. 915. doi: 10.3389/fneur.2018.00915
10. Krupinski E., Reiner B.I. Real-time occupational stress and fatigue measurement in medical imaging practice // *J Digit Imaging.* 2012. Vol. 25, N 3. P. 319–324. doi: 10.1007/s10278-011-9439-1

11. Waite S., Kolla S., Jeudy J., et al. Tired in the reading room: the influence of fatigue in radiology // *J Am Coll Radiol*. 2017. Vol. 14, N 2. P. 191–197. doi: 10.1016/j.jacr.2016.10.009
12. Akerstedt T., Gillberg M. Subjective and objective sleepiness in the active individual // *Int J Neurosci*. 1990. Vol. 52, N 1–2. P. 29–37. doi: 10.3109/00207459008994241
13. Hoddes E., Zarcone V., Smythe H., et al. Quantification of sleepiness: a new approach // *Psychophysiology*. 1973. Vol. 10, N 4. P. 431–436. doi: 10.1111/j.1469-8986.1973.tb00801.x
14. Johns M.W. A new method for measuring daytime sleepiness: the epworth sleepiness scale // *Sleep*. 1991. Vol. 14, N 6. P. 540–545. doi: 10.1093/sleep/14.6.540
15. Rosenthal L., Roehrs T.A., Roth T. The sleep-wake activity inventory: a self-report measure of daytime sleepiness // *Biol Psychiatry*. 1993. Vol. 34, N 11. P. 810–820. doi: 10.1016/0006-3223(93)90070-T
16. Buysse D.J., Reynolds C.F., Monk T.H., et al. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research // *Psychiatry Res*. 1989. Vol. 28, N 2. P. 193–213. doi: 10.1016/0165-1781(89)90047-4
17. Ahsberg E. Dimensions of fatigue in different working populations // *Scand J Psychol*. 2000. Vol. 41, N 3. P. 231–241. doi: 10.1111/1467-9450.00192
18. Miglioretti D.L., Smith-Bindman R., Abraham L., et al. Radiologist characteristics associated with interpretive performance of diagnostic mammography // *J Natl Cancer Inst*. 2007. Vol. 99, N 24. P. 1854–1863. doi: 10.1093/jnci/djm238
19. Maeda E., Yoshikawa T., Hayashi N., et al. Radiology reading-caused fatigue and measurement of eye strain with critical flicker fusion frequency // *Jpn J Radiol*. 2011. Vol. 29, N 7. P. 483–487. doi: 10.1007/s11604-011-0585-7
20. Sheppard A.L., Wolffsohn J.S. Digital eye strain: prevalence, measurement and amelioration // *BMJ Open Ophthalmol*. 2018. Vol. 3, N 1. P. e000146. doi: 10.1136/bmjophth-2018-000146
21. Thomson D.R., Besner D., Smilek D. A critical examination of the evidence for sensitivity loss in modern vigilance tasks // *Psychol Rev*. 2016. Vol. 123, N 1. P. 70–83. doi: 10.1037/rev0000021
22. Mackworth N. Researches on the measurement of human performance. London, UK: The British Institute of Radiology; 1950.
23. See J.E., Howe S.R., Warm J.S., et al. Meta-analysis of the sensitivity decrement in vigilance // *Psychological Bulletin*. 1995. Vol. 117, N 2. P. 230–249. doi: 10.1037/0033-2909.117.2.230
24. Gur D., Rockette H.E., Armfield D.R., et al. Prevalence effect in a laboratory environment // *Radiology*. 2003. Vol. 228, N 1. P. 10–14. doi: 10.1148/radiol.2281020709
25. Evans K.K., Birdwell R.L., Wolfe J.M. If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening // *PLoS One*. 2013. Vol. 8, N 5. P. e64366. doi: 10.1371/journal.pone.0064366
26. Kompaniez-Dunigan E., Abbey C.K., Boone J.M., Webster M.A. Visual adaptation and the amplitude spectra of radiological images // *Cogn Res Princ Implic*. 2018. Vol. 3, N 1. P. 3. doi: 10.1186/s41235-018-0089-4
27. Blatter K., Cajochen C. Circadian rhythms in cognitive performance: methodological constraints, protocols, theoretical underpinnings // *Physiol Behav*. 2007. Vol. 90, N 2–3. P. 196–208. doi: 10.1016/j.physbeh.2006.09.009
28. Monk T.H. The post-lunch dip in performance // *Clin Sports Med*. 2005. Vol. 24, N 2. P. e15–23 xi–xii. doi: 10.1016/j.csm.2004.12.002
29. Krupinski E.A., Berbaum K.S., Caldwell R.T., et al. Long radiology workdays reduce detection and accommodation accuracy // *J Am Coll Radiol*. 2010. Vol. 7, N 9. P. 698–704. doi: 10.1016/j.jacr.2010.03.004
30. Krupinski E.A., Berbaum K.S., Caldwell R.T., et al. Do long radiology workdays affect nodule detection in dynamic CT interpretation? // *J Am Coll Radiol*. 2012. Vol. 9, N 3. P. 191–198. doi: 10.1016/j.jacr.2011.11.013
31. Al-s'adi M., McEntee M., Ryan E. Time of day does not affect radiologists' accuracy in breast lesion detection // *Proc SPIE Med Imag*. 2011. Vol. 7966. P. 1–7. doi: 10.1117/12.877972
32. Cowley H.C., Gale A.G. Time of day effects on mammographic film reading performance. 1997. Vol. 3036. doi: 10.1117/12.271295
33. Stec N., Arje D., Moody A.R., et al. A systematic review of fatigue in radiology: is it a problem? // *AJR Am J Roentgenol*. 2018. Vol. 210, N 4. P. 799–806. doi: 10.2214/AJR.17.18613
34. Krupinski E.A., Schartz K.M., Van Tassel M.S., et al. Effect of fatigue on reading computed tomography examination of the multiply injured patient // *J Med Imaging*. 2017. Vol. 4, N 3. P. 1. doi: 10.1117/1.JMI.4.3.035504
35. Hanna T.N., Zygmunt M.E., Peterson R., et al. The effects of fatigue from overnight shifts on radiology search patterns and diagnostic performance // *J Am Coll Radiol*. 2018. Vol. 15, N 12. P. 1709–1716. doi: 10.1016/j.jacr.2017.12.019
36. Gur D., Bandos A.I., Cohen C.S., et al. The “laboratory” effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations // *Radiology*. 2008. Vol. 249, N 1. P. 47–53. doi: 10.1148/radiol.2491072025
37. Miglioretti D.L., Ichikawa L., Smith R.A., et al. Correlation between screening mammography interpretive performance on a test set and performance in clinical practice // *Acad Radiol*. 2017. Vol. 24, N 10. P. 1256–1264. doi: 10.1016/j.acra.2017.03.016
38. Taylor-Phillips S., Wallis M.G., Jenkinson D., et al. Effect of using the same vs different order for second readings of screening mammograms on rates of breast cancer detection: a randomized clinical trial // *JAMA*. 2016. Vol. 315, N 18. P. 1956–1965. doi: 10.1001/jama.2016.5257
39. Stinton C., Jenkinson D., Adekanmbi V., et al. Does time of day influence cancer detection and recall rates in mammography? // *SPIE Medical Imaging*. 2017. Vol. 10136. doi: 10.1117/12.2254280
40. Wolfe J.M., Van Wert M.J. Varying target prevalence reveals two dissociable decision criteria in visual search // *Curr Biol*. 2010. Vol. 20, N 2. P. 121–124. doi: 10.1016/j.cub.2009.11.066
41. Burnside E.S., Park J.M., Fine J.P., Sisney G.A. The use of batch reading to improve the performance of screening mammography // *AJR Am J Roentgenol*. 2005. Vol. 185, N 3. P. 790–796. doi: 10.2214/ajr.185.3.01850790
42. Fenton J.J., Abraham L., Taplin S.H., et al. Effectiveness of computer-aided detection in community mammography practice // *J Natl Cancer Inst*. 2011. Vol. 103, N 15. P. 1152–1161. doi: 10.1093/jnci/djr206
43. Fenton J.J., Taplin S.H., Carney P.A., et al. Influence of computer-aided detection on performance of screening mammography // *N Engl J Med*. 2007. Vol. 356, N 14. P. 1399–1409. doi: 10.1056/NEJMoa066099

