

Quality assurance in clinical trials—the role of pathology

Christoph Röcken¹

Received: 6 August 2015 / Accepted: 23 September 2015 / Published online: 1 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In the last two decades, our knowledge about cancer genetics and cancer biology increased exponentially. Deep sequencing now allows rapid and cost-effective analysis of entire cancer genomes. Dysregulation of cell growth, cell survival, tissue homeostasis, and immune surveillance have been recognized as hallmarks of cancer. In parallel, diagnostic surgical pathology has been harmonized and consensus diagnostic criteria for cancer classification have been developed by initiatives of the World Health Organization, the International Agency for Research on Cancer, and the Union for International Cancer Control. Pharmaceutical companies developed novel drugs targeting specific molecules in signaling pathways, which has allowed the development of the concept of precision medicine. Now, we are facing a large number of clinical trials which bring together these advances and will explore efficacy of novel treatment regimens. Assessment of the efficacy of a new drug is often coupled with the simultaneous assessment of the capacity of tissue-based biomarkers to predict response of individual patients (companion diagnostics/precision medicine). Patients with histologically similar tumors might respond differently to the same drug. This review summarizes the diverse roles played by surgical pathologists involved in clinical trials, with a special focus on quality assurance of diagnostic, laboratory, and reporting standards.

Keywords Clinical trials · Targeted therapy · Companion diagnostics · Precision medicine · Surgical pathology

✉ Christoph Röcken
christoph.roecken@uk-sh.de

¹ Department of Pathology, Christian-Albrechts-University, Arnold-Heller-Str. 3, Haus 14, 24105 Kiel, Germany

Background

When the search terms “cancer” and “histology” were entered in the search engine of the US National Institutes of Health service (www.clinicaltrials.gov) at the end of June of 2015, more than 15,000 interventional clinical trials came up. This staggering number illustrates how often interventional studies use or rely on information provided by histological data. Of the drugs approved in 2013, 40 % target cancer. In the last 20 years, improvements in conventional chemotherapeutic drugs, which include among others alkylating agents, antimetabolites, and topoisomerase inhibitors, have been slow and limited in impact. However, significant improvements have been achieved in a few cancer types using either antihormonal medication or drugs targeting molecules in specific signaling pathways, e.g., receptor tyrosine kinases and their ligands. Conventional chemotherapy requires accurate histological diagnosis of cancer subtype and the assessment of tumor stage, while “targeted” drugs often need predictive biomarker testing on a tissue specimen. Targeted treatment tends to be palliative rather than curative, but it has expanded progression-free, tumor-specific, and overall survival for an increasing number of cancer types and added further “bullets” in the “war against cancer.” This marketing-type terminology commonly used in public media tends to generate optimism and hope but in doing so tends to ignore major problems encountered, in terms of drug efficacy in clinical trials but also in the development of reliable companion diagnostics. The example of the EXPAND study, which tested capecitabine and cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer, is a good example: adding cetuximab to capecitabine-cisplatin as first-line treatment did not result in any benefit. Cetuximab is a fully humanized antibody directed against the epidermal growth factor receptor

(EGFR) and provides significant survival benefit to patients with advanced colorectal cancer with wild-type KRAS status [1]. As in comparison with colorectal cancer in gastric cancer KRAS mutations are rare while they do express EGFR, expectations were high but cetuximab failed to show any benefit and the relative success in colorectal cancer did not materialize in better clinical management of gastric cancer [2]. Likewise, results of recent studies communicated during the annual meeting of the American Society of Clinical Oncologists failed to show a benefit of MET inhibitors in gastric cancer. In both studies, drug efficacy was tested along with a potential companion diagnostic [3, 4].

Therefore, in spite of a breathtaking number of clinical trials, many of which also attempt to validate companion diagnostics, the number of effective targeted drugs and the spectrum of tumors for which they work remain limited and this also holds true for the accompanying handful of tissue-based biomarkers that have become standard of care in diagnostic surgical pathology. This raises the question which role pathologists might play in planning and executing clinical trials.

Role of the surgical pathologist in clinical trials

Surgical pathology already contributes considerably to clinical trials. Three roles can be distinguished regarding pathology input: support in clinical studies, participation in preclinical investigations, and implementing trial results, notably concerning companion diagnostics, into diagnostic pathology practice [5]. Support in clinical studies consists of providing sound pathological diagnoses through, e.g., central review, as well as providing biomarker test results (specific tissue-based diagnostics conditional for patient inclusion or exclusion). Support in the preclinical phase includes biomarker discovery, i.e., support for or execution of tissue-based translational studies for the development of new clinically relevant biomarkers, which requires well-characterized tissue collections. Implementing trial results concerns roll-out of new tests into daily diagnostic practice, including bedside-to-bench research projects that improve application of a new tissue-based biomarker, but also the establishment and execution of external, objective quality control procedures.

