ORCA – Online Research @ Cardiff

Clinical Trial Endpoints in Metastatic Cancer: Using Individual Participant Data to Inform Future Trials Methodology

Richard M. Goldberg, MD[1], Richard Adams, MD[2], Marc Buyse, ScD[3,4], Cathy Eng, MD[5], Axel Grothey, MD[6], Thierry André, MD[7], Alberto F Sobrero, MD[8], Stuart M. Lichtman, MD[9], Al B Benson, MD[10], Cornelis J.A. Punt, MD[11], Tim Maughan, MD[12], Tomasz Burzykowski, PhD[3,4], Dirkje Sommeijer, MD[13], Everardo D. Saad, MD[3,14], Qian Shi, PhD[15], Elisabeth Coart, PhD[3], Benoist Chibaudel, MD[16], Miriam Koopman, MD[11], Hans-Joachim Schmoll, MD[17], Takayuki Yoshino, MD[18], Julien Taieb, MD[19], Niall C Tebbutt, MD[20], John Zalcberg, MD[21], Josep Tabernero, MD[22], Eric Van Cutsem, MD[23], Alastair Matheson, PhD[24], Aimery de Gramont, MD[16,24]

1. West Virginia University Cancer Institute, Morgantown, WV, USA.

2. Cardiff University and Velindre UNHS Trust, UK.

3. International Drug Development Institute (IDDI), Louvain-la-Neuve, Belgium

4. Hasselt University, Hasselt, Belgium.

5. Vanderbilt-Ingram Cancer Center, Nashville, TN, USA.

6. West Cancer Center and Research Institute, Germantown, TN, USA.

7. Hôpital Saint Antoine, Paris, France.

8. Ospedale S. Martin, Genova, Italy.

9. Memorial Sloan Kettering Cancer Center, New York.

10. Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago IL, USA.

11. University Medical Centre Utrecht, Netherlands.

12. Gray Institute of Radiation Oncology and Biology, University of Oxford, UK.

13. University of Amsterdam Academic Medical Centre and Flevohospital, Almere, The Netherlands.

14. Dendrix Research, Sao Paulo, Brazil.

15. Mayo Clinic, Rochester, MN, USA.

16. Hôpital Franco-Britannique, Paris, France.

17. Martin Luther University, Halle, Germany.

18. National Cancer Center Hospital East, Kashiwa, Japan.

19. Georges Pompidou European Hospital, Paris, France.

20. Austin Health, Heidelberg, Victoria, Australia.

21. Monash University, School of Public Health, Australia.

22. Vall d'Hebron Hospital Campus and Institute of Oncology (VHIO), Barcelona, Spain.

23. University Hospital Gasthuisberg, Leuven, Belgium.

24. Fondation ARCAD, Paris, France.

Corresponding author: Professor Aimery de Gramont

Hôpital Franco-Britannique

4, rue Kléber

92300 Levallois-Perret, Paris, France

Email: adegramont@fondationarcad.org

Phone: +33147316919

Abstract

Meta-analysis based upon individual participant data (IPD) is a powerful methodology for synthesizing evidence by combining information drawn from multiple trials. Hitherto, its principal application has been in questions of clinical management, but an increasingly important use is in clarifying trials methodology, for instance in the selection of endpoints, as discussed in this review. In oncology, the Aide et Recherche en Cancérologie Digestive (ARCAD) Metastatic Colorectal Cancer Database is a leader in the use of IPD-based meta-analysis in methodological research. The ARCAD database contains IPD from over 38,000 patients enrolled in 46 studies and continues to collect Phase III trial data. Here, we review the principal findings of the ARCAD project in respect of endpoint selection and examine their implications for cancer trials. Analysis of the database has confirmed that progression-free survival (PFS) is no longer a valid surrogate endpoint predictive of overall survival in the first-line treatment of colorectal cancer. Nonetheless, PFS remains an endpoint of choice for most first-line trials in metastatic colorectal cancer and other solid tumors. Only substantial PFS effects are likely to translate into clinically meaningful benefits, and accordingly we advocate an oncology research model designed to identify highly effective treatments in carefully defined patient groups. We also review the use of the ARCAD database in assessing clinical response including novel response metrics, and prognostic markers. These studies demonstrate the value of IPD as a tool for methodological studies and provide a reference point for the expansion of this approach within clinical cancer research.

Meta-analysis based upon deidentified individual participant data (IPD) is recognized as the gold standard approach for the synthesis of multiple studies conducted in a specific disease setting (1). One approach to IPD meta-analysis is to construct a database of trials as a standing resource supporting ongoing research projects, an approach pioneered within oncology by the Early Breast Cancer Trialists' Collaborative Group (2). This group has established the oldest, largest and most successful standing IPD resource in oncology, and uses it to address numerous clinical issues in patients with early breast cancer (2). Other similar initiatives in patients with early disease are the Adjuvant Colon Cancer End Points (ACCENT) collaboration (3) and the Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP) collaboration (4). In the advanced disease setting, a pioneering contribution has been made by the Aide et Recherche en Cancérologie Digestive (ARCAD) Metastatic Colorectal Cancer Database (5). The ARCAD database has been used to clarify methodological challenges that are pertinent not only to trials in unresectable locally advanced or metastatic colorectal cancer (designated hereafter as mCRC) but which have relevance to studies in virtually all advanced solid tumors. Our review provides a detailed illustration of the application of data sharing to issues of methodology, with the goal of informing similar initiatives.

The changing oncology trials paradigm

Clinical research in advanced solid tumors is in a period of transition, and studies in colorectal cancer exemplify these changes. Trials in which large, heterogeneous patient populations are treated with one or two uniform interventions are being replaced by studies tailoring treatments to smaller and more stratified groups defined by specific patient and tumor profiles. Clinical management is also evolving, with growing use of targeted therapy including immune checkpoint inhibitors, surgery and ablative techniques, and treatment

4

sequences with greater cumulative effectiveness. A large collection of randomized clinical trials spanning decades of new treatment development provides a reliable source of data to assess therapeutic progress not only within trials, but also over time. In contrast, real-world evidence is often conflicting, probably due to the many confounding factors that cannot be accounted for when analyzing data from various real-world sources. In mCRC, progression on first-line treatment generally manifests itself in less than one year, with patients who receive optimal management with additional lines of therapy surviving for 2-3 years or longer. Outside clinical trials, patients are often older and in poorer health or have poorer access to treatment, and their outcomes may therefore be worse. For instance, a recent study by the Dutch Cancer Registry found that across the Netherlands, in a real-life setting, median OS in mCRC remained approximately 12 months throughout the period 2008-2016 (6). In the United States, the 2-year relative survival rate for patients diagnosed with metastatic disease increased from 21% during the mid-1990s to 37% during 2009-2015 (7).

One achievement of clinical trials conducted in the last 25 years has been a thorough investigation of clinical-trial endpoints, including response rate (RR), progression-free survival (PFS) and their relationship with overall survival (OS). Historically, OS has been the endpoint of choice in advanced cancer trials and remains so for trials of second and subsequent treatment lines, but in the first-line setting the availability of additional interventions after initial progression of disease is increasing overall survival and this has eroded the utility of OS as an informative endpoint. This is due to a number of factors including the confounding effects of crossover in randomized trials, the expanding range of interventions available after first-line progression, greatly varying use of salvage therapies, and the growing incursion of statistical noise as the gap between first progression and death increases (8, 9).

Investigators who are designing clinical trials in mCRC and other solid tumors therefore face a dilemma. On the one hand, the foremost goal of clinical intervention is usually increased survival, and regulatory authorities therefore continue to prefer OS as a primary endpoint; yet assessing treatments by their impact on OS in an unselected patient population requires larger, more complex, expensive, and lengthy trials, with the possibility that positive treatment effects might be missed or poorly characterized. On the other hand, while PFS is representative of the direct antitumor effects of an investigational treatment, its weakened association with OS raises the problem that medicines might gain market access on the basis of PFS data despite having no true impact on OS. This could hamper clinical decision-making and result in an accumulation of expensive novel therapeutics that lack clinically meaningful new benefits for patients. The ARCAD Advanced Colorectal Cancer Database has enabled in-depth analyses of the surrogacy of PFS for OS in mCRC patients treated with contemporary therapies, as discussed further below.