44. ISRCTN Register and Evaluating the age extension of the NHS breast screening Programme – trial registration. 2010.
45. Freemantle N., Ray D., McNulty D., et al. Increased mortality associated with weekend hospital admission: a case for

- expanded seven day services? // *BMJ*. 2015. Vol. 351. P. h4596. doi: 10.1136/bmj.h4596
46. Rimmer A. The BMJ paper and seven day services // *BMJ*. 2016. Vol. 352. P. i1193. doi: 10.1136/bmj.i1193

REFERENCES

1. Hogan H, Healey F, Neale G, et al. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf*. 2012;21(9):737–745. doi: 10.1136/bmjqs-2011-001159
2. Chief Medical Officer. An organisation with a memory: Report of an expert group on learning from adverse events in the NHS. London: The British Institute of Radiology; 2000.
3. Lockley SW, Barger LK, Ayas NT, et al. Effects of health care provider work hours and sleep deprivation on safety and performance. *Jt Comm J Qual Patient Saf*. 2007;33(11 Suppl):7–18. doi: 10.1016/S1553-7250(07)33109-7
4. Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. *N Engl J Med*. 2005;352:125–134. doi: 10.1056/NEJMoa041401
5. Hanna TN, Lamoureux C, Krupinski EA, et al. Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations. *Radiology*. 2018;287(1):205–212. doi: 10.1148/radiol.2017170555
6. Clinical Radiology UK workforce census 2014 report. London: The British Institute of Radiology; 2015.
7. Pigeon WR, Sateia MJ, Ferguson RJ. Distinguishing between excessive daytime sleepiness and fatigue: toward improved detection and treatment. *J Psychosom Res*. 2003;54(1):61–69. doi: 10.1016/S0022-3999(02)00542-1
8. Shahid A, Shen J, Shapiro CM. Measurements of sleepiness and fatigue. *J Psychosom Res*. 2010;69(1):81–89. doi: 10.1016/j.jpsychores.2010.04.001
9. Xu R, Zhang C, He F, et al. How physical activities affect mental fatigue based on EEG energy, connectivity, and complexity. *Front Neurol*. 2018;9:915. doi: 10.3389/fneur.2018.00915
10. Krupinski E, Reiner BI. Real-time occupational stress and fatigue measurement in medical imaging practice. *J Digit Imaging*. 2012;25(3):319–324. doi: 10.1007/s10278-011-9439-1
11. Waite S, Kolla S, Jeudy J, et al. Tired in the reading room: the influence of fatigue in radiology. *J Am Coll Radiol*. 2017;14(2):191–197. doi: 10.1016/j.jacr.2016.10.009
12. Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. *Int J Neurosci*. 1990;52(1-2):29–37. doi: 10.3109/00207459008994241
13. Hoddes E, Zarcone V, Smythe H, et al. Quantification of sleepiness: a new approach. *Psychophysiology*. 1973;10(4):431–436. doi: 10.1111/j.1469-8986.1973.tb00801.x
14. Johns MW. A new method for measuring daytime sleepiness: the epworth sleepiness scale. *Sleep*. 1991;14(6):540–545. doi: 10.1093/sleep/14.6.540
15. Rosenthal L, Roehrs TA, Roth T. The sleep-wake activity inventory: a self-report measure of daytime sleepiness. *Biol Psychiatry*. 1993;34(11):810–820. doi: 10.1016/0006-3223(93)90070-T
16. Buysse DJ, Reynolds CF, Monk TH, et al. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213. doi: 10.1016/0165-1781(89)90047-4
17. Ahsberg E. Dimensions of fatigue in different work-ing populations. *Scand J Psychol*. 2000;41(3):231–241. doi: 10.1111/1467-9450.00192
18. Miglioretti DL, Smith-Bindman R, Abraham L, et al. Radiologist characteristics associated with interpretive performance of diagnostic mammography. *J Natl Cancer Inst*. 2007;99(24):1854–1863. doi: 10.1093/jnci/djm238
19. Maeda E, Yoshikawa T, Hayashi N, et al. Radiology reading-caused fatigue and measurement of eye strain with critical flicker fusion frequency. *Jpn J Radiol*. 2011;29(7):483–487. doi: 10.1007/s11604-011-0585-7
20. Sheppard AL, Wolffsohn JS. Digital eye strain: prevalence, measurement and amelioration. *BMJ Open Ophthalmol*. 2018;3(1):e000146. doi: 10.1136/bmjophth-2018-000146
21. Thomson DR, Besner D, Smilek D. A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychol Rev*. 2016;123(1):70–83. doi: 10.1037/rev0000021
22. Mackworth N. Researches on the measurement of human performance. London, UK: The British Institute of Radiology; 1950.
23. See JE, Howe SR, Warm JS, et al. Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*. 1995;117(2):230–249. doi: 10.1037/0033-2909.117.2.230
24. Gur D, Rockette HE, Armfield DR, et al. Prevalence effect in a laboratory environment. *Radiology*. 2003;228(1):10–14. doi: 10.1148/radiol.2281020709
25. Evans KK, Birdwell RL, Wolfe JM. If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLoS One*. 2013;8(5):e64366. doi: 10.1371/journal.pone.0064366
26. Kompaniez-Dunigan E, Abbey CK, Boone JM, Webster MA. Visual adaptation and the amplitude spectra of radiological images. *Cogn Res Princ Implic*. 2018;3(1):3. doi: 10.1186/s41235-018-0089-4
27. Blatter K, Cajochen C. Circadian rhythms in cognitive performance: methodological constraints, protocols, theoretical underpinnings. *Physiol Behav*. 2007;90(2-3):196–208. doi: 10.1016/j.physbeh.2006.09.009
28. Monk TH. The post-lunch dip in performance. *Clin Sports Med*. 2005;24(2):e15–23xi-xii. doi: 10.1016/j.csm.2004.12.002
29. Krupinski EA, Berbaum KS, Caldwell RT, et al. Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol*. 2010;7(9):698–704. doi: 10.1016/j.jacr.2010.03.004
30. Krupinski EA, Berbaum KS, Caldwell RT, et al. Do long radiology workdays affect nodule detection in dynamic CT interpretation? *J Am Coll Radiol*. 2012;9(3):191–198. doi: 10.1016/j.jacr.2011.11.013
31. Al-s'adi M, McEntee M, Ryan E. Time of day does not affect radiologists' accuracy in breast lesion detection. *Proc SPIE Med Imag*. 2011;7966:1–7. doi: 10.1117/12.877972
32. Cowley HC, Gale AG. Time of day effects on mammographic film reading performance. 1997. doi: 10.1117/12.271295
33. Stec N, Arje D, Moody AR, et al. A systematic review of fatigue in radiology: is it a problem? *AJR Am J Roentgenol*. 2018;210(4):799–806. doi: 10.2214/AJR.17.18613