Diagnostic pathology support in clinical studies

In the past 20 years, major progress has been made in surgical pathology. The World Health Organization (WHO)/International Agency for Research on Cancer (IARC) currently publishes the fourth edition of the classification of tumors. Treatment decisions are based on a sound histological diagnosis and the “blue books,” as they tend to be called, provide a globally accepted consensus-based framework for the

histological and molecular classification of human tumors. Clinical trials would be inconceivable without standardized histological diagnoses, which classify diverse human tumors according to morphology, taking molecular characteristics and natural behavior into account. These histological classifications are work in progress. Novel insight, including concepts based upon next-generation sequencing data but also clinical findings and pharmacotherapeutic progress, all impacts on classification schemes. Lung cancer is a striking example of the dynamic evolution of tumor classification: new modalities of targeted therapy of pulmonary adenocarcinomas have led to international initiatives to profoundly revise their classification [6].

The tumor (T), node (N), metastasis (M) classification of the Union for International Cancer Control (UICC) provides a solid basis for staging of human tumors [7]. It is the most important instrument for tailoring treatment to the needs of the individual cancer patient. The TNM classification has stood the test of time and has never been surpassed in multivariate analyses by any other single prognostic biomarker, often based upon immunohistochemistry or molecular analysis. The TNM classification is used in clinical trials, to select patients who are eligible for inclusion, and in cancer registries to compare outcome between different patient series, across different countries, over different time periods (i.e., ethnicity, medical treatment developments, sociocultural effects), and particularly between different studies. The TNM classification is also continuously improved by the addition of novel criteria and adoption of novel insights into tumor biology and prognostic factors. It is important to emphasize here that every diagnostic pathologist involved in signing out biopsies and surgical specimens of cancer patients contributes to clinical trials, in providing reliable pTNM classification for each case.

Diagnostic standards harmonize and standardize surgical pathology and provide a basis for evidence-based medicine and indirectly for the quality of clinical trials. Standardized classification is particularly important because most clinical trials are multicenter and often span several countries and even continents. Without the support of a solid histopathological diagnosis, including tumor type, grade (if appropriate), and stage, which are all provided by a surgical pathologist who usually is not directly involved in study design and/or execution standardization, clinical trials would be impossible to conduct. Patient accrual by clinicians often depends on work done by surgical pathologists. Beyond standardization of diagnostic criteria, adherence to diagnostic standards needs to be monitored, not only for daily practice but also for clinical trials. To this end, internal and external quality assurance programs have been implemented in many countries. In addition, continuous medical education programs, such as those provided by learned societies including the International Academy of Pathology (www.iapcentral.org), are dedicated to improving quality of diagnostic pathology through education. This

implies that pathologists involved in daily diagnostic practice, which can directly influence quality of trial data, as well as those actively participating in developing and executing clinical studies, should be board-certified specialists with a diagnostic capacity at the level required for standard of care. This may require specific expertise, depending on the particularities of the study or the applied technologies. This becomes imperative when a surgical pathologist is actively involved in a clinical trial as central reviewer providing a second opinion prior to inclusion of a patient.

Participation in preclinical investigations

The question arises whether harmonization of diagnostic standards, continuous medical education, and specialization are sufficient guarantees for the required quality. Tissue-based biomarker testing has become an integral part of histopathology practice in oncology because prognostic biomarkers might determine whether or not the patient needs more, less, or no additional (adjuvant) treatment. Predictive biomarkers are needed to select the right drugs for a specific patient/target, hence the term targeted therapy, the mainstay of precision medicine [8]. As a result, clinical trials increasingly include exploration of tissue-based biomarkers in the quest for prognostic and companion diagnostics. This requires not only diagnostic expertise but also competences in quality assurance laboratory procedures and understanding of diagnostic algorithms beyond basic requirements for board certification. Important issues here are pre-analytical variables, sampling issues, and test and evaluation algorithms.

Pre-analytical variables

An area of particular importance is the impact of pre-analytical variables on the results of a test. A striking example is the Her2/neu retesting controversy in Canada [9]. In 1998, trastuzumab was approved for the treatment of breast cancer but only those that tested positive for Her2/neu expression or amplification on a breast cancer tissue sample. Guidelines for Her2/neu testing of breast cancer were published in 2007 [10]. In spite of these, retesting of breast cancer samples in a quality assurance program revealed major discrepancies, in part due to pre-analytical conditions of tissue treatment including fixation time, as test results are heavily influenced by insufficient fixation. Accumulating evidence led to revised guidelines, which were published in 2013 [11]. In comparison to the significant standardization efforts regarding diagnostic criteria, little effort has been dedicated to standardization of pre-analytical variables. International multicenter trials often fail to provide detailed specification of pre-analytical variables such as fixative and fixation time. While standards of

reporting for tumor marker prognostic studies (REMARK; [12]) and diagnostic accuracy (STARD; [13]) have been published, these mainly refer to standards of reporting and publishing. However, harmonization is also needed in terms of pre-analytical tissue treatment conditions.