Trialists and regulators have taken a pragmatic approach to this dilemma, assessing treatment effects according to a mixture of parameters, including both PFS and OS, as well as RR, disease control rate, duration of response and, increasingly, other factors such as symptom control, secondary curative-intent surgery, and quality of life. In addition, phase IV observational cohort studies and academic clinical trials conducted after regulatory approval continue to play a critical role in optimizing the use of new therapeutics. Given the modest efficacy of most emerging therapeutics in unselected patient populations, it typically takes several years after initial approval before the optimal deployment of new treatments in concert with other interventions is established.

Against this background, our scientific understanding of cancer is now developing rapidly. The disease is being subtyped according to its molecular and genetic characteristics (10); new targeted and investigational treatments including immuno-oncologic agents are

Downloaded from https://academic.oup.com/jnci/advance-article/doi/10.1093/jnci/djab218/6449449 by Cardiff University user on 03 February 2022

being introduced; and a diversity of new markers of response and prognosis are becoming available, as discussed below. We are also now entering an era of increased data sharing, and IPD pooled from multiple trials is providing a more powerful approach than meta-analysis based on aggregate data (11).

Applying IPD to the evaluation and selection of trials endpoints

ARCAD is an international collaboration of leading experts and trialists in colorectal cancer, and currently has some 80 academic members in 16 countries. The collaborative engages directly in research through its database and holds consensus discussions drawing upon its members' clinical and research expertise. At the time of writing, IPD from over 38,000 patients in 46 trials published between 2002 and 2015 are included in ARCAD's Metastatic Colorectal Cancer Database (**Supplementary Table 1**). We are currently in negotiations with the leading investigators of several other trials to provide data from their studies, and the database is now an enduring, shared resource for the CRC research community. The collaboration welcomes proposals and participation from trialists globally. Currently, ARCAD is also constructing a database of European trials in pancreatic cancer as a research resource modelled upon the mCRC database.

Assembling the mCRC database has and continues to require considerable effort including contractual negotiations with trial sponsors and subsequent data processing for integration into the database. Due to intellectual property considerations important to many of the contributors, especially pharmaceutical companies, the database is by agreement not used to investigate the comparative efficacy and toxicity of different specific cancer drugs. Its primary focus is on trial methodology and clinical management of mCRC. In particular, the ARCAD database has helped expand our knowledge of trial endpoints, and prognostic and predictive markers in mCRC.

7

Progression-free survival

PFS is a standard endpoint in cancer studies particularly in patients with advanced disease. It permits more rapid endpoint assessment than OS, supports smaller trials, and is unaffected by crossover or post-progression treatment and events. As discussed above, however, it is no longer clear that increases in PFS in clinical trials are predictive of corresponding increases in OS. Furthermore, PFS is vulnerable to error and bias because of its dependence on tumor measurements, the definition of progression, and censoring. Rules concerning assessment of PFS are heterogeneous, differ between trials and remain the subject of continued debate [see for example references (12-14)]. The use of PFS is also less reasonable in trials involving de-intensification of therapy and reintroduction at progression, as opposed to trials in which patients receive sustained treatment until that intervention fails.

For mCRC trials conducted up to 1999, first-line PFS was shown to be a strong surrogate endpoint for OS for chemotherapy treatments alone on the basis of meta-analysis [reference (15); for an account of the difference between prognostic and predictive markers in individuals, and trial-level surrogate endpoints, see (16)]. However, these trials were conducted during a period in which there were limited treatment options, treatment was less effective than it is today, and survival after progression was typically short. During this time there were few second-line options, they had limited efficacy, and there was minimal biological stratification or use of targeted therapies. Investigations of trials conducted after 2000 have generally failed to demonstrate valid surrogacy of later-line PFS for OS, either in mCRC or other metastatic cancers. For instance, PFS and response-based endpoints were not found to be acceptable surrogates for OS in patients with breast cancer (17). PFS was correlated with OS at both the individual and trial level in studies of *Her2*-amplified breast cancer, as well as in lung cancer, but the relationships were modest and PFS failed to qualify

as a reliable surrogate for OS (18, 19). Haslam et al (20) reviewed studies of surrogate endpoints predicated on meta-analysis across a range of oncology settings and reported that trial-level validation studies generally found at best low correlations with OS.

The ARCAD database has been used in several studies to investigate PFS and OS in in the setting of advanced CRC (21, 22). In an analysis of all trials in the database involving first-line treatment, Shi et al (22) examined IPD for 16,762 patients from 22 first-line studies first published from 2003 to 2012. In total, 12 of these studies tested anti-angiogenic and/or anti-epidermal growth factor receptor (EGFR) agents. Overall, median first-line PFS was 8.3 months, and median OS was 18.2 months, both of which are shorter than we have come to expect in contemporary trials. The correlation between treatment effects on PFS and OS was modest, with an R^2 of 0.45 - 0.69 in different analyses, indicating a correspondingly modest surrogacy (**Figure 1**). Of note, there is no consensus on the strength of the trial level association required for claims of surrogacy, but a threshold for R^2 of 0.75, which corresponds with a correlation coefficient above 0.865 and exceeds that obtained in this analysis, is often used (23). The degree of surrogacy was not improved in analyses limited to trials which investigated biologic agents, nor when limited to trials which treated patients continuously to progression as opposed to trials in which treatment was discontinued or reduced in intensity prior to progression, nor when limited to superiority trials. In most of the studies included in this analysis, more than 50% of patients received subsequent treatment, in many cases with targeted agents, which likely contributed to the long post-progression survival, reducing the strength of formal surrogacy. Median post-progression survival was longer than median first-line PFS irrespective of whether patients received biologic agents in first-line treatment.

In the contemporary setting, both PFS and OS have become longer, and as additional new treatment options emerge, there are likely to result in further increases in OS. This trend

is also apparent in other solid tumors such as non-small cell lung cancer (24). Therefore, at least in first-line treatment, the modest surrogacy of PFS for OS documented in the 2015 ARCAD study is likely to become weaker for trials with new effective targeted agents and in broad patient populations. Conversely, surrogacy of PFS for OS would likely be restored in the case of trials that test highly effective treatments, often in patients selected because their tumors harbor defined genomic vulnerabilities enabling substantial extensions in first-line PFS. This might be achieved through careful biomarker- and clinically- defined patient selection for likely responsive disease subtypes or through the emergence of highly effective new therapeutics in defined patient subgroups, such as immune checkpoint inhibitors, as discussed below.

Regardless of whether surrogacy for OS is achieved, however, PFS remains the best available direct measure of the activity of most new treatments. Furthermore, improvement in PFS may constitute a valuable clinical and patient benefit in its own right: in clinical studies conducted in patients with mCRC and other solid tumors, a prolonged PFS has been found to correlate with enduring QOL and reduction of disease symptoms (e.g. 25-27). PFS therefore remains the preferred primary endpoint for most first-line superiority trials in mCRC. Agents that demonstrate a clinically meaningful PFS treatment effect, that have acceptable tolerability and do not appear to negatively impact OS, should be considered acceptable options for ongoing research and introduction of new therapeutics into clinical use.

Response-based endpoints

Response-based endpoints such as objective response rate (RR) were not validated as trial-level surrogate endpoints for OS even when tested in trials conducted prior to 2000 (28). However, response is a good predictor of outcome at the individual level for therapeutics designed to cause tumor shrinkage, and in recent years there has been interest in the utility of

10

both time to response and magnitude of response as endpoints in the mCRC and other settings (e.g. 29-30). The ARCAD database has therefore been used to conduct several investigations of response-based endpoints, both as predictive markers at the individual level, and as potential trial-level surrogates for OS and PFS.

