DOI: 10.33380/2305-2066-2019-8-2-122-130
УДК 543.544; 615.074

Валидация аналитических методик: графические и расчетные критерии для оценки линейности методик на практике

Н. А. Эпштейн[1]*

1 – Центр регистрации и разработки лекарственных средств ООО «ИРВИН 2», 115446, Россия, г. Москва, Коломенский проезд, д. 13А
***Контактное лицо:** Эпштейн Наум Аронович. **E-mail:** naumepshtein@gmail.com
ORCID ID: orcid.org/0000-0001-8047-4078. **Researcher ID:** A-7114-2019
Статья получена: 16.04.2019. **Статья принята к печати:** 18.04.2019

Резюме

Введение. При валидации аналитических методик одной из обязательных характеристик является «Линейность». Вопросы, связанные с валидацией линейности методик, до сих пор являются актуальными.
Цель. Рассмотреть основные графические и расчетные критерии для подтверждения линейности аналитических методик и их ограничения, а также дать рекомендации.
Материалы и методы. Статистические расчеты выполнены с помощью программы MS Excel. Экспериментальные данные получены методом ВЭЖХ.
Результаты и обсуждение. Подробно рассмотрены основные критерии для доказательства/подтверждения линейности аналитических методик и их ограничения. Показано, что эти критерии не всегда могут дать надежную оценку линейности методики, указаны возможные причины этого и даны рекомендации.
Заключение. При валидации методик необходимо доказывать/подтверждать их линейность с помощью двух, а надежнее – трех критериев, одним из которых обязательно должна быть линейность графика зависимости измеряемого отклика (*Y*) от концентрации определяемого вещества (*C*). С другой стороны, показано, что не следует формально подходить к оценке нелинейности методики при использовании графических и расчетных критериев, основанных на математической статистике, так как они не учитывают возможность практической незначимости небольших отклонений от линейной зависимости *Y* от *C*.

Ключевые слова: валидация, линейность, методики, критерии приемлемости.

Конфликт интересов: конфликта интересов нет.

Для цитирования: Эпштейн Н. А. Валидация аналитических методик: графические и расчетные критерии для оценки линейности методик на практике. *Разработка и регистрация лекарственных средств*. 2019; 8(2): 122–130.

Validation of Analytical Procedures: Graphic and Calculated Criteria for Assessment of Methods Linearity in Practice

N. A. Epshtein[1]*

1 – Center of registration and development of medicines LLC «IRVIN 2», 13A, Kolomensky passage, Moscow, 115446, Russia
***Corresponding author:** Naum A. Epshtein. **E-mail:** naumepshtein@gmail.com
ORCID ID: orcid.org/0000-0001-8047-4078. **Researcher ID:** A-7114-2019
Received: 16.04.2019. **Accepted:** 18.04.2019

Abstract

Introduction. «Linearity» is one of the required characteristics when validating analytical procedures. The issues of the validation of linearity procedures are still relevant.
Aim. To consider the main graphical and calculated criteria for confirmation of the linearity of analytical procedures and their restrictions, as well as to give recommendations.
Materials and methods. Statistical calculations were performed using MS Excel. Experimental data were obtained by HPLC.
Results and discussion. The main criteria for proving/confirming the linearity of analytical procedures and their restrictions are considered in detail. It is shown that these criteria cannot always give a reliable assessment of the linearity of the procedure, the possible reasons for this are indicated and recommendations are given.
Conclusion. When validating the procedures, it is necessary to prove/confirm their linearity by using two, and, more reliably, three criteria, one of which must be the linearity of the plot of measured response (*Y*) vs concentration of the tested substance (*C*). On the other hand, we demonstrate here that a formal approach should not be used for the estimation of non-linearity of the procedure when using graphical and calculation criteria based on mathematical statistics, since they do not take into account the possibility of practical insignificance of small deviations from the linear dependence of *Y* on *C*.

Keywords: validation, linearity, procedures, acceptance criteria.

Conflict of interest: no conflict of interest.

For citation: Epshtein N. A. Validation of analytical procedures: graphic and calculated criteria for assessment of methods linearity in practice. *Drug development & registration*. 2019; 8(2): 122–130.

ВВЕДЕНИЕ

При валидации аналитических методик одной из обязательных характеристик является «Линейность» [1–8]. Согласно ICH «Линейность аналитической методики – это ее способность (в пределах заданного диапазона) получать результаты прямо пропорциональные концентрации (количеству) анализируемого вещества в образце» [1]. Цель валидации линейности методик, согласно фармакопее США, «состоит в том, чтобы иметь модель, линейную или нелинейную, которая близко (<точно>) описывает взаимосвязь концентрация – отклик» (под откликом понимается измеряемый сигнал или результат его обработки); «если линейность недостижима, можно использовать нелинейную модель» [2]. То есть речь идет о желательности, но не обязательности подтверждения линейной модели взаимосвязи концентрация – отклик при валидации методик. Подчеркнем, что *фактически суть теста «Линейность» заключается в проверке (до подтверждения правильности и прецизионности методики) возможности использования линейной модели как наиболее простой и предпочтительной на практике. В то же время по результатам теста «Линейность» может быть показана необходимость использования нелинейной модели, если нелинейная модель более близко (точно), чем линейная, описывает взаимосвязь концентрация – отклик и будет подтверждена правильность и прецизионность методики.*

Линейность методик подтверждают чаще всего в соответствии с ICH [1]: визуально графиком зависимости измеряемого отклика *Y* (например, площади пика в ВЭЖХ) от концентрации *C* определяемого вещества и приемлемым значением коэффициента корреляции *R* линейной регрессионной зависимости *Y* от *C*; приводят регрессионное уравнение или статистические характеристики регрессии. На практике применяют и другие критерии для оценки линейности/нелинейности зависимости *Y* от *C*, основанные на математической статистике [6–24], и для этого есть определенные причины, как будет видно из текста статьи. Неудивительно поэтому, что до сих пор остаются актуальными вопросы, связанные с критериями линейности аналитических методик как при практической валидации методик, так и при экспертизе валидационной документации.

Цель статьи: рассмотреть основные графические и расчетные критерии для подтверждения линейности аналитических методик и их ограничения, а также дать рекомендации.