Fully automated immunostaining devices and certified test kits have improved tissue-based test results, but these will never become truly robust if pre-analytical variables are not harmonized. Commonly, formalin-fixed and paraffin-embedded tissue specimens are used. Formalin fixation is still strongly recommended, in spite of the fact that formalin fixation leads to fragmented suboptimal DNA quality which allows amplification only of small PCR amplicons. It remains the most widely used fixation procedure, allowing also retrospective cohort studies, particularly of cancers with a low prevalence or historical, chemotherapy-naïve patient populations.

Sampling procedure

It is essential that before any molecular testing, a solid histopathological diagnosis including detailed classification has been established. This also applies to the biopsy taken for clinical trials. DNA will be extracted from tissue samples which include a mixture of non-neoplastic and neoplastic cells. The percentage of each of these varies and has to be considered. Supervision and interpretation of molecular testing should not be carried out without surgical pathologists [14]. Lack of supervision by surgical pathologists may carry the risk of testing non-representative or even non-neoplastic tissue samples. The sensitivity of molecular biological assays varies and mutations can be missed when inappropriate assays are applied or the “mutational load” in a tumor subclone is below detection level [15]. The need for quality assurance programs is exemplified by the external quality assessment for *KRAS* mutation testing in colorectal cancer carried out by European Society of Pathology, in which 27 % of the participants genotyped at least 1 of 10 samples incorrectly [16].

Tumor heterogeneity is becoming one of the main obstacles of cancer treatment in the era of precision medicine. Sequencing of multiple samples of primary clear cell renal cell carcinoma (ccRCC) and its metastatic sites revealed mind-boggling complexity, which allowed reconstruction of its evolutionary history in terms of gene abnormalities [17, 18]. Genetically distinct subclones were found in the primary tumor and distant metastases. Cancer is now viewed as a highly dynamic evolutionary disease [19], with continuing mutations favoring the emergence of new (sub)clones with distinct biological properties already in precursor lesions [20]. Subclones may show different patterns of interaction; they may compete, overtake other clones, parasitize, or peacefully co-exist. Based

upon the ccRCC observations, a “trunk and branch” model has been proposed for tumor evolution [21]. A similar level of heterogeneity has, as yet, not been found in other tumor types, but tumor heterogeneity does have a major impact on the interpretation of earlier studies and the design of future clinical trials. Common events, found in every subclone of the tumor region, are represented in the trunk. Heterogeneous somatic events occurring in a limited number of (or even a single) subclones represent the branches of the tree [21]. While targeting alterations in the trunk might be most promising as they occur in all tumor cells, many actionable targets exclusively occur in branches. This subclonal heterogeneity of treatment targets poses problems in diagnosis (tissue sampling error) and therapy (emergence of resistance). Treatment and eradication of a subclone may provide a growth advantage for a competing non-responsive subclone, which finally may kill the patient. This is probably why the efficacy of several targeted therapies was limited in time [21].

Molecular heterogeneity appears to be inherent to tumor evolution and has consequences for sampling procedures. We have studied the impact of sampling protocols on Her2/neu testing in gastric cancer. Her2/neu has been introduced as a predictive biomarker for the treatment of gastric cancer with trastuzumab [22]. Amplification of genes encoding receptor-tyrosine kinases usually occurs in genomically unstable gastric cancers [23], which is why in gastric cancer, Her2/neu expression is intrinsically heterogeneous. We assessed Her2/neu status (according to the gastric cancer scoring system [24]) using a tissue micro array approach, in which a tissue core serves as surrogate for a biopsy procedure, and compared the results with those obtained on whole tissue sections cut from the same paraffin block. On the TMA cores, we obtained a “false-negative” rate of 24 % and a “false-positive” rate of 3 %. Similar observations regarding heterogeneity were made in gastric cancer for expression of MET [25] and microsatellite instability [26]. Heterogeneous expression of predictive biomarkers poses a major challenge in clinical trials but this issue is often neglected. The phenomenon is often better explored in retrospective biomarker studies, once a treatment modality has been formally approved by a regulatory body. As an example, many studies on Her2/neu expression were carried out after the approval of trastuzumab for the treatment of gastric cancer (for a review, see [27]). These issues should be addressed prior to implementation of a trial, and whether (expression of) the molecular target of the treatment is homogeneous (“trunk alteration”) or heterogeneous (“branch alteration”) should be explored in advance and translated into suitable biopsy procedures. However, many studies address treatment in later stages of cancer with palliative intent, and analyses rely on biopsy samples only, as patients may not be eligible for tumor resection. However, an essential requirement for any trial is the exploration of tumor heterogeneity before the trial takes off, to avoid sampling errors and favor representative tissue biopsies, which should then be assessed according to REMARK standards (Table 1) [13].