Considering first the time to tumor response, Sommeijer et al assessed early tumor response, defined as any confirmed complete or partial response by RECIST at 6, 8 or 12 weeks according to the trial, for individual-level correlation with OS and PFS in first-line trials (32). IPD from 13,949 patients enrolled in 15 randomized first-line phase III trials were used, of which 8 trials used anti-angiogenic or anti-EGFR targeted agents. Early response was found to be a strong and independent predictor of both OS and PFS ($P<0.0001$). The individual-level predictive value remained statistically significant after adjusting for age, gender, performance status and location of metastatic disease in the liver or lung. Overall tumor response at 26 weeks proved to be superior to early response parameters in predicting OS.

In a similar study, Sommeijer et al compared markers of early response at 6, 8/9, or 12 weeks with the standard response endpoints of best overall response and confirmed response, in patients treated only with chemotherapy, as predictors of PFS and OS (33). All the patients in this analysis were treated with 5FU-LV or capecitabine alone or with the addition of oxaliplatin and/or irinotecan. In addition to early objective tumor response as above, the study assessed two further metrics of response, early tumor shrinkage (defined as a ≥ 20% decrease from baseline) and non-progression status, defined as either a response or as stable disease. At the individual level, all these measures of early response were found to be statistically significantly predictive of both PFS and OS (**Figure 2**). The strength of association at all time points was equivalent to that seen in the standard RECIST response endpoints of best overall response and confirmed response. A further study examined early

11

response endpoints in trials of anti-angiogenic therapies and chemotherapy, and again found strong correlations between speed of response and outcomes at the individual level (34).

The demonstration of an association between response and outcome for individual patients does not necessarily imply, however, that a corresponding association between response and outcome will be observed at the level of clinical trials, where treatment effects are compared between groups of patients (16, 35). It is in fact rare for individual level correlations between response and outcome to translate into trial level-surrogacy. In a study using the ARCAD database, Coart et al demonstrated that despite their prognostic importance at the individual level, the early response endpoints discussed above were not viable surrogates for PFS or OS at the trial level (36). The study also assessed the two conventional endpoints of best overall response and confirmed response. None of the endpoints investigated had consistently strong trial-level correlations with PFS or OS to qualify them as surrogates. There were different levels of association of response parameters with PFS and OS, ranging from 0.01 to 0.92, with the width of the confidence intervals for the R^2 estimates indicating either a weak trial-level association between PFS and OS, or merely a high level of uncertainty (**Table 1**). The findings were consistent for early and conventional response parameters, and for chemotherapy alone and for chemotherapy that was administered with biologic agents.

A further study of 20 first-line trials investigated both early response and depth of response as potential trial-level surrogates for OS (37). The study estimated time to nadir (i.e. maximum response) and depth of nadir, with the nadir estimated using a linear mixed-effects model rather than based upon direct tumor measurements.

Using this approach, separate analyses were conducted for patients treated with chemotherapy alone, those who received additional anti-angiogenic agents, and those who received additional EGFR pathway-targeted agents. In no setting was time to nadir or depth

of nadir found to be an acceptable surrogate for OS. **Figure 3** presents the results for patients

treated with chemotherapy alone: time to nadir demonstrated only a moderate association,

and depth of nadir a weak association with OS.

Of note, the study also identified differences in tumor-growth kinetics between anti-

angiogenic and anti-EGFR agents. The addition of an anti-angiogenic agent to chemotherapy

appeared to be associated with a later, although not typically deeper nadir, whereas the

addition of an anti-EGFR agent often produced a deeper nadir but with a varying time to best

response. These observations appeared to support the clinical impression that the addition of

an anti-EGFR agent for patients with RAS wild type tumors may achieve a greater depth of

response than the addition of an anti-angiogenic agent.

In summary, analyses of response by ARCAD and others suggest that response-based

endpoints are informative and have a continuing role in clinical trials in mCRC, though they

also have important limitations. ARCAD studies were unable to establish surrogacy of either

conventional or novel response endpoints for PFS or OS, but at the individual level, early

response is a robust indicator of prognosis, and depth of response has direct clinical relevance

in respect of symptom control and in some circumstances an increased likelihood of surgical

resection. Response-based endpoints have long been considered useful in early-phase trials

when selecting regimens for further testing in phase 3 trials and will continue to be useful in

this setting.

Endpoints and thresholds in contemporary research

Currently, our understanding of metastatic cancer and its management is developing

rapidly, and this has important implications for trial methodology and endpoints. Most

importantly, cancers are being subclassified according to an expanding range of driving

genetic aberrations, based upon the consensus molecular sub-classification of disease (9) and

burgeoning knowledge of signal transduction pathways and immune-system responses. In mCRC, the chief focus to date has been the EGFR and vascular endothelial growth factor (VEGF) signaling pathways, for which targeted therapies exist that have been in the clinic for over a decade. Numerous other agents are in development that target these pathways. For a contemporary review of biomarker-guided therapeutics targeting these and other pathways, see Xie et al (38). The most remarkable recent advances have been seen with immune checkpoint inhibitors. In mCRC, only about 5% of all patients have DNA mismatch repair-deficient (dMMR)/high microsatellite instability (MSI-H) tumors in which highly durable responses may be achieved using these inhibitors (39). A major current research interest is to optimize their use and generate strategies that lead to comparable responses in a wider range of CRC and other cancers (40).

The cardinal features of contemporary research in mCRC are the development of new targeted and immune therapies, and the coordinated identification of new biomarkers. Within this approach, clinical trialists design studies to identify and validate biomarkers and to achieve substantial efficacy in selected populations that often represent a small fraction of patients with mCRC. To this end, multi-arm, multi-strategy (MAMS) approaches have been pioneered in trials such as FOCUS4, a molecularly stratified trial (41). With this type of trial design the investigators anticipate more tumor responses and delayed progression, with a restoration of surrogacy between PFS and OS in some settings. Realistically, however, many new agents entering clinical research in mCRC are likely to remain modestly effective, and much research will continue to be concerned with optimizing their use through the integration of interventions that individually are relatively minor advances. Clinical research must therefore identify both highly effective new treatments in bespoke populations *and* more modest advances affording more cumulative incremental benefits, and trial endpoints must serve these objectives.

14

Accordingly, we recommend an investigational model in which highly effective treatments are identified according to strong responses and substantial absolute increases in PFS, that are of sufficient magnitude to be readily detectable in trials of modest sample size that can be rapidly completed. There is debate within the oncology community about the scientific and ethical suitability of conventional parallel-group trial designs for the validation of highly effective treatments, but further evaluation is required before a consensus can be reached on how alternatives such as single arm studies might be employed (42). Single-arm studies with response rate as an endpoint are useful in demonstrating the anti-tumor activity of new agents which induce substantial tumor shrinkage, but single-arm trials provide sufficient evidence for the adoption of a treatment only in exceptional circumstances unlikely to be achieved by most emerging oncology therapeutics. Even when the effect size justifies a small trial, a randomized comparison should therefore be required for regulatory approval. Other alternatives to standard parallel-group trials, such as intra-patient designs for trials of second and third-line therapies may have potential in some settings, but their utility remains to be demonstrated for the vast majority of therapeutic advances (43,44). Treatments that are introduced into clinical practice on the basis of small, rapid randomized trials should be subjected to further randomized trials with patient-relevant clinical endpoints. Such treatments should also undergo stringent scrutiny in real-life settings. More modestly effective treatments can be investigated using MAMS and larger parallel-group trials, with biomarker-directed patient selection and stratification whenever possible. Early stopping rules can be incorporated for futility, but these thresholds should be set cautiously to prevent minor but potentially useful benefits being overlooked.

PFS should in general continue to be the primary endpoint of choice for randomized first-line trials, subject to the detailed qualifications set out in this paper. In order to encourage ambitious research goals, and building on the concept of defining a clinically

15

meaningful outcome (45-47), trialists should determine an upper PFS hazard ratio threshold as a research target. This will permit them to distinguish highly effective treatments from more modestly effective ones. In addition, determining a lower PFS hazard ratio (HR) or response odds ratio threshold will enable trialists to distinguish modestly effective treatments from likely futile ones. In the setting of mCRC, we suggest that if HR is used to quantify the treatment effect, then a hazard ratio threshold of approximately 0.5 or lower should be the target for considering a treatment to be classified as paradigm-changing and a candidate for rapid introduction into clinical use. This is an ambitious goal: for instance, the recent KEYNOTE-177 phase 3 trial of pembrolizumab versus chemotherapy in d-MMR/MSI-H mCRC reported an interim HR for PFS of 0.6 (39). Subsequently, an HR of 0.74 was reported for OS, with crossover to immune checkpoint inhibitors in almost 60% of patients in the control arm (48). Conversely, thresholds of 0.8 or higher in biomarker-defined populations are likely to indicate futility. These thresholds are intended to be approximate, and factors such as toxicity, quality of life benefits and cost should also impact on the evaluation of new treatments.