МАТЕРИАЛЫ И МЕТОДЫ

Статистические расчеты, представленные в статье, выполнены с помощью программы MS Excel, в том числе с помощью статистического пакета, входящего в ее состав. Экспериментальные данные, приведенные в качестве примеров, получены методом ВЭЖХ на жидкостном хроматографе Waters e2695 с детектором Waters 2998 Photodiode Array Detector.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Прежде всего обсудим почему важна именно линейность методики, ведь известно, что для калибровочного графика можно использовать нелинейную модель зависимости Y от C [6, 11–22]. Это можно объяснить следующими причинами. Во-первых, возможностью использования наиболее простого метода – абсолютной калибровки (одной концентрации раствора стандарта) при «пренебрежимой постоянной систематической ошибке (величине отрезка, отсекаемого осью ординат)» [6, с. 102]; во-вторых, при рутинном анализе в случае линейного калибровочного графика достаточно использовать 2-3 точки в то время как при нелинейном калибровочном графике потребуется не менее 5 точек, а это большие затраты времени особенно в случае сравнительно медленных хроматографических методик анализа.

Необходимое количество экспериментальных точек и растворы, используемые для валидации линейности. Для валидации линейности методики в руководстве по валидации ICH рекомендуется использовать не менее 5 экспериментальных точек – концентраций определяемого вещества [1], что согласуется с рекомендацией из математической статистики [25, с. 190]. В то же время пяти точек может оказаться недостаточно для надежного вывода о линейности/нелинейности методики [6]. Для валидации линейности методики ICH допускает получение растворов необходимых концентраций путем разведения раствора стандарта. При этом игнорируется возможность влияния вспомогательных веществ, так как оно косвенно проверяется тестом «Правильность». В то же время, согласно ICH, для подтверждения линейности методики вместо разведения раствора стандарта можно использовать результаты, получаемые на смесях определяемого вещества с плацебо – при тестировании правильности методики. Это обычно сокращает объем работ и существенно повышает надежность выводов о линейности/нелинейности методики, так как учитывается влияние плацебо и увеличивается количество экспериментальных точек – не менее 9, распределенных по возможности равномерно в аналитическом диапазоне методики. На практике мы используем именно такой подход для методик определения содержания основных веществ.

Диапазон (аналитическая область) методики, для которого необходимо валидировать линейность, зависит от типа методики [1, 6]:

- *Для методик количественного определения основных веществ в фармацевтических субстанциях и лекарственных формах*: «не менее, чем от 80 до 120% от номинального содержания» (то есть кон-

центрация испытуемых растворов должна охватывать диапазон, соответствующий не менее 80,0–120,0% от номинального содержания определяемого вещества).

- *Для методик определения однородности дозирования*: «не менее, чем от 70 до 130% от номинального содержания, если для испытания не требуется более широкий интервал (например, для аэрозолей)».
- *Для методик испытаний на растворение*: «±20% (абсолютных) по отношению к верхней / нижней границам указанного в спецификации диапазона».
- *Для методик определения примесей:* «от порога игнорирования примесей (порога регистрации) [26] или предела количественного определения (LOQ) до не менее, чем 120% от нормируемого предела». В фармакопее США иное требование: «от 50% до не менее, чем 120% от нормируемого предела» [2].

Рассмотрим основные критерии для оценки линейности зависимости *Y* от *C*, применяемые при валидации аналитических методик и их ограничения.

Коэффициент корреляции R

При валидации аналитических методик для подтверждения их линейности чаще всего используют коэффициент корреляции Пирсона *R* и график зависимости измеряемого отклика *Y* от концентрации *C* определяемого вещества [1–6]. При этом в качестве критерия приемлемости для коэффициента корреляции R можно ориентироваться на $R \geq 0{,}999$ – для методик определения содержания основных веществ [5, 27] и $R \geq 0{,}98$ – для методик определения содержания примесей [4]. Встречаются и другие критерии приемлемости, например, для методик определения содержания основных веществ $R > 0{,}998$ для АФИ и $R > 0{,}990$ для препаратов [28], $R \geq 0{,}995$ [29]. Наиболее низкие допустимые (критические) значения *R*, принимая во внимание возможность округления чисел, указаны в ГФ XIV: «в большинстве случаев используют линейные зависимости, отвечающие условию $R \geq 0{,}99$, и только при анализе следовых количеств рассматривают линейные зависимости, для которых $R \geq 0{,}9$» [3]; по нашему опыту столь низкие допустимые значения *R* следует рассматривать как недостаточно надежные, приемлемые только при подтверждении правильности и прецизионности методики.

Приведенные выше допустимые значения *R* для определения содержания основных веществ существенно больше, чем критические значения коэффициента корреляции Пирсона R_{crit}, используемые для теоретический оценки наличия линейной связи между *Y* и *C*. Например, для n=5 точек экспериментальных данных для одностороннего теста при вероятности P=95,0% R_{crit}=0,805, при P=99,0% R_{crit}=0,934, а при увеличении количества точек значение R_{crit} только уменьшается (http://www.magniel.com/alex/spearman.html). Это указывает на эмпирический характер упомянутых выше рекомендаций по значениям *R* и может быть объяснено тем, что коэффициент корреляции Пирсона основывается на линейной модели взаимосвязи *Y–C* и, более того, характеризует лишь корреляционную связь *Y* и *C* без учета ее практической значимости [11, 12].

График остатков

В общем случае зависимость измеряемого отклика *Y* от *C* на самом деле может оказаться нелинейной даже при высоком значении коэффициента корреляции; характерный пример с R=0,99952 приведен в [4]. Поэтому в руководстве по валидации аналитических методик ICH рекомендуется использовать *дополнительный критерий линейности – график остатков* [1]. Как правило, он представляет собой графическую зависимость разности экспериментального и вычисленного по уравнению регрессии значения отклика (Y_{exp} – Y_{cal}) от концентрации *C*. Иногда вместо C используют Y_{cal} [9]. В качестве критерия приемлемости рассматривается отсутствие нелинейной тенденции в расположении точек на графике остатков: «график остатков должен демонстрировать случайный характер остатков в постоянном диапазоне при отсутствии какой-либо последовательности или закономерности» [6]. При УФ-детектировании рекомендуется использовать не более, чем 10-кратный диапазон концентраций [6], иначе рискуют столкнуться с нелинейной зависимостью *Y* от *C*.