34. Krupinski EA, Scharzt KM, Van Tassell MS, et al. Effect of fatigue on reading computed tomography examination of the multiply injured patient. *J Med Imaging*. 2017;4(3):1. doi: 10.1117/1.JMI.4.3.035504
35. Hanna TN, Zygmunt ME, Peterson R, et al. The effects of fatigue from overnight shifts on radiology search patterns and diagnostic performance. *J Am Coll Radiol*. 2018;15(12):1709–1716. doi: 10.1016/j.jacr.2017.12.019
36. Gur D, Bandos AI, Cohen CS, et al. The “laboratory” effect: comparing radiologists’ performance and variability during prospective clinical and laboratory mammography interpretations. *Radiology*. 2008;249(1):47–53. doi: 10.1148/radiol.2491072025
37. Miglioretti DL, Ichikawa L, Smith RA, et al. Correlation between screening mammography interpretive performance on a test set and performance in clinical practice. *Acad Radiol*. 2017;24(10):1256–1264. doi: 10.1016/j.acra.2017.03.016
38. Taylor-Phillips S, Wallis MG, Jenkinson D, et al. Effect of using the same vs different order for second readings of screening mammograms on rates of breast cancer detection: a randomized clinical trial. *JAMA*. 2016;315(18):1956–1965. doi: 10.1001/jama.2016.5257
39. Stinton C, Jenkinson D, Adekanmbi V, et al. Does time of day influence cancer detection and recall rates in mammography? *SPIE Medical Imaging*. 2017. Vol. 10136. doi: 10.1117/12.2254280.
40. Wolfe JM, Van Wert MJ. Varying target prevalence reveals two dissociable decision criteria in visual search. *Curr Biol*. 2010;20(2):121–124. doi: 10.1016/j.cub.2009.11.066
41. Burnside ES, Park JM, Fine JP, Sisney GA. The use of batch reading to improve the performance of screening mammography. *AJR Am J Roentgenol*. 2005;185(3):790–796. doi: 10.2214/ajr.185.3.01850790
42. Fenton JJ, Abraham L, Taplin SH, et al. Effectiveness of computer-aided detection in community mammography practice. *J Natl Cancer Inst*. 2011; 103(15):1152–1161. doi: 10.1093/jnci/djr206
43. Fenton JJ, Taplin SH, Carney PA, et al. Influence of computer-aided detection on performance of screening mammography. *N Engl J Med*. 2007;356(14):1399–1409. doi: 10.1056/NEJMoa066099
44. ISRCTN Register and Evaluating the age extension of the NHS breast screening Programme – trial registration. 2010.
45. Freemantle N, Ray D, McNulty D, et al. Increased mortality associated with weekend hospital admission: a case for expanded seven day services? *BMJ*. 2015;351:h4596. doi: 10.1136/bmj.h4596
46. Rimmer A. The BMJ paper and seven day services. *BMJ*. 2016;352:i1193. doi: 10.1136/bmj.i1193

05 ABTOPAX

* Sian Taylor-Phillips, Ph.D.;

ORCID: <http://orcid.org/0000-0002-1841-4346>;

e-mail: s.taylor-phillips@warwick.ac.uk

Chris Stinton;

ORCID: <https://orcid.org/0000-0001-9054-1940>;

e-mail: c.stinton@bham.ac.uk

AUTHORS’ INFO

* Sian Taylor-Phillips, Ph.D.;

ORCID: <http://orcid.org/0000-0002-1841-4346>;

e-mail: s.taylor-phillips@warwick.ac.uk

Chris Stinton;

ORCID: <https://orcid.org/0000-0001-9054-1940>;

e-mail: c.stinton@bham.ac.uk

* Автор, ответственный за переписку / Corresponding author

DOI: <https://doi.org/10.17816/DD70922>

Длительный анамнез бронхоцеле, вызванный типичным карциноидом

К.В. Прусакова, П.В. Гаврилов

Санкт-Петербургский научно-исследовательский институт фтизиопульмонологии, Санкт-Петербург, Российская Федерация

АННОТАЦИЯ

В работе представлен клинический случай с длительным периодом наблюдения одиночного бронхоцеле (бронхогенной ретенционной кисты). При первоначальном комплексном обследовании, включающем такие исследования, как рентгенография, компьютерная томография органов грудной полости, фибробронхоскопия, иммунологические и бактериологические обследования на туберкулёз, данных за онкологическую и инфекционную природу изменений не выявлено. Изменения были расценены как последствия перенесённого неспецифического воспалительного процесса. Через 15 лет при плановом медицинском осмотре по данным рентгенографии органов грудной полости отмечено увеличение размеров бронхоцеле, а также появление округлого образования в медиальных отделах бронхоцеле. С помощью дополнительных методов исследования, таких как компьютерная томография органов грудной полости с внутривенным контрастированием, фибробронхоскопия с биопсией, установлено, что выявленное образование является типичным карциноидом.