Emerging markers, assessments and endpoints

Alongside studies of standard trial endpoints and their variants, IPD meta-analysis can be used to assess a burgeoning variety of prognostic markers, response measures and candidate endpoints currently being investigated in mCRC and other settings. These include the following:

Prognostic markers for trial stratification and treatment planning. Alongside the molecular reclassification of CRC, an increasing number of clinical prognostic markers are emerging from post hoc and meta-analyses of large trial data sets, and with respect to mCRC the ARCAD database is proving instrumental in helping to identify, validate, and characterize

16

their prognostic importance (49,50). Specific parameters investigated include: patient age (51), body-mass index (52), the presence or absence of peritoneal disease (53), whether the primary tumor has been resected (54), primary tumor location on the right versus those originating in the left colon and rectum (55), the presence or absence of lung metastases (56) and the presence or absence of heterogeneous lesion responses (57,58). This ongoing work will continue to inform treatment planning and trial stratification. In terms of patient demographics, we draw particular attention to elderly patients as a population with a high risk of mCRC who are often under-represented in trials, and for whom further research is needed to optimize treatment. Despite increasing incidence in younger patients mCRC remains predominantly a disease of the elderly (7), and age and comorbidities should be a standard consideration in trials design, both in mCRC and other cancers.

Novel assessments of response. A variety of emerging markers of response are entering into trials and clinical practice, such as change in carcinoembryonic antigen (59), circulating tumor cells, mRNA, and circulating tumor DNA (60,61). The latter is rapidly becoming established in many fields of oncology as a useful longitudinal marker of dynamic subclone development and may provide an earlier and more sensitive means for assessing treatment response than measurement of tumor dimensions. However, comparison and standardization of methods for the evaluation of circulating tumor DNA from plasma samples is needed. Diagnostic imaging is also developing rapidly: for instance, the PERCIST response criteria based on positron emission tomography is finding wider use, and magnetic resonance imaging is emerging as a sensitive technique for evaluating response and guiding management, for instance in the early detection of liver metastases (62). Determining the optimal role of these new assessments in the clinic and within clinical trials will require extensive study, and the ARCAD database will provide a research tool as it assimilates relevant data from forthcoming trials.

17

Novel trial endpoints. The evolving oncology landscape will involve an increasingly diverse set of endpoints to characterize the effects of treatments in diverse patient groups and clinical settings. In terms of efficacy, strategy endpoints such as Duration of Disease Control or Time to Failure of Strategy will likely grow in importance as attempts to characterize specific multi-line treatment strategies continue and have been shown by ARCAD to correlate at the trial level with OS (21). For instance, QOL, patient-reported outcomes and related holistic endpoints have been investigated by ARCAD and others and are likely to grow in importance (63-65). The use of cost-benefit analysis will also expand. Circulating tumor DNA kinetics have been demonstrated to be a useful early predictor of response to immunotherapy in non-small cell lung cancer (66) and will be investigated by ARCAD as relevant data are accrued. Novel endpoints may be required for certain classes of therapy such as immune therapies if their effects on response and progression by RECIST1.1 differ markedly from conventional chemotherapy and targeted therapies, and new radiologic criteria of evaluation such as iRECIST criteria are likely to grow in importance (67). There is likely also to be greater use of composite endpoints, for example response rate coupled with QOL or cost metrics. It will also be necessary to further clarify the meaning of clinical benefit in different patient contexts and design endpoints accordingly. For instance, in trials involving very elderly patients, OS can be a misleading measure of a treatment's effectiveness since life expectancy is shorter than in younger patients and many patients die of non-cancer causes. Accordingly, factors such as the patient's QOL, maintenance of function and independence need to be given appropriate weight (68).

Conclusions

Analyses by ARCAD and others that pool individual patient data from multiple large trials conducted during the past 25 years have led to a detailed understanding of the strengths

and limitations of study endpoints in advanced colorectal cancer, including PFS, tumor response, strategy endpoints such as duration of disease control, and OS. Substantial progress has been made in understanding the clinical meaning of these endpoints, the predictive value of early endpoints for OS at the individual level, and their trial-level surrogacy for OS. PFS remains the endpoint of choice for first-line trials in this setting despite lack of strong trial-level association with OS. The ARCAD database is also actively used in identifying and characterizing novel prognostic markers, and has the potential to become a productive platform for validating and exploiting a burgeoning diversity of novel assessment techniques and candidate endpoints. The range of methodological issues and novel assessments investigated by ARCAD in colorectal cancer is typical of many advanced solid tumors, and fundamentally reflects a research environment in which most emerging treatments continue to have only modest efficacy individually, that when deployed in an optimal sequence and aggregated over multiple lines of therapy can lead to meaningful prolongation of OS. Meanwhile, advances in molecular medicine are raising the possibility of optimizing choices and making more substantial progress.

On the basis of these studies, we favor a research paradigm for the introduction and optimization of new therapeutics in which highly promising new treatments are identified in small randomized trials, on the basis of relatively large effects on PFS supported by response data, while less effective treatments are qualified in specific patient groups using larger trials designed to confirm more modest benefits. In addition to clinical trials focused on efficacy, there is now increasing interest in real-life clinical effectiveness trials and monitoring by means of registries or electronic health records. Real-world data should ideally confirm conclusions from clinical trials and meta-analyses about clinical practice, assessment criteria and endpoints, and feed back into new hypotheses and trial designs (**Figure 4**). ARCAD will

explore the possibility of synergies between its IPD approach and real-life data in the investigation of prognostic markers, assessments of response and trial endpoints.

The ARCAD Advanced Colorectal Cancer Database has pioneered wide data sharing within one type of cancer oncology. By applying this model across other disease types, ARCAD members believe that it can contribute to realizing the full promise of data sharing for patients. Ideally, data sharing within medicine should be prospective, with trials designed to facilitate the incorporation of the deidentified IPD into multi-trial databases (69). It should also ideally be used to assess questions of treatment efficacy, toxicity, quality of life, and cost-effectiveness in addition to those of methodology and biomarkers, although substantial institutional, commercial and intellectual property issues are likely to limit progress towards this ideal. In oncology, the EBCTCG remains preeminent in respect of its achievement of these goals and size, and consequently its power as a research tool. In recent years the pharmaceutical industry has collaborated across companies to introduce several initiatives to facilitate data sharing (11,69), but intellectual property concerns are likely to continue to limit the academic use of proprietary data. Of note, real-world patient data are also of high commercial value, and academic researchers may meet similar challenges to those that have been encountered in clinical trials in accessing and sharing this data for independent study. Notwithstanding the challenges of the data sharing era, the ARCAD Metastatic Colorectal Cancer Database demonstrates that highly informative standing databases can be successfully established if the objectives of all contributors are accommodated within a collegial framework which recognizes that the most important stakeholders are our patients.

Funding

20

Notes

Data Availability

No new data were generated or analyzed during the preparation of this review.

References

1. Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA*. 2015;313(16):1657-1665. doi:10.1001/jama.2015.3656.

2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG)

https://www.ctsu.ox.ac.uk/research/the-early-breast-cancer-trialists-collaborative-group-ebctcg. April 26, 2020

3, Renfro LA, Shi Q, Sargent DJ. Mining the ACCENT database: a review and update. *Chin Clin Oncol*. 2013;2(2):18. doi:10.3978/j.issn.2304-3865.2013.03.05.