Если пяти экспериментальных точек недостаточно для однозначного вывода о линейности или нелинейности зависимости *Y* от *C* по графику остатков, то следует увеличить количество экспериментальных точек или, что надежнее, для тестирования линейности методики использовать значения Y полученные в тесте «Правильность». С другой стороны, линейность методики – это понятие в значительной степени условное, о чем говорит многообразие критериев для оценки линейности при валидации методик. В связи с этим не следует формально подходить к оценке линейности/нелинейности методики при использовании критериев, основанных на математической статистике, в том числе графика остатков. Если как на рисунке 1, график остатков указывает на нелинейность методики с теоретической точки зрения, однако, на практике все экспериментальные точки располагаются на линии регрессии или близко к ней, а также очевидно (рисунок 1) или подтверждено расчетом, что нелинейная модель не будет более точно, чем линейная, описывать зависимость *Y* от *C*, то логично сделать вывод: методика линейная и подтвердить/проверить это результатами тестирования правильности и прецизионности методики.

Примечание. Условием для получения корректных графика остатков и обычной (невзвешенной) линейной регрессии является постоянство стандарт-

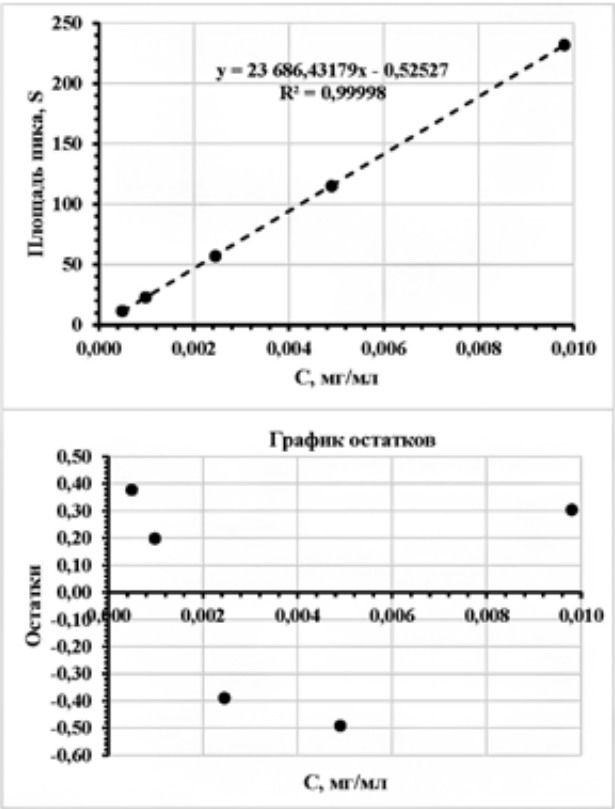

Рисунок 1. Результаты определения зависимости площади пика (S) примеси от ее концентрации (C) – вверху, и соответствующий ему график остатков – внизу, который формально указывает на нелинейную тенденцию, однако, этой тенденцией можно пренебречь с практической точки зрения (см. текст)

Figure 1. The results of determining the dependence of the peak area (S) of the impurity on its concentration (C) – above, and below the corresponding plot of the residuals that formally indicates a non-linear trend. However, this trend can be ignored from a practical perspective (see text)

ного отклонения остатков. Непостоянство стандартного отклонения остатков – гетероскедастичность, можно визуально распознать по клиновидной форме распределения остатков или путем сравнения разброса результатов параллельных измерений в зависимости от концентрации, а также методами математической статистики [6]. При выявлении гетероскедатичности для получения адекватной модели регрессии следует увеличить вклад/весомость меньших концентраций с помощью весовых коэффициентов (взвешенная линейная регрессия) [6].

График зависимости Y/C от C

Для оценки линейности методики вместо графика остатков иногда используют график зависимости *Y/C* от C, называемый «графиком чувствительности» [6]. В соответствии с формулой *Y=bC+a* можно ожидать, что на графике зависимости *Y/C* от C в случае линейности методики точки будут располагаться вблизи линии тренда ≈*b,* характеризующей чувствительность методики и почти параллельной оси концентраций. Однако следует учитывать, что необходимым условием для оценки линейности методики по графику зависимости *Y/C* от *C* является близость к нулю свободного члена регрессии *a*, так как *Y/C=b+a/C≈b* только при *a*≈0. Поэтому перед построением графика зависимости *Y/C* от *C* надо убедиться, что свободный член регрессионного уравнения *a* статистически незначим. Для этого можно использовать *t*-критерий Стьюдента. Значение *a* статистически незначимо (*a*≈0), если [30, с. 177]:

$$t_a = \frac{|a|}{SD_a} < t(P,\ f = n-2), \qquad (1)$$

где t_a – расчетное значение критерия Стьюдента; $|a|$ – абсолютное значение свободного члена линейной зависимости *Y* от *C*; SD_a – стандартное отклонение значения *a*; *t*(*P*, *f*=*n*–2) – табличное значение *t*-критерия Стьюдента при вероятности *P*=95%; *f* – число степеней свободы; *n* – количество экспериментальных точек на регрессионной прямой.

Критерии дисперсионного анализа (ANOVA): *F*-критерии и *p*-значение

Для проверки линейности методики иногда применяют дисперсионный анализ (ANOVA) вместе с визуальным подтверждением линейности зависимости *Y* от *C*. Основными являются два варианта ANOVA.

- **LOF-тест** («Lack of fit test») – тест со статистическим обсчетом результатов параллельных измерений. Его потенциальным преимуществом является независимость критерия приемлемости от принятой модели взаимосвязи *Y и C*. В этом варианте для каждой концентрации проводится несколько повторных измерений, например, по 3 инжекции в случае методики ВЭЖХ [10]. При этом рекомендуется брать не менее 8 концентраций в предполагаемом диапазоне действия методики [6].

