



научный центр
экспертизы средств
медицинского применения

Отчет клинического исследования, имитирующий успех – частые разновидности манипулирования результатами

Горячев Д.В., директор ЦЭК ГЛС

Федеральное государственное бюджетное учреждение
«**Научный центр экспертизы средств медицинского применения**»
Министерства здравоохранения Российской Федерации



Постановка проблемы – кардинальное различие в оценке отчета о клиническом исследовании между экспертами и разработчиками препарата

Экспертам видится очевидное отсутствие подтверждения эффективности/безопасности препарата

Разработчик/компания убеждают, что отчет свидетельствует об эффективности и безопасности препарата





Как часто желание не совпадает с результатом?

Contemporary Clinical Trials Communications 11 (2018) 156–164

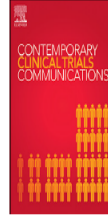


ELSEVIER

Contents lists available at [ScienceDirect](#)

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc



Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review

David B. Fogel

Trials.ai, 4520 Executive Dr., Suite 200, San Diego, CA, 92121, United States

2. Failing to demonstrate efficacy or safety

The primary source of trial failure has been and remains an inability to demonstrate efficacy. Hwang et al. [58] assessed 640 phase 3 trials with novel therapeutics and found that 54% failed in clinical development, with 57% of those failing due to inadequate efficacy. There are many reasons that potentially efficacious drugs can still fail to



Из 640 КИ III фазы новых препаратов (2010-2015) 54% не привели к регистрации: 31% - по причине не доказанной эффективности.

T.J. Hwang, et al. Failure of investigational drugs in late-stage clinical development and publication of trial results, *JAMA Intern. Med* 176 (2016) 1826–1833.



Отчет о клиническом исследовании – универсальный документ, подтверждающий эффективность и безопасность препарата

- Профессиональное клиническое сообщество впервые видит эффекты и вырабатывает отношение к препарату.
- Документ оцениваемый всеми сторонами, **в т.ч. не обладающими компетентностью в проведении КИ.**
- Неотъемлемый элемент доказательной медицины.
- **Становится открытым документом (например, ЕМА)!!!**



Для повышения вероятности успеха КИ работают над дизайном, набором больных, выбором центров...

Contemporary Clinical Trials Communications 11 (2018) 156-164



ELSEVIER

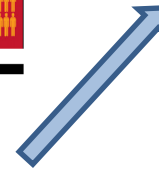
Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc



Неудачный: дизайн, выбор центров, набор, избыточная нагрузка на пациентов, низкое качество выполнения исследования....



Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review

David B. Fogel

Trials.ai, 4520 Executive Dr., Suite 200, San Diego, CA, 92121, United States

А если желаемый результат не достигнут?

Table 1

A list of factors associated with problems or challenges when preparing for or executing a clinical trial, along with the opportunities for artificial intelligence to help alleviate these issues. Abbreviation: NLP = natural language processing.

| Factor | Opportunity | Role for Artificial Intelligence |
|----------------------------|---|---|
| Poor study design | More complete literature review | NLP of available literature, finding similar trials, trials addressing similar issues, or trials addressing different issues utilizing similar techniques, summarized for the study designer |
| | Appropriate endpoints | NLP of available literature, showing endpoints/measures used in other similar studies |
| | Inappropriate eligibility criteria | NLP assessment of similar published trials to determine suitability of eligibility criteria and any potentially important omissions. |
| | Appropriate statistical analysis | NLP of available literature, summarizing statistical methods and associating these methods with successful or failed outcomes. |
| Ineffective site selection | Determination of appropriate sample size | Nonlinear modeling, such as with neural networks, to predict patient drop-out rates and better estimate sample size to avoid becoming underpowered. Agent-based modeling to simulate trial before execution. Use of NLP to mine previously published trials to determine sample sizes used in successful trials |
| | Reducing likelihood of amendments | NLP and knowledge-based processing to present designer with pertinent information to consider. |
| | Inconsistencies in protocol | NLP (including table-based format) to check time and events schedule against text, as well as summary of changes for any amendments. |
| Poor recruitment | Effective measurement of trade-offs for each site | Nonlinear modeling, such as with neural networks, to assess trade-offs site history, staff experience, investigator enthusiasm, available population, expected patient burden, and financial impact. Potential use of fuzzy logic to provide linguistic measurement descriptions. |
| | Improved use of funds | Optimizing communication/advertising to maximize cost effectiveness. Targeting communication to meet patient profile, including sentiment analysis. |
| | Ensuring appropriate eligibility criteria | NLP on prior publications to identify suitable criteria, and also criteria associated with other trial failures. |
| Patient burden/safety | Facilitating locating eligible patients | Database coordination, prompting investigators and patients when appropriate trials are available for specific patients. |
| | Enrolling patients who are likely to complete the trial | NLP and machine learning to profile patients based on prior data on who is more likely to complete a trial, reducing drop-outs. |
| | Minimize travel and wait times | Adaptive patient scheduling, also potentially turn-by-turn driving instructions, using evolutionary algorithms. Incorporate patient profiles to tailor site assignment/schedules to patient constraints |