Test and evaluation algorithms

Another obstacle is the choice of adequate test and evaluation algorithms. Test results of immunohistochemistry-based markers are sensitive to choice of antibody, staining protocol, and the microscopical evaluation procedure. The use of different antibodies and evaluation procedures, added onto dissimilar study populations, almost inevitably generates enormous variability in the final result and what should be considered the true prevalence of expression and significance of a biomarker may be difficult to establish. With regard to Her2/neu, MET, and microsatellite instability in gastric cancer, we carried out a literature review and found that the prevalence ranges from 5 to 29 % for Her2/neu [28], 3.8–85 % for MET [25], and 0–44.5 % for microsatellite instability [26]. Similar observations were made for Her2/neu in breast cancer. In Australia, the Her2/neu-positivity rate decreased over a 4-year period from 23.8 % in 2006 to 14.6 % in 2010 [29]. Thus, not only standardization of laboratory procedures but also evaluation of the staining result is mandatory. In Germany, this has led to the introduction of the Her2-monitor, which serves as an external benchmark [30]. However, such benchmarks are not available for a considerable number of (future) actionable targets.

KRAS (for treatment with cetuximab) and Her2/neu testing (for treatment with trastuzumab) taught us further lessons. Test algorithms cannot be translated from one tumor type to another without solid experimental evidence. Testing for mutations of *KRAS* (and more recently also of *NRAS*, *BRAF*, and *PIK3CA*) in colon cancer allowed identification of patients (with a *KRAS* mutation) who do not respond to cetuximab [31]. This has improved progression-free and tumor-specific survival in the palliative setting [32], but this has not appeared to be applicable in stomach cancer. Mutations in cancer genes do not occur in isolation but in an established genomic landscape, which is different for different cell and tissue types. This “ground state” of the transformed cell may have a profound impact on the effect of a mutation, such as determining whether cell death or clonal expansion might ensue [19]. Somatic alterations occurring in the “omic” landscape of the stomach might not have the effect of the same mutation in the context of the colon. A somatic mutation may be a driver in one but a passenger in a different context [19]. The understanding pathologists have of cell and tissue context makes their contribution to molecular testing particularly valuable.

Once again Her2/neu is a good example. The breast cancer scoring system was found to be unsuitable for gastric cancer, and it was deemed necessary to change the test algorithm [33]. As yet, generally accepted approaches or guidelines as to how for a novel tissue-based biomarker the test algorithm should be developed do not exist. The approaches are characterized by “trial and error.” The rationale underpinning a cutoff value in an immunohistochemical test is rarely provided, often follows statistical reasoning (e.g., splitting at the median) but rarely reflects tumor biology.

Table 1 Recommendations for tumor marker prognostic studies (REMARK)**Introduction**

1. State the marker examined, the study objectives, and any prespecified hypotheses.

Materials and Methods**Patients**

2. Describe the characteristics (e.g., disease stage or comorbidities) of the study patients, including their source and inclusion and exclusion criteria.
3. Describe treatments received and how chosen (e.g., randomized or rule-based).

Specimen characteristics

4. Describe type of biological material used (including control samples) and methods of preservation and storage.

Assay methods

5. Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.

Study design

6. State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g., by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.
7. Precisely define all clinical endpoints examined.
8. List all candidate variables initially examined or considered for inclusion in models.
9. Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.

Statistical analysis methods

10. Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.
11. Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.

Results**Data**

12. Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.
13. Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.

Analysis and presentation

14. Show the relation of the marker to standard prognostic variables.
15. Present univariate analyses showing the relation between the marker and outcome, with the estimated effect (e.g., hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan–Meier plot is recommended.
16. For key multivariable analyses, report estimated effects (e.g., hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.
17. Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.
18. If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.

Discussion

19. Interpret the results in the context of the prespecified hypotheses and other relevant studies; include a discussion of limitations of the study.
20. Discuss implications for future research and clinical value.