В основе LOF-теста лежит рассмотрение определенных остатков (Residuals) и их квадратов, от которых зависит вариабельность полной ошибки измерения [6, 9, 24]:

$$Residual = y_{ij} - \hat{y}_i = Pure\ error + LOF\ error, \qquad (2)$$

$$Pure\ error = y_{i,j} - \overline{y}_i; \quad LOF\ error = \overline{y}_i - \hat{y}_i, \qquad (3)$$

где $y_{i,j}$ – результат *j*-го измерения в точке *i* (т. е. при концентрации C_i); $\bar{y}_i$ – среднее значение измерения и $\hat{y}_i$ – значение *y*, вычисленное по уравнению регрессии, в точке *i*. Из (3) видно, что *Residual* (остаток) представляет собой отклонение измеряемой величины $y_{i,j}$ от предсказываемой $\hat{y}_i$ по уравнению регрессии в точке *i*. *Residual* является суммой ***Pure error (PE)*** – случайной, «чистой ошибки», которая не зависит от принятой

регрессионной модели и представляет собой разность между результатом параллельного определения $y_{i,j}$ и средним значением параллельных определений $\bar{y}_i$ в точке *i*, и ***LOF*** – «ошибки неадекватности подбора (lack of fit error)» [6], представляющей собой разность между результатом параллельного определения $y_{i,j}$ и расчетным значением $\hat{y}_i$ по уравнению регрессии в точке *i*.

Вариабельность полной ошибки измерения SS_T можно представить в виде суммы квадратов для чистой ошибки (SS_{PE}), *LOF* (SS_{LOF}) и фактора, связанного с регрессионной моделью (SS_{REG}) [9]:

$$SS_T = \sum_{i=1}^{k}\sum_{j=1}^{m_i}\left(y_{i,j} - \bar{y}\right)^2 = \tag{4}$$

$$\underbrace{\underbrace{\sum_{i=1}^{k}\sum_{j=1}^{m_i}(y_{i,j} - \bar{y}_i)^2}_{SS_{PE}} + \underbrace{\sum_{i=1}^{k} m_i \cdot (\bar{y}_i - \hat{y}_i)^2}_{SS_{LOF}}}_{SS_{RES}} + \underbrace{\sum_{i=1}^{k} m_i \cdot (\hat{y}_i - \bar{y})^2}_{SS_{REG}},$$

где *k* – количество концентраций; m_i – число параллельных определений при *i*-ой концентрации.

Из (4) при рассмотрении остатков SS_{RES} следует, что SS_{LOF} зависит от принятой регрессионной модели, так как зависит от $\hat{y}_i$ и поэтому может быть использовано для оценки линейности методики. Условием для подтверждения линейности методики является выполнение требования к экспериментальному значению *F*-критерия [6, 9]:

$$F_{exp} = \frac{SS_{LOF}/(k-2)}{SS_{PE}/(n-k)} \le F(P,\ f_1,\ f_2) = F(P,\ k-2,\ n-k), \tag{5}$$

где *k* – количество концентраций; *n* – общее число измерений ($n=m \cdot k$ при количестве параллельных измерений $m_i=m$); $F(P, f_1, f_2)$ – критическое (табличное) значение *F*-критерия Фишера для вероятности *P* (обычно 95%) и степеней свободы $f_1=k-2$ и $f_2=n-k$. Формула (5) является критерием Фишера, так как рассматривается отношение двух дисперсий: $SS_{PE}/(n-k)$ – расчетная оценка дисперсии σ^2, $SS_{LOF}/(k-2)$ – расчетная оценка σ^2+bias^2 (где *bias* – смещение из-за «неадекватности подбора»), если используется неадекватная модель.

Отметим, что этот метод, как и другие варианты *ANOVA,* иногда может приводить к ошибочному выводу о нелинейности методики [10], и поэтому в случае F_{exp}> критического значения *F* следует посмотреть на график остатков, как это рекомендовано в [9]. Так в [10] при использовании метода ВЭЖХ было установлено, что в то время как график остатков указывал на линейную зависимость *Y* от *x*, LOF-тест показал, что $F_{exp}>F(P, k-2, n-k)$ и, как оказалось, это было обусловлено дополнительным источником вариации *Y*, связанным с пробоподготовкой. Это обосновали теоретически и ввели иной экспериментальный F-критерий (**модифицированный LOF-тест** 10]), который подтвердил линейную зависимость *Y* от *x*.

Мы тоже наблюдали ситуацию, сходную с [10], когда при очевидной линейной зависимости площади пика (S) от концентрации и отсутствии нелинейной тенденции на графике остатков (рисунок 2) оказалось, что F_{exp}=25,2455>*F*(95%, 8, 20)=2,4471, то есть LOF-тест ошибочно не подтвердил линейную зависимость *S* от *C*. Так как при этом была очень высокая сходимость результатов параллельных определений площади пика АФИ (значения относительного стандартного отклонения RSD находились в диапазоне от 0,04 до 0,22%), было сделано предположение о том, что именно низкие значения RSD делают определяющим вклад дополнительного источника вариации *S* и приводят к ошибочному выводу LOF-теста. Для проверки этого предположения в таблице Excel увеличили разброс результатов параллельных определений по сравнению с экспериментальными данными: RSD от 0,54 до 1,03%. В результате увеличения RSD значение *Fexp* уменьшилось до F_{exp}=1,5331<*F*(95%, 8, 20)=2,4471 и LOF-тест подтвердил линейную зависимость *Y* от *C* как и ожидалось. При этом лишь незначительно изменились параметры регрессионной зависимости *Y* от *C* и график остатков (рисунок 2). Из полученных результатов следует, что при высокой воспроизводимости результатов измерений – *RSD* меньше ≈0,5% в нашем случае, немодифицированный LOF-тест может приводить к ошибочному выводу о нелинейности аналитической методики. Наоборот, иногда при плохой воспроизводимости результатов параллельных определений можно получить $F_{exp} \le F(P, f_1, f_2)$ и при этом методика может обладать необходимой прецизионностью [24, p. 103].

- **«F/p-тест» линейной зависимости** – тест без статистического обсчета результатов параллельных измерений. В качестве исходных данных используют значения Y экспериментальные (Y_{exp}) и рассчитанные по линейному уравнению регрессии (Y_{cal}). Теоретическая основа этой процедуры подробно рассмотрена в Internet: https://excel2.ru/articles/proverka-znachimosti-regressii-s-pomoshchyu-dispersionnogo-analiza-f-test с примером в Excel https://excel2.ru/sites/default/files/node/808/edit/_f-test_for_simple_regression.xlsx.