Негативные последствия отчетов, не подтверждающих эффективность/безопасность

1. Отказ в регистрации действительно эффективного и безопасного препарата; - **недостатки организации КИ и честная констатация результата.**
2. **Попытка не искажая первичных данных манипулировать результатами:**
 - Регистрация неэффективного препарата на основании декоративной части отчета, констатирующей, что препарат эффективен и безопасен;
 - Регистрация препарата с неопределенной безопасностью по причине ограниченной выборки.



За скобками:

- Имитация полученных данных с изменением первичной документации, скрытой заменой или перебором пациентов, изменением результатов анализов и оценок, изменением кодов пациентов с целью получения необходимого результата.
- Несоответствие административным требованиям (например, отчет включает упоминание центров, в которых исследование не проводилось).
- Содержание отчета КИ, которое не требуется для подтверждения эффективности и безопасности.



Основные группы манипуляций

- Грубые изменения протокола, либо использование изначально несостоятельных протоколов;
- Умеренные изменения протокола, которые, в связи с его изначальноными недостатками приводят к методологической несостоятельности доказательств эффективности/безопасности;
- Ошибка стат. анализа данных (игнорирование необходимых поправок, применение необоснованных критериев оценки);
- Обращение к историческим данным, для замены результатов, полученных в группе сравнения;
- Игнорирование значимой гетерогенности групп после рандомизации*;
- Использование поисковых исследований в качестве подтверждающих; вариант игры в рулетку с целью найти «желанную цель» в виде 5% ошибки первого рода.



Гетерогенность

КОЛЛЕГИЯ ЕВРАЗИЙСКОЙ ЭКОНОМИЧЕСКОЙ КОМИССИИ

РЕКОМЕНДАЦИЯ
от 3 ноября 2020 г. N 19

О РУКОВОДСТВЕ
ПО ПРИМЕНЕНИЮ ПРИНЦИПОВ БИОСТАТИСТИКИ В КЛИНИЧЕСКИХ
ИССЛЕДОВАНИЯХ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ

Следует установить причину любой гетерогенности эффекта вмешательства, характеризующейся выраженными качественными взаимодействиями между эффектом вмешательства и клиническим центром. Если причина гетерогенности эффекта вмешательства не установлена, то **могут потребоваться дополнительные клинические исследования**, прежде чем эффект вмешательства можно будет предсказать с достаточной степенью надежности.



Примеры манипуляций в отчетах (ниже – подробней)

1. При подготовке отчета клинического исследования меняется первичный критерий эффективности;
2. После проведения исследования гипотеза превосходства меняется на гипотезу не меньшей эффективности с подбором границы признания не меньшей эффективности на основании полученных данных;
3. Исследования второй фазы трактуются достаточными для регистрации препарата по стандартным процедурам;
4. Метод статистического анализа не соответствует характеристике ПКТ (например, *категориальная шкала анализируется с использованием средних*);
5. Исследование завершается ранее запроотолированного момента анализа с приведением данных, демонстрирующих достижение цели;
6. Подмена статистической значимостью клинической значимости.



1. Не обоснованно меняется первичный критерий эффективности. Возможный вариант последовательности событий

1. Запланирован признанный и обоснованный первичный критерий эффективности, однако при получении результата нулевая гипотеза не отвергнута.
2. Перебираются возможные варианты первичных конечных точек и среди множества вариантов находят отреагировавшую в нужном направлении.
3. СОПРОВОЖДЕНИЕ: подача отчета в регистрационном досье сопровождается мощной проработкой масс-медиа для создания контекста очевидной эффективности препарата + применение аргументов далеких от научной методологии оценки результатов клинических исследований.