4. ICECaP Working Group, Sweeney C, Nakabayashi M, et al. The Development of Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP) [published correction appears in J Natl Cancer Inst. 2016 Sep 14;108(9):null]. *J Natl Cancer Inst*. 2015;107(12):djv261. Published 2015 Sep 25. doi:10.1093/jnci/djv261.

5. Sargent DJ, Buyse M, Matheson A, Goldberg RM, de Gramont A; ARCAD Clinical Trials Program. The ARCAD clinical trials program: an update and invitation. *Oncologist*. 2012;17(2):188-191. doi:10.1634/theoncologist.2011-0332.

6. Hamers PAH, Elferink MAG, Stellato RK, et al. Informing metastatic colorectal cancer patients by quantifying multiple scenarios for survival time based on real-life data. *Int J Cancer*. 2021;148(2):296-306. doi:10.1002/ijc.33200

7. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(3):145-164. doi:10.3322/caac.21601.

8. Saad ED, Buyse M. Statistical controversies in clinical research: end points other than overall survival are vital for regulatory approval of anticancer agents. *Ann Oncol*. 2016;27(3):373-378. doi:10.1093/annonc/mdv562

9. Broglio KR, Berry DA. Detecting an overall survival benefit that is derived from progression-free survival. *J Natl Cancer Inst*. 2009;101(23):1642-1649. doi:10.1093/jnci/djp369

10. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21(11):1350-1356. doi:10.1038/nm.3967.

11. Rockhold F, Bromley C, Wagner EK, Buyse M. Open science: The open clinical trials data journey. *Clin Trials*. 2019;16(5):539-546. doi:10.1177/1740774519865512

12. Sridhara R, Mandrekar SJ, Dodd LE. Missing data and measurement variability in assessing progression-free survival endpoint in randomized clinical trials. *Clin Cancer Res*. 2013;19(10):2613-2620. doi:10.1158/1078-0432.CCR-12-2938.

13. Denne JS, Stone AM, Bailey-Iacona R, Chen TT. Missing data and censoring in the analysis of progression-free survival in oncology clinical trials. *J Biopharm Stat*. 2013;23(5):951-970. doi:10.1080/10543406.2013.813515.

14. Zhang JJ, Sun Z, Yuan H, Wang M. Alternatives to the Kaplan-Meier estimator of progression-free survival [published online ahead of print, 2020 Sep 5]. *Int J Biostat*. 2020;/j/ijb.ahead-of-print/ijb-2019-0095/ijb-2019-0095.xml. doi:10.1515/ijb-2019-0095.

15. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol*. 2007;25(33):5218-5224. doi:10.1200/JCO.2007.11.8836.

16. Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A. Biomarkers and surrogate end points--the challenge of statistical validation. *Nat Rev Clin Oncol*. 2010;7(6):309-317. doi:10.1038/nrclinonc.2010.43.

17. Burzykowski T, Buyse M, Piccart-Gebhart MJ, et al. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol*. 2008;26(12):1987-1992. doi:10.1200/JCO.2007.10.8407.

18. Michiels S, Pugliano L, Marguet S, et al. Progression-free survival as surrogate end point for overall survival in clinical trials of HER2-targeted agents in HER2-positive metastatic breast cancer. *Ann Oncol*. 2016;27(6):1029-1034. doi:10.1093/annonc/mdw132.

19. Michiels S, Saad ED, Buyse M. Progression-Free Survival as a Surrogate for Overall Survival in Clinical Trials of Targeted Therapy in Advanced Solid Tumors. *Drugs*. 2017;77(7):713-719. doi:10.1007/s40265-017-0728-y.

20. Haslam A, Hey SP, Gill J, Prasad V. A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. *Eur J Cancer*. 2019;106:196-211. doi:10.1016/j.ejca.2018.11.012.

21. Chibaudel B, Bonnetain F, Shi Q, et al. Alternative end points to evaluate a therapeutic strategy in advanced colorectal cancer: evaluation of progression-free survival, duration of disease control, and time to failure of strategy--an Aide et Recherche en Cancerologie Digestive Group Study. *J Clin Oncol*. 2011;29(31):4199-4204. doi:10.1200/JCO.2011.35.5867.

22. Shi Q, de Gramont A, Grothey A, et al. Individual patient data analysis of progression-free survival versus overall survival as a first-line end point for metastatic colorectal cancer in modern randomized trials: findings from the analysis and research in cancers of the digestive system database. *J Clin Oncol*. 2015;33(1):22-28. doi:10.1200/JCO.2014.56.5887.

23. Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res*. 2008;17(3):303-340. doi:10.1177/0962280207082719

24. Rutkowski J, Saad ED, Burzykowski T, Buyse M, Jassem J. Chronological Trends in Progression-Free, Overall, and Post-Progression Survival in First-Line Therapy for Advanced NSCLC. *J Thorac Oncol*. 2019;14(9):1619-1627. doi:10.1016/j.jtho.2019.05.030

25. Siena S, Peeters M, Van Cutsem E, et al. Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. *Br J Cancer*. 2007;97(11):1469-1474. doi:10.1038/sj.bjc.6604053.

26. Griebsch I, Palmer M, Fayers PM, Ellis S. Is progression-free survival associated with a better health-related quality of life in patients with lung cancer? Evidence from two randomised trials with afatinib. *BMJ Open*. 2014;4(10):e005762. Published 2014 Oct 31. doi:10.1136/bmjopen-2014-005762.

27. Fallowfield LJ, Fleissig A. The value of progression-free survival to patients with advanced-stage cancer. *Nat Rev Clin Oncol*. 2011;9(1):41-47. Published 2011 Oct 18. doi:10.1038/nrclinonc.2011.156.

28. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Meta-Analysis Group in Cancer. *Lancet*. 2000;356(9227):373-378. doi:10.1016/s0140-6736(00)02528-9.

29. Piessevaux H, Buyse M, Schlichting M, et al. Use of early tumor shrinkage to predict long-term outcome in metastatic colorectal cancer treated with cetuximab. *J Clin Oncol*. 2013;31(30):3764-3775. doi:10.1200/JCO.2012.42.8532.

30. Cremolini C, Loupakis F, Antoniotti C, et al. Early tumor shrinkage and depth of response predict long-term outcome in metastatic colorectal cancer patients treated with first-line chemotherapy plus bevacizumab: results from phase III TRIBE trial by the Gruppo Oncologico del Nord Ovest. *Ann Oncol*. 2015;26(6):1188-1194. doi:10.1093/annonc/mdv112.

31. Heinemann V, Stintzing S, Modest DP, Giessen-Jung C, Michl M, Mansmann UR. Early tumour shrinkage (ETS) and depth of response (DpR) in the treatment of patients with metastatic colorectal cancer (mCRC). *Eur J Cancer*. 2015;51(14):1927-1936. doi:10.1016/j.ejca.2015.06.116.

32. Sommeijer DW, Shi Q, Meyer J, et al. Prognostic value of early objective tumor response (EOTR) to first line systemic therapy in metastatic colorectal cancer (mCRC): IPD (IPD)

meta-analysis of randomized trials from the ARCAD database. *Journal of Clinical Oncology* Supplement (May 20, 2013) 3520-3520. doi: 10.1200/jco.2013.31.15_suppl.3520.

33. Sommeijer DW, Shi Q, Saad ED, et al. ASCO Early predictors of prolonged overall survival (OS) in patients (pts) on first-line chemotherapy (CT) for metastatic colorectal cancer (mCRC): An ARCAD study with IPD (IPD) on 10,962 pts. *Journal of Clinical Oncology* 2014; 32, no. 15_suppl (May 20, 2014) 3538-3538. doi: 10.1200/jco.2014.32.15_suppl.