Для подтверждения линейности зависимости *Y* от *C* требуется, чтобы расчетная величина значения F-критерия (F_o), получаемая с помощью функции =ЛИНЕЙН(диапазон Y;диапазон X;ИСТИНА;ИСТИНА):

$$F_o = \frac{SSR/1}{SSE/(k-2)}, \tag{6}$$

где

$$SSR = \sum_{i=1}^{k}(\hat{y}_i - \bar{y})^2;\ SSE = \sum_{i=1}^{k}(y_i - \hat{y}_i)^2, \tag{7}$$

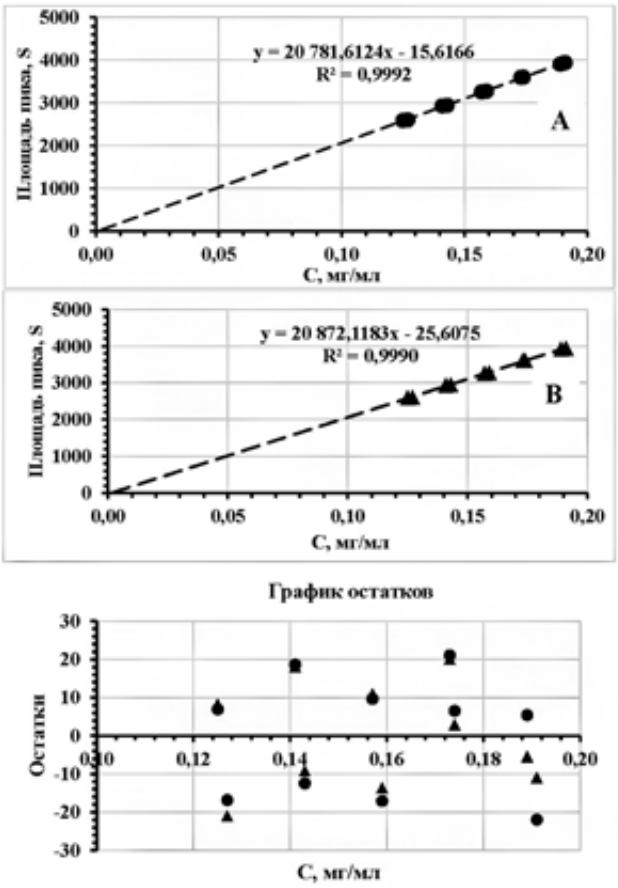

Рисунок 2. Экспериментальные (*A*) и скорректированные (*B*) – путем увеличения относительного стандартного отклонения (RSD) параллельных измерений, результаты определения зависимости площади пика (S) АФИ от концентрации (C). В случае (A) из-за очень высокой воспроизводимости параллельных измерений (RSD от 0,04 до 0,22%) LOF-тест не подтвердил линейность, хотя точки на графике остатков (●) указывали на линейность методики. При увеличении RSD параллельных измерений до 0,54–1,03% – случай (B), имело место лишь незначительное изменение параметров регрессионной зависимости Y от C и графика остатков (▲) по сравнению с (A), однако, этого оказалось достаточно, чтобы LOF-тест подтвердил линейность методики (пояснения см. в тексте).

Figure 2. Experimental (*A*) and corrected by increasing the relative standard deviation (RSD) parallel measurements – (*B*) results of determining the dependence of the peak area (S) of the API on its concentration. In case (A), due to the very high reproducibility of parallel measurements (RSD, 0.04 to 0.22%) the LOF test did not confirm linearity, although the points on the plot of the residuals (●) pointed to the linearity of the procedure. When increasing RSD of parallel measurements up to 0.54–1.03% – case (B), there was only a slight change in the parameters of the regression dependence of Y against C and of the plot of the residuals (▲) compared to (A). However, this turned out to be sufficient so that the LOF test to confirm the linearity of the method (see text for explanation).

была больше критического значения критерия Фишера F_{crit} (вычисляемого с использованием функции =F.ОБР.ПХ(α,1, k – 2), где k – количество экспериментальных данных – точек на графике, *a* расчетное значение критерия *p* (вычисляемого с использованием функции =F.РАСП.ПХ(F_o; 1; k – 2) было бы меньше заданного уровня значимости α (обычно α=0,05 или α=0,01, что соответствует вероятности 95% или 99% соответственно)[1]. Отметим, что *SSR* соответствует SS_{REG} в (4), а *SSE* соответствует SS_{LOF} при m_i=1 и i=1. Использование этого варианта ANOVA обычно не вызывает затруднений, однако, он основан на линейной модели, а это снижает надежность его выводов. Например, для регрессионной зависимости *Y* от *C* на рисунке 1, в отличие от графика остатков (нелинейность), результаты этого теста, полученные в Excel, указали на отсутствие нелинейности:

	ANOVA – оценка линейности		
	F-критерий	p	
F_o	1,45956E+05		
α=	0,5		
=F.ОБР.ПХ(α, 1, k – 2)	10,12796	3,95484E–08	=F.РАСПЮПХ(F_o; 1; k – 2)
Линейность	ИСТИНА	ИСТИНА	

Остаточное стандартное отклонение SD_0 и индекс корреляции R_c

На Украине разработан альтернативный ICH подход для валидации линейности методик [8], основанный на использовании нормализованных координат и предельно допустимой неопределенности методики анализа [7, 8]. При валидации линейности методик в качестве контролируемых характеристик рассматривают остаточное стандартное отклонение SD_0 и индекс корреляции R_c. Условием для применимости SD_0 и подхода в целом, является статистическая/практическая незначимость свободного члена регрессионной зависимости *a* (то есть $a \approx 0$). Используют следующий алгоритм:

Методом наименьших квадратов рассчитывают линейную регрессионную зависимость между нормализованными экспериментальными значениями отклика Y_n и концентрации C_n:

$$Y_{ni}=bC_{ni}+a, \qquad (8)$$

где $Y_{ni}=Y_i/Y_{st}$, $C_{ni}=C_i/C_{st}$, индексы *i* и *st* относятся соответственно к *i*-му испытуемому раствору и к стандартному раствору, индекс *n* указывает на то, что рассматриваются нормализованные значения. Затем проверяют статистическую/практическую незначимость свободного члена регрессионного уравнения (8). После подтверждения того, что $a \approx 0$, проверяют выполнение требования к остаточному стандартному отклонению SD_0 по формуле:

$$SD_0 / b \le \frac{\Delta_{As}}{t(95\%,\ n-2)}, \qquad (9)$$

[1]Англоязычный эквивалент F.ОБР.ПХ(α, 1, k – 2) это F.INV.RT(α, 1, k – 2), а F.РАСП.ПХ(F_{crit}; 1; k – 2) это F.DIST.RT(F_{crit}; 1; k – 2).