1. Не обоснованно меняется первичный критерий эффективности - Результат

- Истощение экспертного ресурса на длительное монотонное объяснение причин отказа.
- Мобилизация ресурса компании на создание внешнего фона очевидной возможности регистрации препарата
- Декларация тезисов:
 - «Критерий можно использовать для подтверждения эффективности и значимость продемонстрирована»
 - «Такое раньше регистрировалось»
 - «Препарат безопасен»
 - «Для подтверждения нужна нереально большая выборка»
 - «Мы уже подали препарат на регистрацию за рубежом»
 - «А вот, как раз, письмо ведущего специалиста с его подписью»
 - «Мы выбрали критерий, который имеет клиническую значимость»



1. Не обоснованно меняется первичный критерий эффективности – почему нет признания эффективности терапии, если выбран один из вторичных критериев?

Выбор из 5 возможных критериев (вторичных точек) при отсутствии реальной эффективности препарата дает возможность получить успех с 25% вероятностью $(1-0,95^5)$, при выбранной значимости для каждого из критериев 0,05.



2. После проведения исследования гипотеза превосходства меняется на гипотезу НМЭ с подбором границы признания не меньшей эффективности на основании полученных данных

VII. OVERALL CONCLUSIONS

The interpretation of superiority trials as non-inferiority trials and vice versa is best approached by expressing the results as a confidence interval for the difference between the test treatment and control. There is no fundamental problem associated with the use of this confidence interval as a basis for either mode of interpretation. For a well-designed and conducted trial, there are few difficulties connected with the change from non-inferiority to superiority that cannot be addressed by appropriate analysis. However, there are more severe difficulties associated with the switch from superiority to non-inferiority because of the possible need to find a basis for, and agree on, a margin of equivalence after seeing the outcome, and because of the inherent difficulties of non-inferiority trials. There are consequences for the design of a superiority trial in which non-inferiority might be an acceptable outcome.



The European Agency for the Evaluation of Medicinal Products
Evaluation of Medicines for Human Use

London, 27 July 2000
CPMP/EWP/482/99



COMMITTEE FOR PROPRIETARY MEDICINAL PRODUCTS
(CPMP)

POINTS TO CONSIDER ON SWITCHING BETWEEN SUPERIORITY
AND NON-INFERIORITY

«Однако, существуют серьезные проблемы, связанные с переходом от КИ превосходства к КИ НМЭ из-за необходимости найти основание и согласовать предел эквивалентности после получения результата, а также из-за присущих сложностей в трактовке таких исследований».



2. После проведения исследования гипотеза превосходства меняется на гипотезу НМЭ с подбором границы признания не меньшей эффективности на основании полученных данных

- Наиболее часто непреодолимым препятствием становится невозможность обосновать границу НМЭ.
- Отсутствует совокупность результатов КИ препарата сравнения с плацебо для обоснования границы признания НМЭ.



3. Исследования второй фазы трактуются достаточными для регистрации препарата по стандартным процедурам

- Поискное исследование выявляет различие с группой сравнения и на этой основе компания считает объем данных достаточным для легализации результата.
- Проблемы: недостаточная численность для оценки безопасности, отсутствие учета множественности сравнений, гетерогенность на малых выборках, необходимость подтверждения в повторном «эксперименте» (КИ).



3. Исследования второй фазы трактуются достаточными для регистрации препарата по стандартным процедурам

- Поискное исследование мини-макс выявляет превышение относительно заложенной границы, на этой основе компания считает объем данных достаточным для легализации результата (онкология)
- Проблемы: отсутствие группы сравнения во II фазе делает невозможным формирование четкого вывода об эффективности, т.к. конкретная популяция не может быть идентична популяциям исторического контроля.



3. Необходимость наличия подтверждающего исследования в рамках правил ЕЭК

КОЛЛЕГИЯ ЕВРАЗИЙСКОЙ ЭКОНОМИЧЕСКОЙ КОМИССИИ

РЕКОМЕНДАЦИЯ

от 3 ноября 2020 г. N 19

О РУКОВОДСТВЕ

ПО ПРИМЕНЕНИЮ ПРИНЦИПОВ БИОСТАТИСТИКИ В КЛИНИЧЕСКИХ
ИССЛЕДОВАНИЯХ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ

Подтверждающее клиническое исследование

13. Подтверждающие клинические исследования необходимы для получения надежного доказательства эффективности или безопасности исследуемого лекарственного препарата. В таких исследованиях ключевая интересующая гипотеза напрямую следует из основной цели исследования, она заранее сформулирована и является гипотезой, которая впоследствии проверяется по завершении исследования. В подтверждающем исследовании одинаково важно оценить с должной точностью (прецизионностью) величину эффектов, связанных с интересующим вмешательством, и сопоставить эти эффекты с их клинической значимостью.