34. Saad ED, Coart E, Sommeijer, et al. Early predictors of improved long-term outcomes in first-line antiangiogenics plus chemotherapy (anti-ANG/CT) in metastatic colorectal cancer (mCRC): Analysis of individual patient (pt) data from the ARCAD database. *Journal of Clinical Oncology* 2014 (32 Supplement) 3578. doi: 10.1200/jco.2014.32.15_suppl.3578

35. Burzykowski T, Molenberghs G, Buyse M. *The Evaluation of Surrogate Endpoints*. New York, NY: Springer; 2005.

36. Coart E, Saad ED, Shi Q, et al. Trial-level association between response-based endpoints (rbes) and progression-free (pfs)/overall survival (os) in first-line therapy for metastatic colorectal cancer (mcrc) in the arcad database. *Journal of Clinical Oncology* 2015; 33 (Supplement): 666. doi: 10.1200/jco.2015.33.3_suppl.666

37. Burzykowski T, Coart E, Saad ED, et al. Evaluation of Continuous Tumor-Size-Based End Points as Surrogates for Overall Survival in Randomized Clinical Trials in Metastatic Colorectal Cancer. *JAMA Netw Open*. 2019;2(9):e1911750. Published 2019 Sep 4. doi:10.1001/jamanetworkopen.2019.11750.

38. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct Target Ther*. 2020;5(1):22. Published 2020 Mar 20. doi:10.1038/s41392-020-0116-z.

26

39. André T, Shiu KK, Kim TW, et al. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med*. 2020;383(23):2207-2218. doi:10.1056/NEJMoa2017699.

40. Morse MA, Hochster H, Benson A. Perspectives on Treatment of Metastatic Colorectal Cancer with Immune Checkpoint Inhibitor Therapy. *Oncologist*. 2020;25(1):33-45. doi:10.1634/theoncologist.2019-0176.

41. Kaplan R, Maughan T, Crook A, et al. Evaluating many treatments and biomarkers in oncology: a new design. *J Clin Oncol*. 2013;31(36):4562-4568. doi:10.1200/JCO.2013.50.7905.

42. Gill J, Prasad V. When are randomized controlled trials needed to assess novel anticancer drugs? An illustration based on the development of selpercatinib, a RET inhibitor. *Ann Oncol*. 2020;31(3):328-330. doi:10.1016/j.annonc.2019.11.021.

43. Texier M, Rotolo F, Ducreux M, Bouché O, Pignon JP, Michiels S. Evaluation of Treatment Effect with Paired Failure Times in a Single-Arm Phase II Trial in Oncology. *Comput Math Methods Med*. 2018;2018:1672176. Published 2018 Jan 11. doi:10.1155/2018/1672176.

44. Italiano A, Nanda S, Briggs A, et al. Larotrectinib versus Prior Therapies in Tropomyosin Receptor Kinase Fusion Cancer: An Intra-Patient Comparative Analysis. *Cancers (Basel)*. 2020;12(11):3246. Published 2020 Nov 4. doi:10.3390/cancers12113246.

45. Sobrero AF, Pastorino A, Sargent DJ, Bruzzi P. Raising the bar for antineoplastic agents: how to choose threshold values for superiority trials in advanced solid tumors. *Clin Cancer Res*. 2015;21(5):1036-1043. doi:10.1158/1078-0432.CCR-14-1505.

46. Cherny NI, Dafni U, Bogaerts J, et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol*. 2017;28(10):2340-2366. doi:10.1093/annonc/mdx310

47. Cherny NI, de Vries EGE, Dafni U, et al. Comparative Assessment of Clinical Benefit Using the ESMO-Magnitude of Clinical Benefit Scale Version 1.1 and the ASCO Value Framework Net Health Benefit Score. *J Clin Oncol*. 2019;37(4):336-349. doi:10.1200/JCO.18.00729

48. André T, Shiu KK, Kim TW, et al. Final overall survival for the phase III KN177 study: Pembrolizumab versus chemotherapy in microsatellite instability-high/mismatch repair deficient (MSI-H/dMMR) metastatic colorectal cancer (mCRC). J Clin Oncol 39: 2021 (suppl; abstr 3500)

49. Renfro LA, Goldberg RM, Grothey A, et al. Clinical Calculator for Early Mortality in Metastatic Colorectal Cancer: An Analysis of Patients From 28 Clinical Trials in the Aide et Recherche en Cancérologie Digestive Database. *J Clin Oncol*. 2017;35(17):1929-1937. doi:10.1200/JCO.2016.71.5771.

50. Sjoquist KM, Renfro LA, Simes RJ, et al. Personalizing Survival Predictions in Advanced Colorectal Cancer: The ARCAD Nomogram Project. *J Natl Cancer Inst*. 2018;110(6):638-648. doi:10.1093/jnci/djx253.

51. Lieu CH, Renfro LA, de Gramont A, et al. Association of age with survival in patients with metastatic colorectal cancer: analysis from the ARCAD Clinical Trials Program. *J Clin Oncol*. 2014;32(27):2975-2984. doi:10.1200/JCO.2013.54.9329.

52. Renfro LA, Loupakis F, Adams RA, et al. Body Mass Index Is Prognostic in Metastatic Colorectal Cancer: Pooled Analysis of Patients From First-Line Clinical Trials in the ARCAD Database. *J Clin Oncol*. 2016;34(2):144-150. doi:10.1200/JCO.2015.61.6441.

53. Franko J, Shi Q, Meyers JP, et al. Prognosis of patients with peritoneal metastatic colorectal cancer given systemic therapy: an analysis of individual patient data from prospective randomised trials from the Analysis and Research in Cancers of the Digestive

System (ARCAD) database. *Lancet Oncol*. 2016;17(12):1709-1719. doi:10.1016/S1470-2045(16)30500-9.

54. van Rooijen KL, Shi Q, Goey KKH, et al. Prognostic value of primary tumour resection in synchronous metastatic colorectal cancer: Individual patient data analysis of first-line randomised trials from the ARCAD database. *Eur J Cancer*. 2018;91:99-106. doi:10.1016/j.ejca.2017.12.014.

55. Salem ME, Yin J, Weinberg BA, et al. Clinicopathological differences and survival outcomes with first-line therapy in patients with left-sided colon cancer and rectal cancer: Pooled analysis of 2879 patients from AGITG (MAX), COIN, FOCUS2, OPUS, CRYSTAL and COIN-B trials in the ARCAD database. *Eur J Cancer*. 2018;103:205-213. doi:10.1016/j.ejca.2018.08.020.

56. Henriques J, Vernerey D, de Gramont A, et al. Prognosis of lung metastases in patients with metastatic colorectal cancer: an ARCAD meta-analysis *Annals Of Oncology* 2016 Jun; 27:122. doi: https://doi.org/10.1093/annonc/mdw198.12.

57. Ou F-S, Hubbard JM, Kasi PM, et al. Heterogeneity in early lesion changes on treatment as a marker of poor prognosis in patients (pts) with metastatic colorectal cancer (mCRC) treated with first line systemic chemotherapy ± biologic: Findings from 9,092 pts in the ARCAD database. *Journal of Clinical Oncology* 2017 35:15 (Supplement) 3535-3535. doi: 10.1200/JCO.2017.35.15_

58. Ou FS, Lou Y, Van Cutsem Eet al. Evaluation of lesion-based response at 12 weeks (LBR12) of treatment (Rx) in metastatic colorectal cancer (mCRC): Findings from 9,092 patients (pts) in the ARCAD database. *Journal of Clinical Oncology* 2018; 36 (Supplement). doi:10.1200/JCO.2018.36.4_suppl.612

59. Gulhati P, Yin J, Pederson L, et al. Threshold Change in CEA as a Predictor of Non-Progression to First-Line Systemic Therapy in Metastatic Colorectal Cancer Patients With Elevated CEA. *J Natl Cancer Inst*. 2020;112(11):1127-1136. doi:10.1093/jnci/djaa020.

60. Araujo DV, Bratman SV, Siu LL. Designing circulating tumor DNA-based interventional clinical trials in oncology. *Genome Med*. 2019;11(1):22. Published 2019 Apr 19. doi:10.1186/s13073-019-0634-x.

61. Reece M, Saluja H, Hollington P, et al. The Use of Circulating Tumor DNA to Monitor and Predict Response to Treatment in Colorectal Cancer. *Front Genet*. 2019;10:1118. Published 2019 Nov 21. doi:10.3389/fgene.2019.01118.