Несмотря на то что бронхоцеле в большинстве случаев является доброкачественным изменением, из разнообразия причин, вызывающих его развитие, следует выделить обструкцию бронха новообразованием. Среди новообразований лёгкого типичный карциноид составляет всего 1–2%, характеризуется крайне медленным ростом и отсутствием специфичной клинической симптоматики. Несмотря на это, типичный карциноид относится к злокачественным нейроэндокринным образованиям I типа. В 10–15% случаев выявляются метастазы, преимущественно в медиастинальные лимфатические узлы, а также в печень, кости, реже в мягкие ткани.

Данное клиническое наблюдение говорит о том, что даже при отрицательных результатах первичного обследования локально расположенного бронхоцеле такие изменения требуют онкологической настороженности и периодических обследований в динамике.

Ключевые слова: клинический случай; бронхоцеле; типичный карциноид; компьютерная томография.

Как цитировать

Прусакова К.В., Гаврилов П.В. Длительный анамнез бронхоцеле, вызванный типичным карциноидом // *Digital Diagnostics*. 2021. Т. 2, № 2. С. 223–230.
DOI: <https://doi.org/10.17816/DD70922>

DOI: <https://doi.org/10.17816/DD70922>

Long-term bronchocele anamnesis, triggered by typical carcinoid

Kseniya V. Prusakova, Pavel V. Gavrilov

Saint-Petersburg State Research Institute of Phthisiopulmonology, Saint Petersburg, Russian Federation

ABSTRACT

The paper presents a case of a single bronchocele (bronchogenic retention cyst) caused by a typical carcinoid that was observed for a long time. During the initial complex examination, including computed tomography with intravenous contrast, fibrobronchoscopy, and immunological and bacteriological examinations of tuberculosis, there were no changes for the oncological and infectious nature. The changes were interpreted as the result of a postponed nonspecific inflammatory process. Most of them were monitored using chest X-ray and the changes were stable. After 15 years, a control chest X-ray revealed an increase in the size of the compaction in the lung and the appearance of a mass with calcification in the medial sections of the compaction zone. Additional examination, including computed tomography with biopsy, determined that the obstruction of the bronchus was caused by a neoplasm [according to histological examination (typical carcinoid)].

It should be noted that the initial detection of negative study results requires oncological alertness and periodic examinations in dynamics.

Keywords: case report; bronchocele; typical carcinoid; computed tomography.

To cite this article

Prusakova KV, Gavrilov PV. Long-term bronchocele anamnesis, triggered by typical carcinoid. *Digital Diagnostics*. 2021;2(2):223–230. DOI: <https://doi.org/10.17816/DD70922>

Received: 23.05.2021

Accepted: 23.06.2021

Published: 01.07.2021

DOI: <https://doi.org/10.17816/DD70922>

由典型的类癌引起的支气管囊肿的悠久历史

Kseniya V. Prusakova, Pavel V. Gavrilo

Saint-Petersburg State Research Institute of Phthisiopulmonology, Saint Petersburg, Russian Federation

简评

本文提出了一个长期观察单个支气管囊肿（支气管源性保留囊肿）的临床病例。在最初的全局检查中，包括放射照相术，胸腔计算机断层扫描，纤维支气管镜检查，结核病的免疫学和细菌学检查等研究，没有发现改变的肿瘤和感染性的数据。这些变化被视为转移非特异性炎症过程的后果。15年后，在常规体检期间，根据胸腔的射线照相，注意到支气管的大小增加，以及支气管囊肿内侧部分圆形出现。在其他研究方法的帮助下（例如胸腔静脉造影的计算机断层扫描，活检的纤维支气管镜检查），确定检测到的形成是典型的类癌。

尽管在大多数情况下，支气管囊肿是一种良性变化，但从导致其发展的各种原因来看，有必要通过肿瘤来区分支气管囊肿的阻塞。在肺部肿瘤中，典型的类癌仅为1-2%，其特征在于极其缓慢的生长和没有特定的临床症状。尽管如此，典型的类癌属于第一类型恶性神经内分泌形成。在10-15%的病例中，检测到转移，主要在纵隔淋巴结中，以及在肝脏，骨骼中，在软组织中较少。

这一临床观察表明，即使对局部定位的支气管囊肿的初步检查结果为阴性，这种变化也需要肿瘤警觉性和动态的定期检查。

关键词：临床病例；支气管囊肿；典型类癌；计算机断层扫描。

引用本文：

Prusakova KV, Gavrilo PV. 由典型的类癌引起的支气管囊肿的悠久历史。 *Digital Diagnostics*. 2021;2(2):223-230.

DOI: <https://doi.org/10.17816/DD70922>

收到: 23.05.2021

接受: 23.06.2021

发布日期: 01.07.2021

ВВЕДЕНИЕ

Бронхоцеле (бронхогенная ретенционная киста, мукоцеле) является относительно частой находкой при рентгенологических исследованиях органов грудной клетки. Морфологическим субстратом бронхоцеле является локальное расширение бронхов с заполнением дыхательных путей слизистым содержимым вследствие продолжения выделения секрета слизистой оболочкой и проксимальной обструкции дыхательных путей [1]. При рентгенографии и компьютерной томографии бронхоцеле визуализируется в виде трубчатых разветвлённых структур V- или Y-образной формы, связанных с бронхиальным деревом (симптом «пальцев в перчатке») [2]. Структура содержимого однородная, но в 30% случаев в структуре визуализируются плотные включения — кальцинаты [2, 3]. При компьютерной томографии с внутривенным контрастированием накопления контрастного препарата не происходит.

В части случаев бронхоцеле могут принимать овальные или округлые очертания, что зависит от калибра обтурированного бронха, количества содержимого в просвете и от состояния окружающей лёгочной паренхимы.

Одиночные локальные ретенционные кисты протекают бессимптомно. Спектр причин, вызывающих развитие ретенционных кист, очень широк: врождённая патология (бронхиальная атрезия, секвестрация лёгкого, муковисцидоз); инфекционная патология (неспецифические воспалительные процессы, туберкулёз, микобактериоз, аллергический бронхолёгочный аспергиллёз); обструкция бронха образованием (доброкачественным, злокачественным), инородным телом или рубцовая деформация бронха. Дифференциальную диагностику затрудняет то, что бронхоцеле, вызванные различными причинами, имеют схожую рентгенологическую семиотику [2].