в которой SD_0 – остаточное стандартное отклонение, которое можно определить в Excel с помощью функции =СТОШYX(диапазон Y; диапазон X); Δ_{As} – предельно допустимая неопределенность методики анализа, вычисляемая по простым формулам [7, 8]; $t(95\%, n-2)$ – значение одностороннего t-критерия Стьюдента для вероятности 95% и числа степеней свободы $n-2$; n – объем выборки (число экспериментальных точек на графике)[1].

При выполнении условия (9) вычисляют индекс корреляции по формуле:

$$R_c = \sqrt{1 - \frac{SD_0^2}{RSD_{Range}^2}}, \qquad (10)$$

в которой SD_0 – остаточное стандартное отклонение, а RSD_{Range} – стандартное отклонение концентраций, вычисляемое по формуле:

$$RSD_{Range} = \sqrt{\sum \frac{(C_{ni} - \bar{C}_n)^2}{(n-1)}} \cdot 100\%, \qquad (11)$$

где C_{ni} и $\bar{C}_n$ – соответственно концентрация i-го раствора и средняя концентрация растворов в нормализованных координатах, n – число экспериментальных точек. Затем проверяют выполнение требования к индексу корреляции:

$$R_c \geq \sqrt{1 - \left(\frac{\Delta_{As}/t(95\%,\ n-2)}{RSD_{Range}} \right)^2}. \qquad (12)$$

Методика рассматривается как линейная при выполнении требований (9) и (12). Заметим, что индекс корреляции R_c не зависит от модели $Y=f(C)$ и поэтому имеет существенное преимущество перед коэффициентом корреляции Пирсона R при оценке линейности методик. В книге [8] этот подход наиболее подробно изложен с теоретической и с практической точки зрения и приведена наиболее полная таблица с критическими значениями рассмотренных выше критериев (правые части формул (12) и (9) – при b≈1), облегчающая проведение валидации методик для тестов «Количественное определение», «Однородность дозирования» и «Растворение». Несмотря на то, что этот подход основывается на теоретических предпосылках и использует критерии «практической незначимости», он по результатам поиска в Internet применяется пока только при валидации методик в Украине и частично в РФ и в некоторых других постсоветских странах; большинство исследователей используют рекомендации ICH/USP [1, 2].

[1]Англоязычный эквивалент СТОШYX(диапазон Y; диапазон X) это STEYX(known_y's, known_x's).

Дополнительные рекомендации:

- Если при валидации методики будет установлена ее нелинейность, следует попробовать уменьшить диапазон или изменить область рассматриваемых концентраций; использовать, при необходимости, взвешенную линейную регрессию.
- Для проверки того, что определенная точка экспериментальных данных является «выбросом» относительно регрессионной модели можно использовать t-тест [11]. Вычисляют экспериментальное значение t-критерия:

$$t_{exp} = \frac{\left| y_i - \hat{y}_i \right|_{max}}{SD_0 \cdot \sqrt{1 + \frac{1}{n} + \frac{(y_j - \bar{y})^2}{(n-1) \cdot SD_y^2}}}, \qquad (13)$$

где $\left| y_i - \hat{y}_i \right|_{max}$ – максимальное абсолютное значение разности между экспериментальным y_i и расчетным $\hat{y}_i$ (по регрессионной зависимости) значением отклика среди всех точек данных (от i=1 до n); n – количество экспериментальных точек; y_j – экспериментальное значение отклика в точке «выброса» j; $\bar{y} = \sum y_i / n$; SD_0 – остаточное стандартное отклонение и SD_y – стандартное отклонение экспериментальных значений отклика y, вычисляемые по формулам:

$$SD_0 = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}}, \quad SD_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}. \qquad (14)$$

Сравнивают t_{exp} с табличным значением критерия Стьюдента t (95%, $f=n-2$). Если t_{exp} больше табличного значения t (95%, $f=n-2$), то с вероятностью 95% эта точка не соответствует принятой регрессионной модели и является «выбросом», требующим повторного определения y_j. Альтернативой этому тесту является F-тест, приведенный в [31].

ЗАКЛЮЧЕНИЕ

Рассмотрены основные критерии для оценки линейности при валидации аналитических методик и их ограничения. Показано, что эти критерии по отдельности не могут дать надежную оценку линейности методики. Поэтому на практике при валидации методик следует доказывать/подтверждать их линейность с помощью двух, а надежнее – трех критериев, одним из которых должен быть график зависимости измеряемого отклика Y от концентрации определяемого вещества. С другой стороны, как показано выше, не следует формально подходить к оценке нелинейности методики при использовании графических и расчетных критериев, основанных на математической статистике, так как они не учитывают возможность практической незначимости небольших отклонений от линейной зависимости Y от C.

