

Критерии экспертизы программ

Термин «эффективная программа» (effective program) неоднозначен. Программа может очень эффективно решать какую-либо проблему или демонстрировать положительную динамику по какому-либо показателю (такому, например, как «снижение потребления алкоголя»), но при этом не улучшать другие параметры (такие как «повышенная тревожность» и т.п.). Чтобы профессионалам было проще искать программы, которые решают проблемы, имеющие отношение к их работе, Национальный реестр программ и практик с доказанной эффективностью (NREPP - The National Registry of Evidence-based Programs and Practices; США) составляет рейтинги программ на уровне социальных результатов/изменений в обществе (outcome level). Благодаря такому подходу мы достаточно точно определяем, какую именно потребность эффективно удовлетворяет программа.

В процессе изучения каждого значимого социального результата, вошедшего в доказательную базу оценочных исследований программы, аккредитованные эксперты NREPP пользуются специальным Инструментом рейтинга социальных результатов (The NREPP Outcome Rating Instrument). Социальные результаты программ оцениваются по четырем параметрам.

Эти четыре параметра охватывают следующие аспекты:

1. **Строгость (Rigor)**. Строгость является сильной стороной методологии исследования, и включает следующие элементы:
 - a. **Дизайн/набор в группы (Design/Assignment)**. Эксперты обращают внимание на то, как формировались группы участников – (1) группа, включающая тех, кто получал терапию/вмешательство, и (2) контрольная группа. Лучшие дизайнеры стремятся к тому, чтобы в момент начала интервенции эти две группы были сходными (или эквивалентными). Если группы различаются, то оценка не сможет выяснить, каким образом произошли изменения, зафиксированные через определенный период времени - под воздействием программы или в силу различий между группами.
 - b. **Изначальное распределение по группам (ITT - Intent-to-Treat / OGA – Original Group Assignment)**. В процессе изучения данного элемента эксперты смотрят, в какой мере анализ опирался на информацию о том, в какую группу изначально были направлены люди/кластеры - независимо от того, что они получили в итоге, и какой была степень данного вмешательства. Например, даже если некоторые участники, направленные в экспериментальную группу, так и не получили терапию/лечение/услугу, их все равно следует отнести к группе, получающей вмешательство.
 - c. **Статистическая точность (Statistical Precision)**. Данный элемент обращает внимание экспертов на размеры групп, сформированных для целей данного исследования. Чем больше размеры выборки, тем выше вероятность, что оценка способна распознать социальный эффект программы (impact).
 - d. **Претестовая эквивалентность (Pretest Equivalence)**. Изучая данный элемент, эксперты должны использовать материалы исследования, чтобы сделать вывод о существенности различий между группами по наблюдаемым переменным на этапе начального тестирования.
 - e. **Претестовая корректировка (Pretest Adjustment)**. Когда группы очень похожи, но не эквивалентны, исследователям следует учитывать изначально существующие различия и, в частности, в процессе анализа данных корректировать баллы начального тестирования. Такой подход помогает понять, являются ли какие-либо изменения в показателях социальных результатов следствием выполнения программы, или они наблюдаются в силу различий, изначально зафиксированных на этапе начального тестирования.
 - f. **Аналитический метод (Analysis Method)**. Исследователи могут анализировать информацию разными способами, поэтому очень важно, чтобы они выбрали аналитический метод, который наилучшим образом учитывает специфику данных.