62. Mao Y, Chen B, Wang H, et al. Diagnostic performance of magnetic resonance imaging for colorectal liver metastasis: A systematic review and meta-analysis. *Sci Rep*. 2020;10(1):1969. Published 2020 Feb 6. doi:10.1038/s41598-020-58855-1.

63. Bonnetain F, Borg C, Adams RR, et al. How health-related quality of life assessment should be used in advanced colorectal cancer clinical trials. *Ann Oncol*. 2017;28(9):2077-2085. doi:10.1093/annonc/mdx191.

64. Zalcberg JR, Shi Q, Ferraro DA, et al. Impact of overall severity of adverse events (AEs) on long term outcomes in metastatic colorectal cancer (mCRC) patients (pts) treated with first line systemic chemotherapy: Findings from 3,971 pts in the ARCAD database. *J Clin Oncol* 2017; 35 (Supplenmt) May 2017 3582-3582.doi: 10.1200/JCO.2017.35.15_suppl.3582 [Abstract]

65. Ferraro DA, Zalcberg JR, Shi Q, et al. Associations of incidence of common adverse events (AEs) and survival outcomes in metastatic colorectal cancer (mCRC) patients (pts) treated with first line chemotherapy: Findings from 9,812 pts in the ARCAD database. *J Clin Oncol* 2018; 36 (Supplement) Abstract 612.

66. Guibert N, Jones G, Beeler JF, et al. Targeted sequencing of plasma cell-free DNA to predict response to PD1 inhibitors in advanced non-small cell lung cancer. *Lung Cancer*. 2019;137:1-6. doi:10.1016/j.lungcan.2019.09.005.

67. Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics [published correction appears in Lancet Oncol. 2019 May;20(5):e242]. *Lancet Oncol*. 2017;18(3):e143-e152. doi:10.1016/S1470-2045(17)30074-8.

68. Mohile SG, Xian Y, Dale W, et al. Association of a cancer diagnosis with vulnerability and frailty in older Medicare beneficiaries. *J Natl Cancer Inst*. 2009;101(17):1206-1215. doi:10.1093/jnci/djp239.

69. Kawahara T, Fukuda M, Oba K, Sakamoto J, Buyse M. Meta-analysis of randomized clinical trials in the era of individual patient data sharing. *Int J Clin Oncol*. 2018;23(3):403-409. doi:10.1007/s10147-018-1237-z.

Table

Table 1. Trial-level association (coefficient of determination, R^2) between response-based endpoints and progression-free survival (PFS/overall survival (OS) in first-line therapy[a]

Treatment	Early tumor shrinkage		Early objective tumor response		Early non-progression rate		Best overall response	Confirmed Response
	At 6/8/9 weeks	At 12 weeks	At 6/8/9 weeks	At 12 weeks	At 6/8/9 weeks	At 12 weeks		
All[b]								
PFS	0.25 (0.00-0.72)	0.01 (0.00-0.16)	0.31 (0.00-0.93)	0.01 (0.00-0.14)	—	—	0.16 (0.00-0.57)	0.06 (0.00-0.34)
OS	0.25 (0.00-0.95)	0.07 (0.00-0.49)	0.92 (0.00-1.00)	0.23 (0.00-0.80)	0.26 (0.00-0.78)[e]	0.65 (0.32-0.97)[e]	0.07 (0.00-0.56)	0.05 (0.00-0.54)
Chemotherapy alone[c]								
PFS	0.42 (0.00-1.00)	0.17 (0.00-0.74)	0.56 (0.10-1.00)[e]	0.05 (0.00-0.44)	—	—	0.34 (0.00-1.00)	0.03 (0.00-1.00)
OS	0.20 (0.00-0.97)	0.08 (0.00-0.58)	0.59 (0.16-1.00)[e]	0.18 (0.00-0.88)	0.26 (0.00-0.78)[e]	0.65 (0.32-0.97)[e]	0.07 (0.00-0.56)	0.05 (0.00-0.54)
Chemotherapy + targeted agents[d]								
PFS	0.13 (0.00-0.67)	0.08 (0.00-0.72)	0.13 (0.00-0.71)	0.01 (0.00-0.13)[e]	—	—	0.13 (0.00-0.66)	0.17 (0.00-0.74)
OS	0.19 (0.00-1.00)	0.24 (0.00-1.00)	0.00 (0.00-0.06)[e]	0.00 (0.00-0.02)[e]	0.17 (0.00-1.00)	0.73 (0.00-1.00)	0.26 (0.00-0.66)	0.18 (0.00-0.60)

[a] From Coart et al (36), with permission.

[b] All patients; 19 trials, 37 treatment comparisons, 12167 patients

[c] Chemotherapy only; 10 trials, 18 treatment comparisons, 5123 patients.

[d] Chemotherapy plus at least one targeted therapy; 9 trials, 19 treatment comparisons, 7044 patients

[e] Unadjusted R^2. Unless indicated, R^2 is adjusted for estimation errors.

Figure Titles and Legends

Figure 1. Correlation of PFS and overall survival in 22 first-line advanced colorectal cancer trials published from 2003 to 2012. From Shi Q, de Gramont A, Grothey A, et al. Individual patient data analysis of progression-free survival versus overall survival as a first-line end point for metastatic colorectal cancer in modern randomized trials: findings from the analysis and research in cancers of the digestive system database. *J Clin Oncol.* 2015;33(1):22-28 (22), with permission. (**A**) Correlation between progression-free survival (PFS) at 6 months and overall survival (OS) at 12 months at the treatment arm level. (**B**) Correlation between treatment effects on PFS and OS. Circles indicate treatment arm with nonbiologic agents only; triangles indicate treatment arm with biologic agents; diagonal lines indicate 95% prediction limits. (B) Log scale was used for x- and y-axes. Horizontal line corresponds to the hazard ratio (HR) for OS of 1. The vertical line corresponds to the HR for PFS of 1. PFS = progression-free survival; OS = overall survival; HR = hazard ratio.

Figure 2. Association between early responses and overall survival in 16 first-line trials.
From Sommeijer DW, Shi Q, Saad ED, et al. ASCO Early predictors of prolonged overall survival (OS) in patients (pts) on first-line chemotherapy (CT) for metastatic colorectal cancer (mCRC): An ARCAD study with IPD (IPD) on 10,962 pts. *Journal of Clinical Oncology* 2014; 32, no. 15_suppl (May 20, 2014) 3538-3538., with permission. Best Overall Response, complete or partial response by RECIST within initial 26 weeks of treatment; C-index (C) used for comparing the prediction accuracy of candidate surrogate endpoints, with values closer to 1 indicating better prediction accuracy; Confirmed Response, complete or partial response by RECIST confirmed at least 4 weeks later during initial 26 weeks of treatment. The P values were calculated using 2-sided tests. The error bars represent the 95% confidence intervals (CIs) corresponding to the hazard ratios. ETS = early tumor

33

shrinkage; EOTR = early objective tumor response; EnPD = no early progression; RECIST = Response Evaluation Criteria in Solid Tumours.

Figure 3. Trial-level association between response-based endpoints (time to nadir and depth of nadir) and overall survival. Adapted from Burzykowski T, Coart E, Saad ED, et al. Evaluation of Continuous Tumor-Size-Based End Points as Surrogates for Overall Survival in Randomized Clinical Trials in Metastatic Colorectal Cancer. *JAMA Netw Open*. 2019;2(9):e1911750. (37), with permission. Hazard ratios (HRs) of overall survival (OS) associated with time to nadir and depth of nadir in chemotherapy-only-treated patients. The difference in nadir is the difference between the model-estimated mean relative tumor-size change at nadir (relative to baseline) in each contrast. The line indicates weighted regression; the sizes of the circles are proportional to the total sample sizes of the corresponding contrasts. HR = hazard ratio; OS = overall survival.