3. Расхождение результатов 2 и 3 фаз

22 Case Studies Where Phase 2 and Phase 3 Trials had Divergent Results

Table of Contents

| | |
|---|----|
| I. Overview | 2 |
| II. Clinical Trials: Understanding Medical Product Testing | 2 |
| III. Flexibility in Clinical Trial Design | 3 |
| IV. Case Studies | 5 |
| A. Phase 3 Trials Demonstrating Lack of Efficacy in a Promising Experimental Therapy | 5 |
| 1. Bitopertin | 5 |
| 2. Brivanib | 6 |
| 3. Capsaicin Topical Patch (Qutenza) | 8 |
| 4. Darapladib | 9 |
| 5. Dexmecamylamine | 10 |
| 6. Exhale Drug-Eluting Stent | 11 |
| 7. Experimental HSV-2 Vaccine | 12 |
| 8. Glutamic Acid Decarboxylase Vaccine | 13 |
| 9. Imiquimod (Aldara 5% Cream) | 14 |
| 10. Iniparib | 15 |
| 11. Lithium | 16 |
| 12. MAGE-A3 vaccine | 17 |
| 13. NicVAX Vaccine | 18 |
| 14. Velimogene Aليplasmid (Allovetin-7) | 19 |
| B. Phase 3 Trials Demonstrating Lack of Safety in a Promising Experimental Therapy | 20 |
| 15. Olanzapine Pamoate (Zyprexa Relprevv) | 20 |
| C. Phase 3 Trials Demonstrating Lack of Efficacy and Lack of Safety in a Promising Experimental Therapy | 21 |
| 16. Aliskiren (Rasilez, Tekturna) | 21 |
| 17. CoStar Drug-Eluting Stent | 22 |
| 18. Figitumumab | 23 |
| 19. Recombinant Factor VIIa (NovoSeven) | 24 |
| 20. Semagacestat | 25 |
| 21. Torcetrapib | 26 |
| 22. V710 vaccine | 27 |
| V. Discussion | 28 |
| VI. Conclusions | 29 |
| Appendix A: RCTs and Clinical Trial Design Considerations | 31 |
| Appendix B: Methods | 32 |
| Appendix C: Summary Table | 34 |
| References | 36 |

FDA U.S. FOOD & DRUG
ADMINISTRATION

22 CASE STUDIES WHERE PHASE 2 AND PHASE 3 TRIALS HAD DIVERGENT RESULTS

January 2017



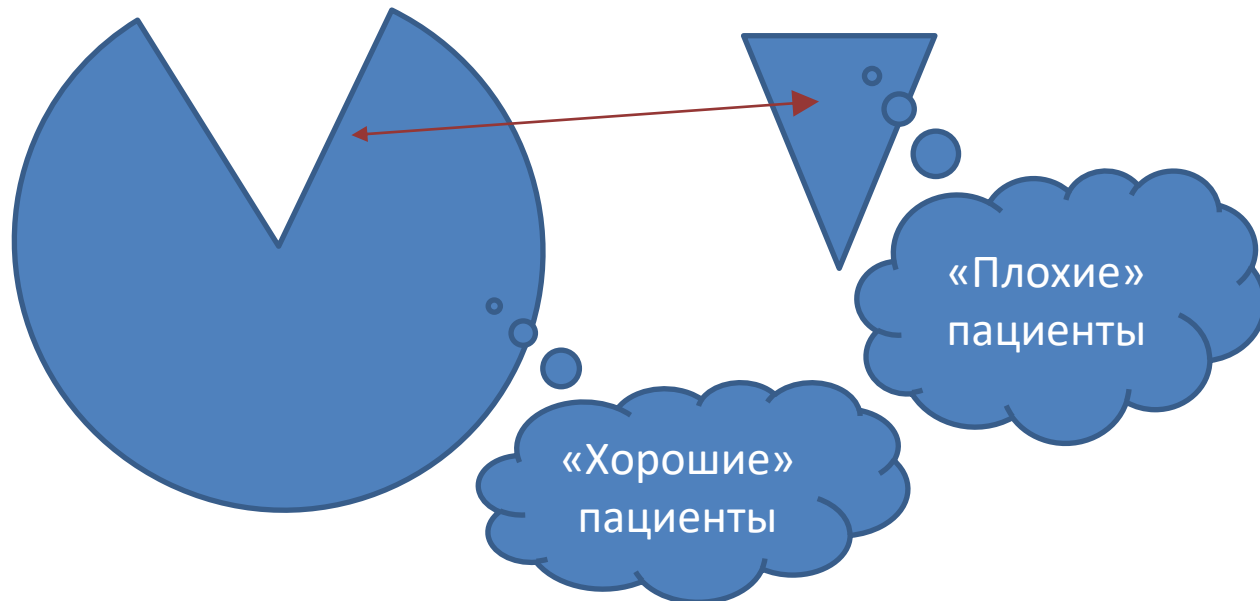
4. Метод статистического анализа не соответствует характеристике ПКТ или основным принципам