ЛИТЕРАТУРА

1. ICH Harmonised Tripartite Guideline. Validation of Analytical Procedures: Text and Methodology Q2(R1). ICH, 2005; 13.
2. United States Pharmacopoeia. <1225> Validation of compendial procedures. USP40–NF35. 2017; 1640–1646.
3. Государственная фармакопея Российской Федерации. XIV издание. Москва. 2018; 1: 276–288.
4. Green J. M. A Practical Guide to Analytical Method Validation. *Analytical Chemistry News & Features*. 1996; 1: 305A–309A. DOI: https://doi.org/10.1021/ac961912f.
5. Эпштейн Н. А. Оценка пригодности (валидация) ВЭЖХ методик в фармацевтическом анализе (Обзор). *Химико-фармацевтический журнал*. 2004; 38(4): 40–56. DOI: https://doi.org/10.1023/b:phac.0000038422.27193.6c
6. Эрмер Й., Миллер Д. Х. Мак Б. Валидация методик в фармацевтическом анализе. Примеры наилучших практик. М.: *ВИАЛЕК*. 2013; 512.
7. Руководство для предприятий фармацевтической промышленности. Методические рекомендации. Ч. I. Руководство по валидации методик анализа лекарственных средств / Под ред. Н. В. Юргеля, А. Л. Младенцева, А. В. Бурдейна и др. М.: Изд-во *«Спорт и культура – 2000»*. 2007; 5–92.
8. Гризодуб А. И. Стандартизованные процедуры валидации методик контроля качества лекарственных средств. Харьков: *Государственное предприятие «Украинский научный фармакопейный центр качества лекарственных средств»*, 2016; 396.
9. Massart D. L., Vandeginste B. G. M., Buydens L. M. C., De Jong S., Lewi P. J., Smeyers-Verbeke J. Straight Line Regression and Calibration, in: Handbook of Chemometrics and Qualimetrics: Part A, volume 20A of Data Handling in Science and Technology. Amsterdam, Netherlands: *Elsevier*. 1998; 171–230. DOI: https://doi.org/10.1016/S0922-3487(97)80038-X.
10. Vial J., Jardy A. Taking into account both preparation and injection in HPLC linearity studies. *J. Chromatogr. Sci*. 2000; 38: 189–194.
11. Burke S. Regression and Calibration. *LC-GC Europe Online Supplement statistics and data analysis*. 2001: 13–18.
12. Kiser M. M., Dolan J. W. Selecting the Best Curve Fit. *LC-GC North America*. 2004; 22(2): 138–143.
13. Mark H., Workman J. Chemometrics in Spectroscopy. How to Test for Non-linearity. *Spectroscopy*. 2005; 20(9): 26–35.
14. Mark H., Workman J. Linearity in calibration: Other tests for non-linearity // *Spectroscopy*. 2005; 20(4): 38–39.
15. Scheilla Vitorino Carvalho de Souza, Junqueira R. G. A procedure to assess linearity by ordinary least squares method. *Analytica Chimica Acta*. 2005; 552(1-2): 25–35. DOI: https://doi.org/10.1016/j.aca.2005.07.043.
16. Mark H., Workman J. Chemometrics in Spectroscopy. Linearity in Calibration: Quantifying Nonlinearity, Part II. *Spectroscopy*. 2006; 21(1): 44-54.
17. Bruggemann L., Quapp W., Wennrich R. Test for non-linearity concerning linear calibrated chemical measurements. *Accred Qual Assur*. 2006; 11: 625–631. DOI: https://doi.org/10.1007/s00769-006-0205-x.
18. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part I: theory of nonlinearity and quantification of curvature. *Bioanalysis*. 2015; 7 (14): 1731–1743. DOI: https://doi.org/10.4155/bio.15.103.
19. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part II: a Monte Carlo justification for the use of nonlinear regressions. *Bioanalysis*. 2015; 7 (14): 1745–1761. DOI: https://doi.org/10.4155/bio.15.104.
20. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part III: examples of nonlinear drug assays. *Bioanalysis*. 2015; 7 (14): 1763–1774. DOI: https://doi.org/10.4155/bio.15.105.
21. Raposo F. Evaluation of analytical calibration based on least-squares linear regression for instrumental techniques: A tutorial review. *TrAC Trends in Analytical Chemistry*. 2016; 77: 167–185. DOI: https://doi.org/10.1016/j.trac.2015.12.006.
22. Rawski R. I., Sanecki P. T., Kijowska K. M., Skital P. M., Saletnik D. E. Regression Analysis in Analytical Chemistry. Determination and Validation of Linear and Quadratic Regression Dependencies. *S. Afr. J. Chem*. 2016; 69: 166–173. DOI: https://doi.org/10.17159/0379-4350/2016/v69a20.
23. Мымриков Антон. Действительно ли ваша калибровка является прямой линией? Available at: https://pharm-community.com/2017/8336/ (accessed 25.12.2017).
24. Chemometrics in Chromatography ed. by Komsta L., Vander Heyden Y., Sherma J. Boca Raton: *CRC Press*. 2018; 506. DOI: https://doi.org/10.1201/9781315154404.
25. Доерфель К. Статистика в аналитической химии. М.: *Мир*. 1966; 211. DOI: https://doi.org/10.1002/bimj.19690110612.
26. Эпштейн Н. А. О допустимых значениях порога игнорирования примесей и отношения сигнал/шум при проверке чувствительности хроматографической системы. *Ведомости НЦЭСМП*. 2017; 7(2): 85–91.
27. Reviewer Guidance: Validation of Chromatographic Methods. Center for Drug Evaluation and Research (CDER). Washington. 1994. Available at: https://www.fda.gov/downloads/drugs/guidances/ucm134409.pdf.
28. Rosario LoBrutto and Tarun Patel. Method Validation in HPLC for Pharmaceutical Scientists. Ed. By Kazakevich Y.V., LoBrutto R. Hoboken, New Jersey: *John Wiley & Sons, Inc*. 2007; 455–502. DOI: https://doi.org/10.1002/9780470087954.ch9.
29. Gao Quanyin and Sanvordeker D.R. Analytical methods development and methods validation for oral solid dosage forms in Generic Drug Product Development: Solid Oral Dosage Forms, 2-nd Ed. Shargel, Leon: *CRC Press/Taylor & Francis Group*. 2014; 31–50.
30. Дерффель К. Статистика в аналитической химии. М.: *Мир*, 1994; 268.
31. Njaka N. A., Elise O. R., Herinirina N. R., Lucienne V. R., Manovantsoatsiferana H., R. A., Randrianarivony E. Dealing with Outlier in Linear Calibration Curves: A Case Study of Graphite Furnace Atomic Absorption Spectrometry. *World Journal of Applied Chemistry*. 2018; 3 (1): 10–16.