Дифференцировать бронхоцеле следует с артериовенозной мальформацией в лёгких, эндобронхиальным метастазом. Компьютерная томография с внутривенным контрастированием в таком случае является предпочтительным методом диагностики [2].

В большинстве случаев бронхоцеле вызвано доброкачественными изменениями в лёгких и не требует динамического наблюдения, однако при наличии локально расположенного бронхоцеле следует исключить обтурационный генез образованием или инородным телом. Для этого лучевые методы исследования рекомендовано дополнить фибробронхоскопией с биопсией [4, 5].

В настоящий момент не разработан диагностический алгоритм, позволяющий оптимальным способом выявить причину развития бронхоцеле, так же как нет единых рекомендаций по дальнейшему наблюдению пациентов с впервые выявленными бессимптомными ретенционными кистами (бронхоцеле).

ОПИСАНИЕ СЛУЧАЯ

Пациент, мужчина, 56 лет, обратился в отделение лучевой диагностики для проведения компьютерной томографии органов грудной полости.

Из анамнеза известно, что 15 лет назад он проходил обследование по поводу пневмонии. Несмотря на положительную динамику, по данным клинических исследований, на фоне курса антибактериальной терапии, рентгенологические данные не соответствовали типичному течению регрессии инфильтративных изменений в лёгких при пневмонии. По данным рентгенографии органов грудной полости, в среднем отделе правого лёгкого определялся участок уплотнения трубчатой разветвлённой структуры (рис. 1, а). В качестве дополнительных исследований были проведены компьютерная

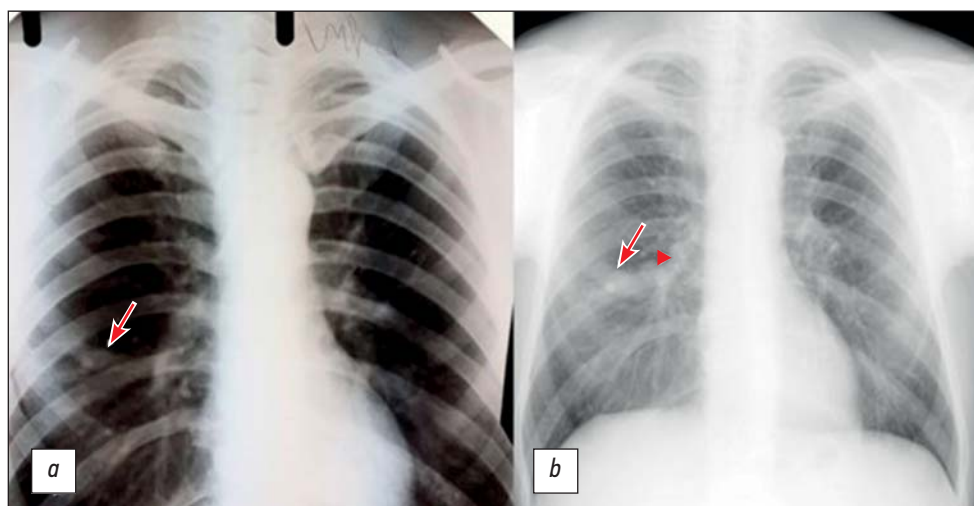


Рис. 1. Пациент, 56 лет, рентгенограмма органов грудной полости: *а* — при первичном исследовании в возрасте 41 года в среднем отделе правого лёгкого определяется участок уплотнения разветвлённой трубчатой структуры (стрелка); *б* — через 15 лет отмечено увеличение размеров бронхоцеле (стрелка) и появление округлого образования в медиальных отделах бронхоцеле (головка стрелки).



Рис. 2. Тот же пациент. Выборочный скан компьютерной томографии органов грудной полости: однородная V-образная структура в средней доле правого лёгкого с чёткими контурами (стрелка).

томография органов грудной полости с внутривенным контрастированием, фибробронхоскопия, иммунологические и бактериологические исследования, по результатам которых данных за туберкулёз и онкологический процесс не получено. Результат компьютерной томографии был представлен в виде выборочных сканов на плёночном носителе, на которых была выявлена локальная единичная разветвлённая структура с ровными чёткими

контурами, расположенная по ходу субсегментарных бронхов средней доли правого лёгкого (симптом «пальцев в перчатке»), с однородным содержимым (рис. 2). Был установлен диагноз бронхогенной ретенционной кисты (бронхоцеле) средней доли правого лёгкого. В дальнейшем контрольные исследования проводились при помощи рентгенографии органов грудной полости ежегодно, выявленные изменения были стабильны.

В настоящий момент при прохождении медицинского осмотра для устройства на работу с вредными условиями труда по данным рентгенографии органов грудной полости выявлено увеличение размеров ранее определяемого бронхоцеле (рис. 1, *b*), а также новое округлое образование в медиальных отделах бронхоцеле с обызвествлениями по контуру образования (см. рис. 1, *b*). Для уточнения характера изменений пациенту выполнена компьютерная томография органов грудной полости с внутривенным контрастированием, по данным которой в средней доле правого лёгкого сохраняется единичная разветвлённая структура V-образной формы с чётким контуром и однородным содержимым, расположенная по ходу субсегментарных бронхов (симптом «пальцев в перчатке»). У основания бронхоцеле определяется округлое образование с ровным чётким контуром, практически полностью перекрывающее просвет бронха В4, и единичными кальцинатами по периферии с признаками накопления контрастного препарата в венозную фазу от +29НУ до +112НУ (рис. 3).