- g. **Дополнительные угрозы внутренней валидности** (Other Threats to Internal Validity). Изучая этот элемент, эксперты стремятся установить дополнительные факторы, которые могли бы повлиять на динамику показателей социальных результатов наряду с воздействием, оказанным программой. Например, если в одной группе проводилось мероприятие, которое не имело отношения к интервенции и не распространялось на другую группу, то изменения в показателях социальных результатов могли оказаться следствием данной программы ИЛИ данного мероприятия. Точный ответ дать невозможно.
- h. **Надежность измерений** (Measurement Reliability). Данный элемент предлагает оценить согласованность показателей, например - выяснить, приводят ли изучаемые испытания к одинаковым результатам. (Чтобы присудить определенный балл качеству измерений, данный элемент следует рассматривать в сочетании с «валидностью».)
- i. **Валидность измерений** (Measurement Validity). Данный элемент предлагает выяснить, позволяет ли система измерений описать или «оцифровать» то, что планируется измерить, либо то, что намерены измерить исследователи. (Чтобы присудить определенный балл качеству измерений, данный элемент следует рассматривать в сочетании с «надежностью».)
- j. **Выбывание** (Attrition). Данный элемент предлагает выяснить, сколько участников вышли из состава одной или обеих групп, а также посмотреть, сколько участников включено в анализ. Если участники выбывают в ходе исследования, то это обстоятельство может отразиться на эквивалентности групп.
2. **Размер эффекта** (Effect Size). Размер эффекта – это способ измерения, позволяющий выяснить: удалось ли программе получить социальный эффект; насколько значительным он оказался; и каким был его характер - интервенция помогла группе, подвергнутой вмешательству, или причинила ей вред. Размеры эффекта вычисляются в том случае, если оценочные исследования предоставляют необходимые данные.
3. **Корректность выполнения программы** (Program Fidelity). Эксперты изучают результаты оценочных исследований, чтобы установить: выполнена ли программа в соответствии с изначальным планом, и получила ли вмешательство/услуги целевая аудитория. Данный параметр состоит из следующих элементов:
- a. **Пользователи услуг** (Service utilization). Данный элемент предлагает выяснить, в какой мере изначально указанная группа получила запланированные услуги. Иными словами, удалось ли программе выйти на целевую аудиторию?
- b. **Оказание услуг** (Service Delivery). Данный элемент предлагает выяснить, получили ли участники услуги или терапию конкретного вида, в надлежащем объеме, и должного качества. Иными словами, речь идет о том, в какой мере основные услуги или компоненты программы согласуются с тем, что указано в рамочной концепции.
4. **Рамочная концепция** (Conceptual Framework). Данный параметр имеет отношение к тому, насколько четко сформулированы компоненты программы. Он состоит из следующих элементов:
- a. **Цели программы** (Program Goals). Данный элемент оценивает четкость формулировки программных целей. Цель программы описывает изменения, которые стремится получить данная инициатива. Кроме того, чтобы считаться значимыми, цели должны быть реалистичными и достижимыми, равно как и результаты программы. Другими словами, должна существовать рациональная связь между тем, какая работа проводится в рамках программы, и что планируется получить в итоге.
- b. **Компоненты программы** (Program Components). Данный элемент предлагает выяснить, какое значение имеют виды деятельности или компоненты программы для достижения поставленной цели. Обычно программы включают в себя много разных компонентов или видов работ/услуг, однако ключевые компоненты имеют отношение к основным функциям и направлениям деятельности, которые являются обязательными условиями для получения

желаемых социальных результатов. Помимо этого, они непосредственно связаны с Теорией изменений данной программы.

- c. **Теория изменений** (Theory of Change). Данный элемент определяет степень правдоподобности теории воздействия на общество, разработанной для данной программы. Теория изменений должна объяснять, почему программа стимулирует процесс преобразований. Предположение о том, что программа создаст определенную пользу для целевой аудитории, зависит от появления причинно-следственного механизма, который начинается с взаимодействия целевого объекта с программой и заканчивается улучшением жизненных обстоятельств целевой группы населения. Например, программа может предлагать психотерапевтические услуги, которые улучшают навыки жизнедеятельности и снижают уровень стресса и тревожности.

В процессе рейтинга элементов на основе перечисленных параметров формируется заключение о сильных сторонах (строгости) методологии, корректности выполнения программы, размере эффекта и рамочной концепции. Каждому элементу, входящему в Инструмент рейтинга социальных результатов NREPP (за исключением размера эффекта), присваивается численное значение. Для обеспечения единообразия в процессе экспертизы параметры сопровождаются определениями и другими инструкциями, которыми рецензенты пользуются в ходе изучения элементов. Также, эксперты обращают внимание на любые другие сведения, которые стоит взять на заметку в виду их особой значимости. Предлагаем ознакомиться с [Процессом экспертизы](#) более подробно, а также узнать, как эти параметры используются для составления рейтинга социальных результатов.

Дополнительная информация о терминологии, используемой в описании параметров, вошедших в Инструмент рейтинга социальных результатов NREPP, содержится в [Глоссарии](#).

Источник: Управление делами в сфере профилактики и лечения наркотической зависимости и психических расстройств (The Substance Abuse and Mental Health Services Administration; США): http://nrepp.samhsa.gov/04e_reviews_program.aspx