- Обычно это порождается нечеткостью статистической части протокола.
- Примеры: категориальная шкала (например, при COVID) анализируется с использованием средних);
- Расчет идет по особой популяции или исключительно по популяции РР или ИТТ, без учета особенностей исследования;



5. Исследование завершается на неполной выборке с приведением данных, демонстрирующих достижение цели

- При анализе данных выбирается «оптимальная» для необходимого результата популяция





6. Клиническая значимость \neq статистическая значимость

КОЛЛЕГИЯ ЕВРАЗИЙСКОЙ ЭКОНОМИЧЕСКОЙ КОМИССИИ

РЕКОМЕНДАЦИЯ
от 3 ноября 2020 г. N 19

О РУКОВОДСТВЕ ПО ПРИМЕНЕНИЮ ПРИНЦИПОВ БИОСТАТИСТИКИ В КЛИНИЧЕСКИХ ИССЛЕДОВАНИЯХ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ

13. Подтверждающие клинические исследования необходимы для получения надежного доказательства эффективности или безопасности исследуемого лекарственного препарата. В таких исследованиях ключевая интересующая гипотеза напрямую следует из основной цели исследования, она заранее сформулирована и является гипотезой, которая впоследствии проверяется по завершении исследования. В подтверждающем исследовании одинаково важно оценить с должной точностью (прецизионностью) величину эффектов, связанных с интересующим вмешательством, и сопоставить эти эффекты с их клинической значимостью.

15. Надежное доказательство для подтверждения соответствующих гипотез подразумевает, что результаты подтверждающих клинических исследований будут свидетельствовать о наличии у исследуемого лекарственного препарата клинической пользы. В связи с этим подтверждающие клинические исследования должны быть



6. Клиническая значимость \neq статистическая значимость

EVIDENCE-BASED DERMATOLOGY: STUDY

Statistical Significance and Clinical Relevance

The Importance of Power in Clinical Trials in Dermatology

Sachin S. Bhardwaj, MD; Fabian Camacho, MS; Amy Derrow, MS;
Alan B. Fleischer, Jr, MD; Steven R. Feldman, MD, PhD

May 7, 1997

Penciclovir Cream for the Treatment of Herpes Simplex Labialis

A Randomized, Multicenter, Double-blind, Placebo-Controlled Trial

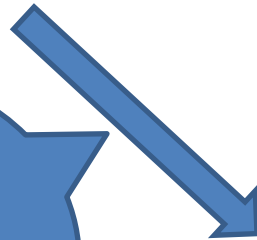
Spotswood L. Spruance, MD; Ted L. Rea, MD; Christopher Thoming, MD; [et al](#)

» Author Affiliations

JAMA. 1997;277(17):1374-1379. doi:10.1001/jama.1997.03540410052030

$P < 0,01$

Различие в
0,7 суток
????



Patients. —Otherwise healthy individuals with a history of frequent episodes of herpes simplex labialis. A total of 2209 patients were enrolled and given study medication, and 1573 initiated treatment for a recurrence.

Results. —Healing of classical lesions (vesicles, ulcers, and/or crusts) was 0.7 day faster for penciclovir-treated patients compared with those who received vehicle control cream (median, 4.8 days vs 5.5 days; hazard ratio [HR], 1.33; 95% confidence interval [CI], 1.18-1.49; $P < .001$). Pain (median, 3.5 days vs 4.1 days; HR, 1.22; 95% CI, 1.09-1.36; $P < .001$) and lesion virus shedding (median, 3 days vs 3 days; HR, 1.35; 95% CI, 1.10-1.64; $P = .003$) also resolved more quickly for penciclovir-treated patients compared with patients who applied the vehicle control.



Призыв к честности:

Данные не подтверждающие эффективность и безопасность методом, заложенным в протоколе, или очевидным подходом, описанным в руководствах, не следует компенсировать их «нечестной» обработкой и сомнительными манипуляциями.

Вывод Вашего «хорошего, но не эффективного» препарата приводит к вторичной волне вывода на рынок в качестве лекарственных средств фармацевтического балласта.



**Сделаешь новый
лекарственный
препарат?**





Спасибо за внимание!



научный центр
экспертизы средств
медицинского применения