REFERENCES

1. ICH Harmonised Tripartite Guideline. Validation of Analytical Procedures: Text and Methodology Q2(R1). ICH, 2005; 13.
2. United States Pharmacopoeia. <1225> Validation of compendial procedures. USP40–NF35. 2017; 1640–1646.
3. State Pharmacopoeia of the Russian Federation XIV edition, 2018, V.1; 276-288 (in Russ.).
4. Green J. M. A Practical Guide to Analytical Method Validation. *Analytical Chemistry News & Features*. 1996; 1: 305A-309A. DOI: https://doi.org/10.1021/ac961912f.
5. Epstein N. A. Evaluation of the suitability (validation) of HPLC techniques in pharmaceutical analysis (Review). Chemical Pharmaceutical Journal. 2004; 38 (4): 40–56. DOI: https://doi.org/10.1023/b:phac.0000038422.27193.6c
6. Ermer Y., Miller D.H. Mac B. Validation of methods in pharmaceutical analysis. Examples of best practices. M.: *VIALEK*. 2013; 512 (in Russ.).
7. Guide for enterprises of the pharmaceutical industry. Guidelines. Part I. Guidelines for the validation of methods for analyzing drugs. Ed. N. V. Yurgelya, A. L. Mladentseva, A. V. Burdeina i dr. M.: *Publishing house «Sport and Culture – 2000»*. 2007; 5–92 (In Russ.).
8. Grizodub A. I. Standardized procedures for the validation of drug quality control methods. Khar'kov: Gosudarstvennoe predpriyatie *«Ukrainskiy nauchnyy farmakopeynyy tsentr kachestva lekarstvennykh sredstv»*. 2016; 396 (in Russ.).
9. Massart D. L., Vandeginste B. G. M., Buydens L. M. C., De Jong S., Lewi P. J., Smeyers-Verbeke J. Straight Line Regression and Calibration, in: Handbook of Chemometrics and Qualimetrics: Part A, volume 20A of Data Handling in Science and Technology. Amsterdam, Netherlands: *Elsevier*. 1998: 171–230. DOI: https://doi.org/10.1016/S0922-3487(97)80038-X.
10. Vial J. and Jardy A. Taking into account both preparation and injection in HPLC linearity studies. *J. Chromatogr. Sci*. 2000; 38:189–194.
11. Burke S. Regression and Calibration. *LC-GC Europe Online Supplement statistics and data analysis*. 2001: 13–18.
12. Kiser M. M., Dolan J. W. Selecting the Best Curve Fit. *LC-GC North America*. 2004; 22(2): 138–143.
13. Mark H., Workman J. Chemometrics in Spectroscopy. How to Test for Non-linearity. *Spectroscopy*. 2005; 20(9): 26–35.

14. Mark H., Workman J. Linearity in calibration: Other tests for non-linearity // *Spectroscopy*. 2005; 20(4):38–39.
15. Scheilla Vitorino Carvalho de Souza, Junqueira R. G. A procedure to assess linearity by ordinary least squares method. *Analytica Chimica Acta*. 2005; 552(1-2): 25–35. DOI: https://doi.org/10.1016/j.aca.2005.07.043.
16. Mark H., Workman J. Chemometrics in Spectroscopy. Linearity in Calibration: Quantifying Nonlinearity, Part II. *Spectroscopy*. 2006; 21(1): 44–54
17. Bruggemann L., Quapp W., Wennrich R. Test for non-linearity concerning linear calibrated chemical measurements. *Accred Qual Assur*. 2006; 11: 625–631. DOI: https://doi.org/10.1007/s00769-006-0205-x.
18. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part I: theory of nonlinearity and quantification of curvature. *Bioanalysis*. 2015; 7(14): 1731–1743. DOI: https://doi.org/10.4155/bio.15.103.
19. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part II: a Monte Carlo justification for the use of nonlinear regressions. *Bioanalysis*. 2015; 7(14): 1745-1761. DOI: https://doi.org/10.4155/bio.15.104.
20. Burrows J, Watson K. Linearity of chromatographic systems in drug analysis part III: examples of nonlinear drug assays. *Bioanalysis*. 2015; 7(14): 1763–1774. DOI: https://doi.org/10.4155/bio.15.105.
21. Raposo F. Evaluation of analytical calibration based on least-squares linear regression for instrumental techniques: A tutorial review. T*rAC Trends in Analytical Chemistry*. 2016; 77: 167–185. DOI: https://doi.org/10.1016/j.trac.2015.12.006.
22. Rawski R. I., Sanecki P. T., Kijowska K. M., Skital P. M., Saletnik D. E. Regression Analysis in Analytical Chemistry. Determination and Validation of Linear and Quadratic Regression Dependencies. *S. Afr. J. Chem*. 2016; 69: 166–173. DOI: https://doi.org/10.17159/0379-4350/2016/v69a20.
23. Mymrikov Anton. Is your calibration a straight line? Available at: https://pharm-community.com/2017/8336/ (accessed 25.12.2017).
24. Chemometrics in Chromatography ed. by Komsta L., Vander Heyden Y., Sherma J. Boca Raton: *CRC Press*. 2018; 506. DOI: https://doi.org/10.1201/9781315154404.
25. Doerffel K. Statistik in der analytischen Chemie. Leipzig: VEB Deutscher Verlag für Grundstoffindustrie. 1966; 211. DOI: https://doi.org/10.1002/bimj.19690110612.
26. Epshtein N.A. About admissible values of disregard limits for impurities and signal-to-noise ratio when checking chromatographic system sensitivity // *The Bulletin of the Scientific Center for Expert Evaluation of Medicinal Products*. 2017; 7(2): 85–91 (In Russ.).
27. Reviewer Guidance: Validation of Chromatographic Methods. Center for Drug Evaluation and Research (CDER). Washington. 1994. Available at: https://www.fda.gov/downloads/drugs/guidances/ucm134409.pdf.
28. Rosario LoBrutto and Tarun Patel. Method Validation in HPLC for Pharmaceutical Scientists. Ed. By Kazakevich Y. V., LoBrutto R. Hoboken, New Jersey: John Wiley & Sons, Inc. 2007; 455–502. DOI: https://doi.org/10.1002/9780470087954.ch9.
29. Gao Quanyin and Sanvordeker D.R. Analytical methods development and methods validation for oral solid dosage forms in Generic Drug Product Development: Solid Oral Dosage Forms, 2-nd Ed. Shargel, Leon: CRC Press/Taylor & Francis Group. 2014; 31–50.
30. Doerffel K. Statistik in der analytischen Chemie. Leipzig: Deutscher Verlag für Grundstoffindustrie GmbH: 1990; 256.
31. Njaka N. A., Elise O. R., Herinirina N. R., Lucienne V. R., Manovantsoatsiferana H., R.A., Randrianarivony E. Dealing with Outlier in Linear Calibration Curves: A Case Study of Graphite Furnace Atomic Absorption Spectrometry. *World Journal of Applied Chemistry*. 2018; 3 (1): 10–16.