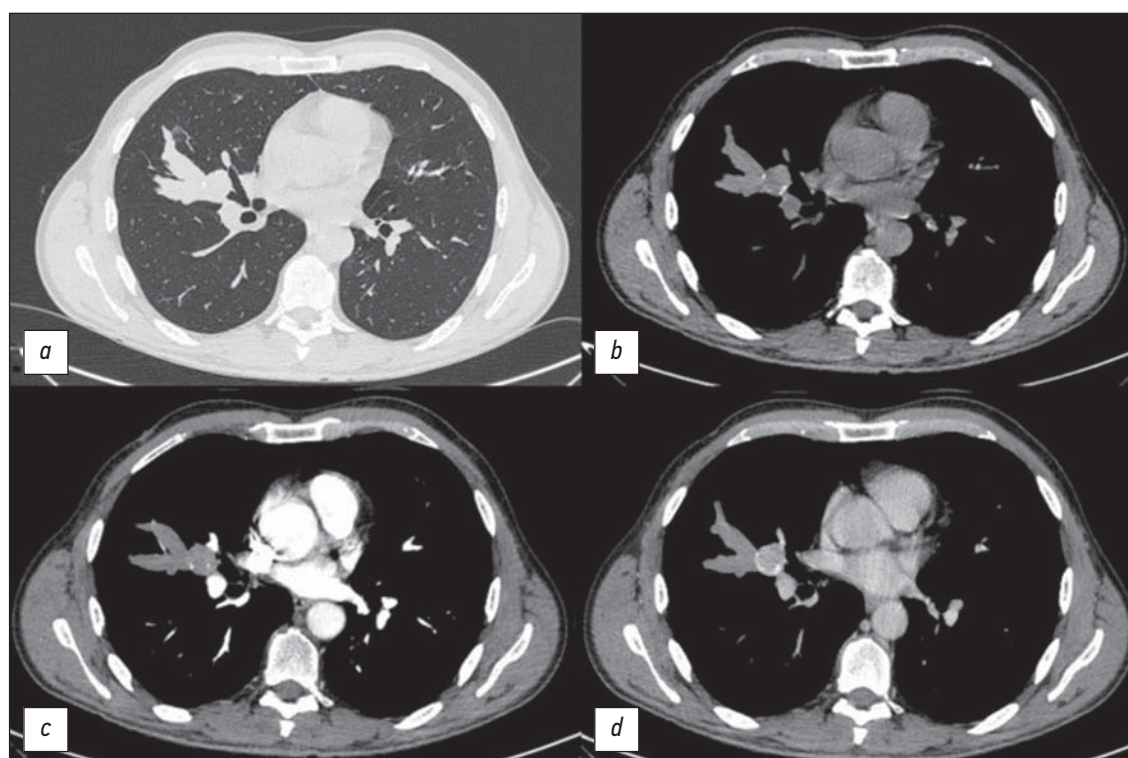


Рис. 3. Тот же пациент. Компьютерная томография органов грудной полости в аксиальной плоскости: *a* — лёгочное окно, нативная фаза (округлое образование в основании бронхоцеле); *b* — средостенное окно (единичные кальцинаты по периферии образования); *c* — средостенное окно, артериальная фаза; *d* — средостенное окно, венозная фаза (признаки накопления контрастного препарата образованием).

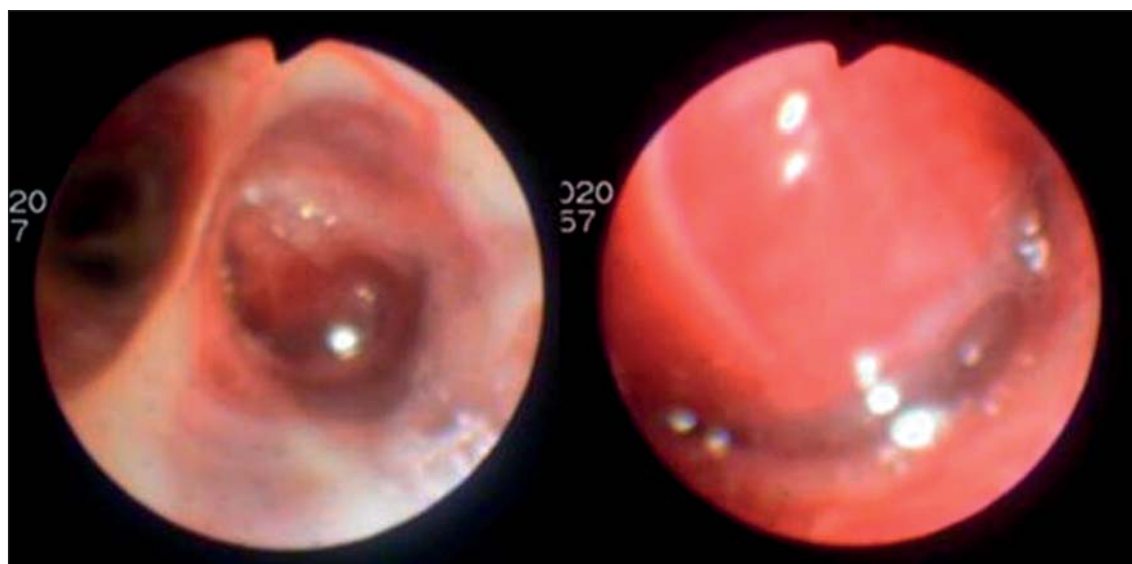


Рис. 4. Тот же пациент. Фибробронхоскопия: образование устья В4 справа, полностью перекрывающее просвет бронха.

Изменения были характерны для бронхоцеле, вызванного обструкцией бронха новообразованием. В качестве дополнительных методов исследования выполнена фибробронхоскопия с биопсией. При бронхоскопии отмечается округлое образование устья В4, полностью перекрывающее просвет бронха (рис. 4). Образование малоподвижное, при контакте ранимое, слизистая на поверхности гиперемизированная, отёчная. По результатам биопсии установлено, что гистологическая картина образования соответствует типичному карциноиду. При иммуногистохимическом исследовании опухолевые клетки интенсивно экспрессируют CD56 и не экспрессируют TTF1. Индекс пролиферативной активности Ki67 2%.

Пациенту выполнено оперативное лечение в объёме резекции средней доли правого лёгкого. При контрольном обследовании через год по данным компьютерной томографии органов грудной полости признаков рецидива карциноида не выявлено.

ОБСУЖДЕНИЕ

Наиболее распространёнными причинами образования множественных бронхоцеле являются муковисцидоз, аллергический бронхолегочный аспергиллёз и туберкулёз. Одиночные локальные ретенционные кисты чаще вызваны обструкцией бронха новообразованием (доброкачественным или злокачественным) [2, 6].

Среди новообразований лёгкого типичный карциноид составляет всего 1–2% [7]. В 70% случаев опухоль локализуется в главных бронхах, чаще в правом лёгком, преимущественно в средней доле [8]. Средний возраст людей с типичным карциноидом составляет 40–50 лет. При такой форме новообразования лёгкого не выявлено достоверной взаимосвязи с влиянием канцерогенов и курением [9, 10].

В большинстве случаев карциноид бронхов протекает бессимптомно и обнаруживается в качестве случайной находки при плановом исследовании, однако в 2–5% случаев карциноиды бронхов могут продуцировать нейромедиаторы и пептидные гормоны (серотонин, адренокортикотропный гормон, соматостатин и брадикинин) [11]. К клиническим проявлениям карциноидного синдрома относятся периодические приступы жара или чувство прилива крови к голове, шее и рукам, бронхоспазм, диарея, психические расстройства [11–13].