The goal of REMARK is to encourage transparent and complete reporting so that the relevant information will be available to others to help them to judge the usefulness of the data and understand the context in which the conclusions apply [12]. The guideline should also be applied to tissue-based studies of putative prognostic and predictive biomarkers

Table 2 Statement for reporting studies of diagnostic accuracy (STARD)

No.	Item
1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading ‘sensitivity and specificity’).
2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.
3	Describe the study population: the inclusion and exclusion criteria, setting and locations where data were collected.
4	Describe participant recruitment: was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?
5	Describe participant sampling: was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.
6	Describe data collection: was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?
7	Describe the reference standard and its rationale.
8	Describe technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.
9	Describe definition of and rationale for the units, cutoffs and/or categories of the results of the index tests and the reference standard.
10	Describe the number, training and expertise of the persons executing and reading the index tests and the reference standard.
11	Describe whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.
12	Describe methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95 % confidence intervals).
13	Describe methods for calculating test reproducibility, if done.
14	Report when study was done, including beginning and ending dates of recruitment.
15	Report clinical and demographic characteristics of the study population (e.g., age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).
16	Report the number of participants satisfying the criteria for inclusion that did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).
17	Report time interval from the index tests to the reference standard, and any treatment administered between them.
18	Report distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.
19	Report a cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results report the distribution of the test results by the results of the reference standard.
20	Report any adverse events from performing the index tests or the reference standard.
21	Report estimates of diagnostic accuracy and measures of statistical uncertainty (e.g., 95 % confidence intervals).
22	Report how indeterminate results, missing responses and outliers of the index tests were handled.
23	Report estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.
24	Report estimates of test reproducibility, if done.
25	Discuss the clinical applicability of the study findings.

Study design and patient selection threaten internal and external validity of a study of diagnostic accuracy. The STARD aimed to be a useful resource to improve reporting of diagnostic accuracy studies. Complete and informative reporting can only lead to better decisions in health care, which also applies now increasingly to clinical trials exploring tissue-based prognostic and predictive biomarkers [13]

Implementing trial results: roll-out and quality assurance

In many countries, diagnostic standards are maintained through certification and accreditation. In the UK and the USA, clinical pathology is accredited according to the internationally recognized standard ISO/IEC 15189:2012 (www.ukas.com). In Germany, departments of pathology are accredited according to the ISO/IEC 17020:2012 (www.dakks.de). Accreditation programs confirm by an independent third party the compliance to standards and competence in laboratory and diagnostic procedures, which includes continuous medical

education. The National Accreditation Body in Germany has published guidelines for the validation of immunohistochemical tests [34]. The College of American Pathologists has published principles for analytic validation of immunohistochemical assays [35]. As a result, standards in diagnostic pathology are high, but somehow these standards are not always applied in clinical trials or validation studies, at least in part because what has become mandatory in diagnostic surgical pathology might not be applied in a research laboratory. In our literature review of Her2/neu [23], MET [25], and MSI testing [26], we did not find a single publication reporting participation in an external quality

Table 3 Statement of the German Society for Pathology

1. General aspects

Study initiator and director are known; contact details are provided Society, individual researcher/research network (outside a biomedical society), research institute (e.g., university), industry, other

Details of study financing are provided
Society/association (e.g., ESMO, national organizations of anatomical and surgical pathologists), public third-party funds, private third-party funds, industry, other

Type of study is declared
Therapeutic trial, biomarker study, combined therapeutic trial and biomarker study, diagnostic study
Study type 1: histopathological validation for clinical studies (with and without therapeutic relevance)
Study type 2: request of samples for scientific studies for biomarker validation
Study type 3: combination of study types 1 and 2

Aim of the study is defined

Type and amount of requested material is stated
Fresh tissue, paraffin block samples, slides with section/cytological preparations, paraffin sections (unstained), extracts (DNA, RNA, protein extracts), assessment reports (diagnoses), individual information (results), other

2. Legal aspects

Informed patient consent has been given^b
Informed patient consent for forwarding of material, the potential future location of study material and the planned analyses; the patient has been made aware of the possible consequences of forwarding and use of the material

Study-associated trusteeship of material is resolved^b
Within the context of the study/project, responsibilities have been defined and the responsible persons named

The future location or return of study material is defined^b
The location of study material after completion of the investigations has been defined according to conditions; in the case of nonreturned material, the terms of availability and return for future studies/health care purposes are defined

3. Data protection requirements are fulfilled^b4. Ethics committee approval has been obtained^b

5. Pathology-specific standards

The study pathologist is a consultant pathologist^b
Additional study-specific specialist knowledge is particularly desirable

In the case of a histopathological finding relevant to clinical diagnosis, a pathology report will be generated

The head pathologist will be informed of the assessment results before further steps are taken
Ensures optimal continued care of the patient and avoids conflicts between primary pathologist, study pathologist, and clinics

Table 3 (continued)

Archiving of results and material according to national accreditation standards

See e.g., [34]

6. Compensation of expenses

Appropriate compensation for resource and personnel costs
Invoicing and reimbursement modalities are resolved

Surgical pathologists are increasingly asked to provide study material obtained during routine clinical service. This opposes surgical pathologists with a multitude of unresolved issues. The German Society for Pathology generated a structured criteria catalog, which may serve as a source of information and a decision-making aid for studies and other scientific investigations by both, academic centers and other pathology institutes, as well as for their partners [5]

AIO Arbeitsgemeinschaft Internistische Onkologie (Working Group for Internal Oncology), Deutsche Gesellschaft für Pathologie e.V. (German Society for Pathology)

^a Noncompliance

^b An exclusion criterion

assurance program. Certification or accreditation of a surgical pathology research laboratory appears not to exist.

This opens up the question whether certification and accreditation of surgical pathology research laboratories is necessary or even at all feasible. In a translational study aimed at finding a prognostic or predictive tissue-based biomarker, external quality assurance is not appropriate; good scientific and laboratory practice including appropriate positive and negative controls are essential. However, subsequent confirmatory studies anticipating clinical/diagnostic use (which is the case for the vast majority of Her2/neu studies in gastric cancer after formal approval of trastuzumab) should significantly benefit from external quality assurance programs.

Several examples illustrate the gap between phase III clinical trials leading to formal approval of a new drug in combination with a companion diagnostic and subsequent roll-out of the new test into clinical and surgical pathological practice. Narrowing this gap is necessary. Drug development, clinical trials, and roll-out are expensive and time consuming but, when not carried out according to the guidelines recommended by REMARK and STARD (Tables 1 and 2, respectively) [12, 13], may end up to be disastrous for the patient. Quality assurance programs should become mandatory also in validation of biomarker studies and roll-out of new biomarker tests. A minimum requirement during roll-out of clinical trial results with companion diagnostics would be implementation of an external quality assurance program, as soon as the drug that requires a companion diagnostic is approved by the European Medicines Agency. This has become the standard approach of the German Quality Initiative for Pathology (QuIP; a collaborative initiative of the German Society for Pathology (Table 3) and the Bundesverband Deutscher Pathologen e.V.).

Common pitfalls and key issues

Recently, the US Food and Drug Administration (FDA) published common pitfalls and challenges in their companion diagnostic review and approval process [36]. These include the use of multiple tests to test patient specimens to select for a trial, the lack of analytical validation prior to use in the trial, inappropriate specimen types used to validate the assay, missing samples, the instability of analytes during storage, the assessment of drug efficacy in only one subset defined by the test, and the retrospective assignment of cutoffs in a pivotal trial. Almost all also apply to tissue-based diagnostics. The proposals to improve on these key issues provided by the FDA are particularly relevant:

- 1) A companion diagnostic development plan should be included as part of the development plan of the drug under study when biomarker-based conclusions about drug safety and efficacy are anticipated.
- 2) The final version of the test should be used to screen patients for the trial.
- 3) Drug and device claims rely on prespecified device design, and analytical validation prior to initiation of a study is critical to planning patient enrollment.
- 4) A plan for appropriate banking and annotating of patient specimens (both test negative and test positive) and assuring storage that does not impact on test results will be critical to future bridging studies [36].

A problem with the latter is that it will often interfere with regulations impeding on patient service outside clinical trials. Many patients are enrolled in clinical trials only after a diagnostic biopsy or therapeutic resection specimen was obtained, and such tissue samples were obtained for diagnostic and not for research purposes.

Do clinical trials need to interfere with diagnostic pathology service?

Surgical pathologists not involved in clinical trials are increasingly challenged in daily practice by the requirements of an ever increasing number of clinical trials. They are requested to provide tissue samples entrusted to them for diagnostic purposes only or additional study-related information, or even perform additional studies on the tissue samples without having any primary research/study intention. This confronts pathologists in their role as tissue trustee with a multitude of unresolved issues, including heterogeneity of the research landscape (who finances the study, what are the aims of the study, does patient informed consent cover the request), study-associated tissue collections, proprietorship, study results, data protection, and compensation of expenses (for a review see [5]). The German Society of Pathology recently summarized the problems increasingly

encountered in diagnostic pathology outside clinical trials and published a proposal [5], and explicitly recommend that pathologists should be actively involved in study planning, implementation and data analysis, and the use of cell and tissue material. By doing so, pathologists assume their dual role as competent diagnostician and expert advisor in specific issues regarding tissue-based research and diagnostics. In such a position, pathologists would be instrumental in avoiding protocols, procedures, and regulations that are disadvantageous for health care, hamper the study, and disturb interactions between involved parties. The study pathologist would ensure that pathology-specific standards are met comprehensively in cell- and tissue-based research and clinical studies including diagnostics and that the study requirements are not in conflict with principles of good practice. When central pathology review is part of the study, the study pathologist must have an outstanding level of expertise and recognized professional status. Trials and studies which include tissue-based analyses but not tissue based analytical methods (e.g., histopathology, immunohistochemistry, *in situ* hybridization) also require expert histological characterization of tissue samples in order to assure adequate sample selection and based upon relevant parameters (e.g., tumor tissue vs. non-neoplastic tissue, necrotic vs. vital tissue, percentage of tumor tissue in the sample). This is the only way potentially irrelevant results or inadequate interpretation of the results can be avoided [5].