На рентгенограммах типичный карциноид визуализируется в виде округлого или овального образования с чёткими ровными (иногда дольчатыми) контурами. Достаточно часто (до 30% случаев) встречаются эксцентрично расположенные или диффузные обызвествления [2, 3].

При компьютерной томографии типичный карциноид визуализируется в виде округлого образования с чёткими ровными или дольчатыми контурами. При внутривенном контрастировании отмечается накопление контрастного препарата, в некоторых случаях возможно проследить питающую артерию, подходящую к образованию из системы бронхиальных артерий [6]. По отношению к бронху карциноид располагается интрабронхиально, экстрабронхиально и смешанно по типу «айсберга», вызывая при этом частичную или полную обтурацию просвета бронха [2, 3].

Несмотря на то, что при первичном комплексном обследовании причина развития бронхоцеле в данном клиническом примере не установлена, при ретроспективной оценке данных компьютерной томографии, представленных на плёночном носителе, можно предположить наличие образования в основании бронхоцеле (см. рис. 2). При экстрабронхиальном его расположении изменения при фибробронхоскопии могут быть не выявлены.

Денситометрические параметры образования, расположенного в основании ретенционной кисты, могут не слишком отличаться от слизи, и при небольших его

размерах визуализация может быть затруднена. Центральные формы карциноида могут быть заподозрены при выявлении признаков обструкции (ателектазы, «воздушные ловушки», бронхоцеле).

Дифференцировать типичный карциноид следует с нейроэндокринными образованиями легких II типа (атипичный карциноид), бронхогенной кистой, бронхоцеле.

Типичный карциноид отличается крайне медленным ростом. Согласно результатам исследования D. Raz и соавт. [14], среднее время удвоения типичных карциноидных опухолей составляет 7 лет, поэтому тяжело судить о наличии динамики по данным ежегодной профилактической рентгенографии лёгких, так как сложно визуально уловить незначительное увеличение размеров образования. Авторами высказано предложение, что при наличии локально расположенного бронхоцеле неустовленной природы, несмотря на видимое отсутствие динамики по данным рентгенографии, контрольные исследования при помощи компьютерной томографии органов грудной полости с внутривенным контрастированием необходимо проводить с некоторой периодичностью для достоверной оценки динамики изменений и исключения обструкции бронха новообразованием.

Компьютерная томография является предпочтительным методом диагностики, однако, учитывая особенности расположения типичных карциноидов, в качестве дополняющих методов исследования многие авторы рекомендуют фибробронхоскопию с трансбронхиальной биопсией [4, 5, 15].

Золотым стандартом лечения типичных карциноидов является хирургическая резекция, так как данная патология имеет низкую чувствительность к химио- и лучевой терапии. В случае полного эндобронхиального расположения карциноида в центральных отделах резекция может быть выполнена трансбронхиальным способом [6, 8, 13].

ЗАКЛЮЧЕНИЕ

Несмотря на то, что в большинстве случаев бронхоцеле является доброкачественным изменением, при выявлении локально расположенного бронхоцеле необходимо исключить онкологическую природу обструкции бронха, для чего рекомендовано проведение

компьютерной томографии органов грудной полости с внутривенным контрастированием и дополнительно фибробронхоскопии с биопсией.

Следует помнить, что некоторые типы новообразований, таких как типичный карциноид, характеризуются крайне медленным ростом, и даже при отрицательных результатах первичного обследования локально расположенного бронхоцеле данные изменения требуют онкологической настороженности и периодических обследований в динамике.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при подготовке и публикации статьи.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. К.В. Прусакова — сбор материала, написание статьи; П.В. Гаврилов — обработка полученных результатов, финальное редактирование публикации. Авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Информированное согласие на публикацию. Пациент подписал форму добровольного информированного согласия на публикацию медицинской информации в обезличенной форме в журнале Digital Diagnostics.

Funding source. This publication was not supported by any external sources of funding.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. K.V. Prusakova — collecting material, writing an article; P.V. Gavrilov — processing of the results obtained, final editing of the publication. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

Consent for publication. Written consent was obtained from the patient for publication of relevant medical information and all of accompanying images within the manuscript.

СПИСОК ЛИТЕРАТУРЫ

1. Hansell D.M., Bankier A.A., MacMahon H., et al. Fleischner Society: glossary of terms for thoracic imaging // *Radiology*. 2008. Vol. 246, N 3. P. 697–722. doi: 10.1148/radiol.2462070712
2. Martinez S., Heyneman L.E., McAdams H.P., et al. Mucoid impactions: finger-in-glove sign and other CT and radiographic features // *Radiographics*. 2008. Vol. 28, N 5. P. 1369–1382. doi: 10.1148/rq.285075212
3. Nguyen E.T. The gloved finger sign // *Radiology*. 2003. Vol. 227, N 2. P. 453–454. doi: 10.1148/radiol.2272011548
4. Farrell C., Goggins M., Casserly M. Unexpected diagnosis resulting from presentation with chronic obstructive pulmonary disease (COPD) exacerbation // *International Journal of Case Reports and Images*. 2019;43–47. doi: 10.36811/jcri.2019.110007
5. Kulkarni G.S., Gawande S.C., Chaudhari D.V., Bhoir A.P. Bronchial carcinoid: case report and review of literature // *MVP J Med Sci*. 2016. Vol. 3, N 1. P. 71–78. doi: 10.18311/mvpjms/2016/v3/i1/740
6. Yadav V., Rathi V. Bronchial carcinoid with bronchocele masquerading as Scimitar syndrome on chest radio-

graph // *Radiol Case Rep.* 2021. Vol. 16, N 3. P. 710–713. doi: 10.1016/j.radcr.2021.01.013

7. Jeung M.Y., Gasser B., Gangi A., et al. Bronchial carcinoid tumors of the thorax: spectrum of radiologic findings // *Radiographics.* 2002. Vol. 22, N 2. P. 351–365. doi: 10.1148/radiographics.22.2.g02mr01351

8. Paladugu R.R., Benfield J.R., Pak H.Y., et al. Bronchopulmonary Kulchitzky cell carcinomas. A new classification scheme for typical and atypical carcinoids // *Cancer.* 1985. Vol. 55, N 6. P. 1303–1311. doi: 10.1002/1097-0142(19850315)55:6<1303::aid-cnrcr2820550625>3.0.co;2-a

9. Grote T.H., Macon W.R., Davis B., et al. Atypical carcinoid of the lung. A distinct clinicopathologic entity // *Chest.* 1988. Vol. 93, N 2. P. 370–375. doi: 10.1378/chest.93.2.370

10. Harpole D.H., Feldman J.M., Buchanan S., et al. Bronchial carcinoid tumors: a retrospective analysis of 126 patients // *Ann Thorac Surg.* 1992. Vol. 54, N 1. P. 50–54; discussion 54–5. doi: 10.1016/0003-4975(92)91139-z

11. Кузнецов Н.С., Латкина Н.В., Добрева Е.А. АКТГ — эктопированный синдром: клиника, диагностика, лечение // *Эндокринная хирургия.* 2012. Т. 6, № 1. С. 24–36.