Compliance with ethical standards

Conflict of interest The author declares that he has no competing interests.

References

1. De Rock W, Piessevaux H, De SJ, Janssens M, De HG, Personeni N, Biesmans B, Van Laethem JL, Peeters M, Humblet Y, Van CE, Tejpar S (2008) KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. *Ann Oncol* 19:508–515
2. Lordick F, Kang YK, Chung HC, Salman P, Oh SC, Bodoky G, Kurteva G, Volovat C, Moiseyenko VM, Gorbunova V, Park JO, Sawaki A, Celik I, Gotte H, Melezinkova H, Moehler M (2013) Capecitabine and cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer (EXPAND): a randomised, open-label phase 3 trial. *Lancet Oncol* 14:490–499
3. Cunningham D, Tebbutt NC, Davidenko I, Murad AM, Al-Batran S, Ilson DH, Tjulandin S, Gotovkin E, Zhang Y, Hoang T, Sidhu R, Catenacci DVT (2015) Phase 3, randomized, double-blind, multicenter, placebo-controlled trial of rilotumumab plus epirubicin, cisplatin, and capecitabine as first-line therapy in patients with advanced MET-positive gastric or gastroesophageal junction cancer: RILOMET-1 Study. *J Clin Oncol* 33 (suppl.): Abstr. No. 4000
4. Shah MA, Bang YJ, Lordick F, Tabernero J, Chen M, Hack SP, Phan S, Shames DS, Cunningham D (2015) METGastric: a phase III study of onartuzumab plus mFOLFOX6 in patients with metastatic HER2-negative (HER2-) and MET-positive (MET+)

- adenocarcinoma of the stomach or gastroesophageal junction (GEC) *J Clin Oncol* 33 (suppl.): Abstr. No. 4012
5. Röcken C, Höfler H, Hummel M, Meyermann R, Zietz C, Schirmacher P (2014) Participation in and support of clinical studies and other scientific investigations—Statement of the German Society for Pathology. *Pathol Res Pract* 210:705–712
 6. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Powell CA, Beer D, Riely G, Garg K, Austin JH, Rusch VW, Hirsch FR, Jett J, Yang PC, Gould M (2011) International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. *Proc Am Thorac Soc* 8:381–385
 7. Sobin LH, Gospodarowicz M, Wittekind C (2010) International union against cancer (UICC) TNM classification of malignant tumours, 7th edn. Wiley-Liss, New York
 8. Collins FS, Varmus H (2015) A new initiative on precision medicine. *New Engl J Med* 372:793–795
 9. Gregory DM, Parfrey PS (2010) The breast cancer hormone receptor retesting controversy in Newfoundland and Labrador, Canada: lessons for the health system. *Healthc Manag Forum* 23:114–118
 10. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van d V, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 25:118–145
 11. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF, American Society of Clinical O, College of American P (2013) Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 31:3997–4013
 12. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM (2005) Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol* 23:9067–9072
 13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HC (2003) Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Am J Clin Pathol* 119:18–22
 14. Weichert W, Schewe C, Lehmann A, Sers C, Denkert C, Budczies J, Stenzinger A, Joos H, Landt O, Heiser V, Röcken C, Dietel M (2010) KRAS genotyping of paraffin-embedded colorectal cancer tissue in routine diagnostics: comparison of methods and impact of histology. *J Mol Diagn* 12:35–42
 15. Dubbink HJ, Deans ZC, Tops BB, van Kemenade FJ, Koljenovic S, van Krieken HJ, Blokx WA, Dinjens WN, Groenen PJ (2014) Next generation diagnostic molecular pathology: critical appraisal of quality assurance in Europe. *Mol Oncol* 8:830–839
 16. Tembuysler L, Ligtenberg MJ, Normanno N, Delen S, van Krieken JH, Dequeker EM (2014) Higher quality of molecular testing, an unfulfilled priority: results from external quality assessment for KRAS mutation testing in colorectal cancer. *J Mol Diagn* 16:371–377
 17. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR, Nohadani M, Eklund AC, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal PA, Swanton C (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New Engl J Med* 366:883–892
 18. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, Fisher R, McGranahan N, Matthews N, Santos CR, Martinez P, Phillimore B, Begum S, Rabinowitz A, Spencer-Dene B, Gulati S, Bates PA, Stamp G, Pickering L, Gore M, Nicol DL, Hazell S, Futreal PA, Stewart A, Swanton C (2014) Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* 46:225–233
 19. Yates LR, Campbell PJ (2012) Evolution of the cancer genome. *Nat Rev Genet* 13:795–806
 20. Humphries A, Cereser B, Gay LJ, Miller DS, Das B, Gutteridge A, Elia G, Nye E, Jeffery R, Poulson R, Novelli MR, Rodriguez-Justo M, McDonald SA, Wright NA, Graham TA (2013) Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. *Proc Natl Acad Sci U S A* 110:E2490–E2499
 21. Swanton C (2012) Intratumor heterogeneity: evolution through space and time. *Cancer Res* 72:4875–4882
 22. Bang YJ, Van CE, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, Hill J, Lehle M, Ruschoff J, Kang YK (2010) Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 376:687–697
 23. Cancer Genome Atlas Research N (2014) Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513:202–209
 24. Rüschoff J, Dietel M, Baretton G, Arbogast S, Walch A, Monges G, Chenard MP, Penault-Llorca F, Nagelmeier I, Schlake W, Höfler H, Kreipe HH (2010) HER2 diagnostics in gastric cancer—guideline validation and development of standardized immunohistochemical testing. *Virchows Arch* 457:299–307
 25. Metzger ML, Behrens HM, Böger C, Haag J, Krüger S, Röcken C (2015) MET in gastric cancer—discarding a 10 % cutoff rule. *Histopathology*. doi:10.1111/his.12745
 26. Mathiak M, Warneke VS, Behrens HM, Haag J, Böger C, Krüger S, Röcken C (2015) Clinico-pathological characteristics of microsatellite instable gastric carcinomas revisited. *Appl Immunohistochem Mol Morphol*. doi:10.1097/PAL.0000000000000264
 27. Lordick F, Röcken C (2013) The identification of predictive factors for perioperative chemotherapy in esophago-gastric cancer. *Ann Oncol* 24:1135–1138
 28. Warneke VS, Behrens HM, Böger C, Becker T, Lordick F, Ebert MP, Röcken C (2013) Her2/neu testing in gastric cancer: evaluating the risk of sampling errors. *Ann Oncol* 24:725–733
 29. Bilous M, Morey AL, Armes JE, Bell R, Button PH, Cummings MC, Fox SB, Francis GD, Waite B, McCue G, Raymond WA, Robbins PD, Farshid G (2012) Assessing HER2 amplification in breast cancer: findings from the Australian In Situ Hybridization Program. *Breast Cancer Res Treat* 134:617–624
 30. Choritz H, Busche G, Kreipe H (2011) Quality assessment of HER2 testing by monitoring of positivity rates. *Virchows Arch* 459:283–289
 31. De RW, Claes B, Bernasconi D, De SJ, Biesmans B, Fountzilias G, Kalogeris KT, Kotoula V, Papamichael D, Laurent-Puig P, Penault-Llorca F, Rougier P, Vincenzi B, Santini D, Tonini G, Cappuzzo F, Frattini M, Molinari F, Saletti P, De DS, Martini M, Bardelli A, Siena S, Sartore-Bianchi A, Tabernero J, Macarulla T, Di FF, Gangloff AO, Ciardiello F, Pfeiffer P, Qvortrup C, Hansen TP, Van CE, Piessevaux H, Lambrechts D, Delorenzi M, Tejpar S (2010) Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* 11:753–762

32. Amado RG, Wolf M, Peeters M, Van CE, Siena S, Freeman DJ, Juan T, Sikorski R, Suggs S, Radinsky R, Patterson SD, Chang DD (2008) Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* 26:1626–1634
33. Rüschoff J, Dietel M, Baretton G, Arbogast S, Walch A, Monges G, Chenard MP, Penault-Llorca F, Nagelmeier I, Schlake W, Hofler H, Kreipe HH (2010) HER2 diagnostics in gastric cancer-guideline validation and development of standardized immunohistochemical testing. *Virchows Arch* 457:299–307
34. Germany's National Accreditation Body (2013) Guideline of the Pathology/Neuropathology sector committee for the validation of examination methods in immunohistology. <http://dakks.de/en/content/guideline-pathologyneuropathology-sector-committee-validation-examination-methods>
35. Fitzgibbons PL, Bradley LA, Fatheree LA, Alsabeh R, Fulton RS, Goldsmith JD, Haas TS, Karabakhtsian RG, Loykasek PA, Marolt MJ, Shen SS, Smith AT, Swanson PE, College of American Pathologists P, Laboratory Quality C (2014) Principles of analytic validation of immunohistochemical assays: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med* 138:1432–1443
36. Roscoe DM, Hu YF, Philip R (2015) Companion diagnostics: a regulatory perspective from the last 5 years of molecular companion diagnostic approvals. *Expert Rev Mol Diagn* 15:869–880