12. Бурякина С.А., Кармазановский Г.Г., Волеводз Н.Н., и др. КТ-признаки нейроэндокринных опухолей легких и их взаимосвязь с АКТГ-эктопическим синдромом // *REJR.* 2018. Т. 8, № 4. С. 56–72. doi: 10.21569/2222–7415-2018-8-4-56-72

13. Трахтенберг А.Х., Колбанов К.И., Франк Г.А., и др. Особенности диагностики и лечения карциноидных опухолей легких // *Атмосфера. Пульмонология и аллергология.* 2009. № 1. С. 2–6.

14. Raz D.J., Nelson R.A., Grannis F.W., Kim J.Y. Natural history of typical pulmonary carcinoid tumors: a comparison of nonsurgical and surgical treatment // *Chest.* 2015. Vol. 147, N 4. P. 1111–1117. doi: 10.1378/chest.14-1960

15. Kaifi J.T., Kayser G., Ruf J., Passlick B. The diagnosis and treatment of bronchopulmonary carcinoid // *Dtsch Arztebl Int.* 2015. Vol. 112, N 27–28. P. 479–485. doi: 10.3238/arztebl.2015.0479

REFERENCES

1. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. *Radiology.* 2008;246(3):697–722. doi: 10.1148/radiol.2462070712

2. Martinez S, Heyneman LE, McAdams HP, et al. Mucoid impactions: finger-in-glove sign and other CT and radiographic features. *Radiographics.* 2008;28(5):1369–1382. doi: 10.1148/rg.285075212

3. Nguyen ET. The gloved finger sign. *Radiology.* 2003;227(2):453–454. doi: 10.1148/radiol.2272011548

4. Farrell C, Goggins M, Casserly M. Unexpected diagnosis resulting from presentation with chronic obstructive pulmonary disease (COPD) exacerbation. *International Journal of Case Reports and Images.* 2019;43–47. doi: 10.36811/jcri.2019.110007

5. Kulkarni GS, Gawande SC, Chaudhari DV, Bhoyar AP. Bronchial carcinoid: case report and review of literature. *MVP J Med Sci.* 2016;3(1):71–78. doi: 10.18311/mvpjms/2016/v3/i1/740

6. Yadav V, Rath V. Bronchial carcinoid with bronchocele masquerading as Scimitar syndrome on chest radiograph. *Radiol Case Rep.* 2021;16(3):710–713. doi: 10.1016/j.radcr.2021.01.013

7. Jeung MY, Gasser B, Gangi A, et al. Bronchial carcinoid tumors of the thorax: spectrum of radiologic findings. *Radiographics.* 2002;22(2):351–365. doi: 10.1148/radiographics.22.2.g02mr01351

8. Paladugu RR, Benfield JR, Pak HY, et al. Bronchopulmonary Kulchitzky cell carcinomas. A new classification scheme for typi-

cal and atypical carcinoids. *Cancer.* 1985;55(6):1303–1311. doi: 10.1002/1097-0142(19850315)55:6<1303::aid-cnrcr2820550625>3.0.co;2-a

9. Grote TH, Macon WR, Davis B, et al. Atypical carcinoid of the lung. A distinct clinicopathologic entity. *Chest.* 1988;93(2):370–375. doi: 10.1378/chest.93.2.370. PMID: 2827965

10. Harpole DH, Feldman JM, Buchanan S, et al. Bronchial carcinoid tumors: a retrospective analysis of 126 patients. *Ann Thorac Surg.* 1992;54(1):50–54; discussion 54–5. doi: 10.1016/0003-4975(92)91139-z

11. Kuznetsov NS, Latkina NV, Dobrova EA. ACTH-ectopic syndrome: clinic, diagnosis, treatment. *Endocrine surgery.* 2012;6(1):24–36. (In Russ).

12. Buryakina SA, Karmazanovsky GG, Volevodz NN, et al. CT-signs of neuroendocrine lung tumors and their relationship with ACTH-ectopic syndrome. *REJR.* 2018;8(4):56–72. (In Russ). doi: 10.21569/2222–7415-2018-8-4-56-72

13. Trachtenberg AH, Kolbanov KI, Frank GA, et al. Features of diagnosis and treatment of lung carcinoid tumors. *Atmosphere. Pulmonology and allergology.* 2009;(1):2–6. (In Russ).

14. Raz DJ, Nelson RA, Grannis FW, Kim JY. Natural history of typical pulmonary carcinoid tumors: a comparison of nonsurgical and surgical treatment. *Chest.* 2015;147(4):1111–1117. doi: 10.1378/chest.14-1960

15. Kaifi JT, Kayser G, Ruf J, Passlick B. The diagnosis and treatment of bronchopulmonary carcinoid. *Dtsch Arztebl Int.* 2015;112(27–28):479–485. doi: 10.3238/arztebl.2015.0479

ОБ АВТОРАХ

* **Гаврилов Павел Владимирович**, к.м.н.;

адрес: Россия, 191036, Санкт-Петербург, Лиговский пр., д. 2-4;

e-mail: spbniifrentgen@mail.ru;

ORCID: <https://orcid.com/0000-0003-3251-4084>

Прусакова Ксения Владимировна;

e-mail: ksenya.rush@mail.ru;

ORCID: <https://orcid.com/0000-0002-3934-6290>

AUTHORS' INFO

* **Pavel V. Gavrilov**, MD, Cand. Sci. (Med.);

address: 2-4, Ligovskiy pr., Saint-Petersburg, 191036, Russia;

e-mail: spbniifrentgen@mail.ru;

ORCID: <https://orcid.com/0000-0003-3251-4084>

Kseniya V. Prusakova, MD;

e-mail: ksenya.rush@mail.ru;

ORCID: <https://orcid.com/0000-0002-3934-6290>

* Автор, ответственный за переписку / Corresponding author