Figure 4. Hypothesis development, clinical trials, treatment qualification and research feedback in advanced colorectal cancer. Trial hypotheses (left box) are formed on the basis of basic biological knowledge, available treatments and phase 1 clinical trials results, and informed by the choice of evaluation criteria. Phase 2-3 trials (center box) may include multi-arm multi strategy (MAMS) trials, small, rapid trials for potentially highly effective treatments, and larger and longer trials for more modestly effective treatments. In trials investigating first-line treatments, if hazard ratio (HR) is used to quantify the effect, a PFS HR of approximately <0.5 compared with standard of care should be considered a major advance likely leading to rapid qualification of a treatment in the specific patient group in which the effect was achieved. Candidate treatments with negligible response or HRs approximately >0.8 for PFS should in general be discarded. The clinical value of treatments

34

should be upgraded or downgraded according to toxicity, QOL and cost considerations (right

box). Clinical trials data should be collated in multi-trial databases – in the advanced

colorectal cancer setting, the ARCAD database. Real-life performance of treatments should

be monitored through registries and electronic patient records. Analysis of this information

should feed back into novel hypotheses and clinical trials, refine the use of evaluation criteria,

and potentially directly inform the introduction of new therapeutics. PFS = progression-free

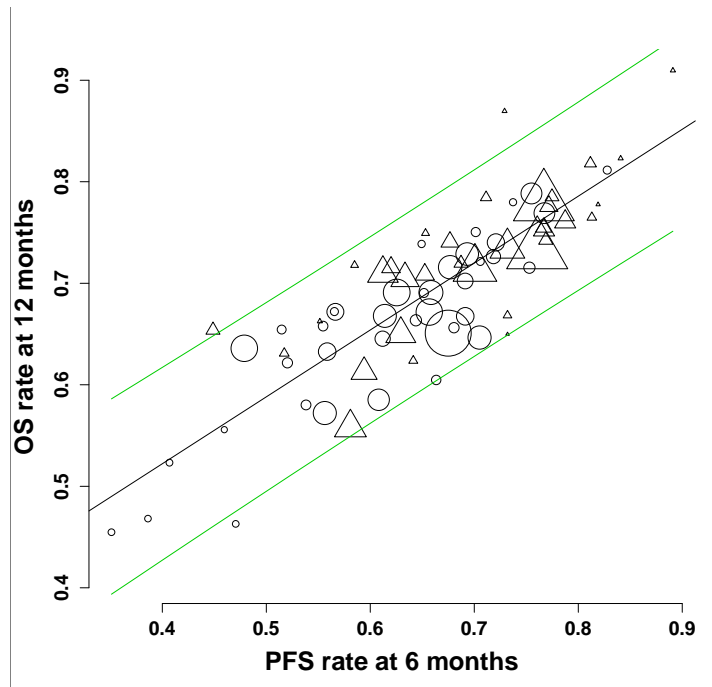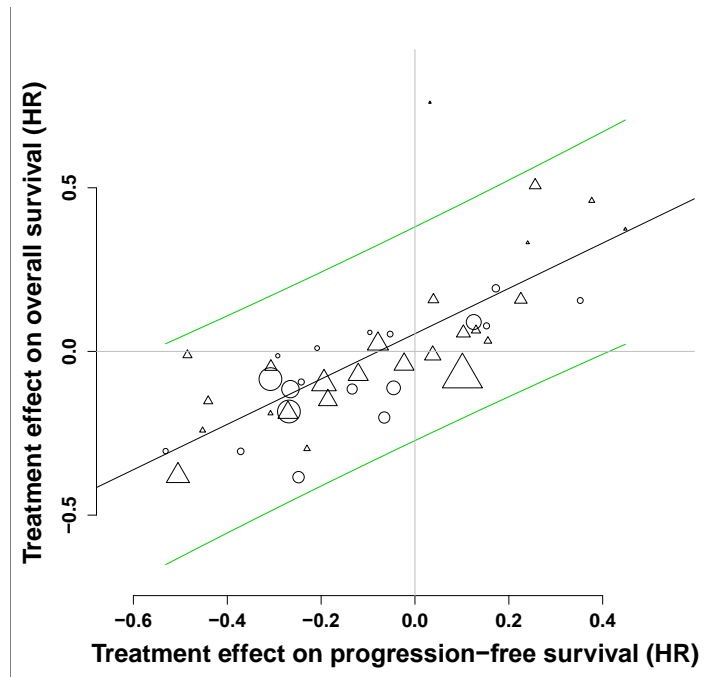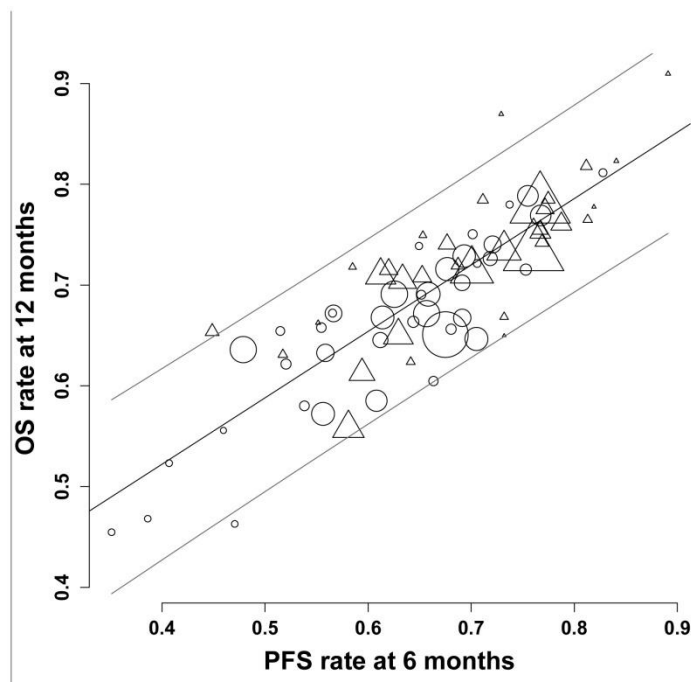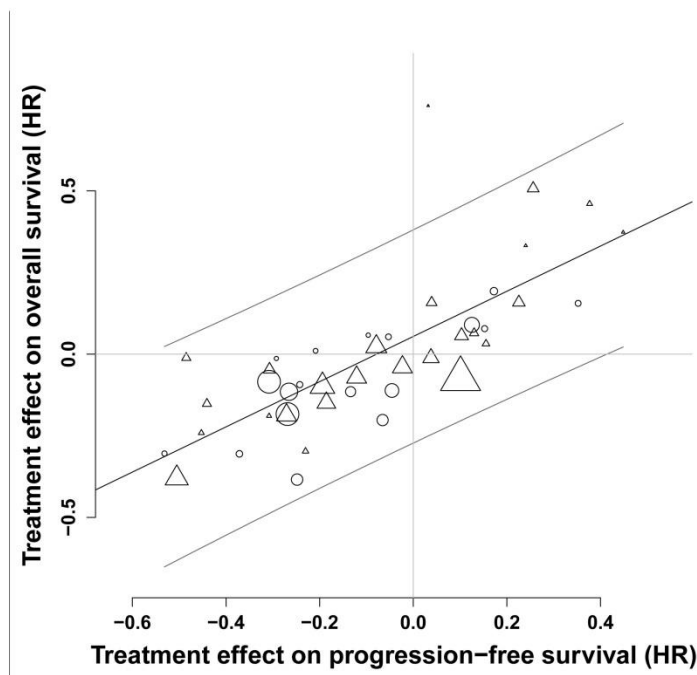survival; MAMS = multi-arm, multi-strategy; QOL = quality of life.

Figure 1 color--FINAL

Figure 1

A

B

.

Figure 1 BW--FINAL

Figure 1

A

B

Figure 2--FINAL

Figure 2

	Events/Total	HR (95% CI)	P-value	C (95% CI)
ETS				
6 weeks	1726/2484	0.61 (0.56-0.68)	<.0001	0.56 (0.52-0.61)
8/9 weeks	954/1175	0.67 (0.59-0.77)	<.0001	0.56 (0.52-0.60)
12 weeks	3524/4633	0.63 (0.59-0.68)	<.0001	0.57 (0.53-0.61)
EOTR				
6 weeks	1736/2488	0.61 (0.54-0.69)	<.0001	0.55 (0.52-0.58)
8/9 weeks	829/1101	0.60 (0.52-0.69)	<.0001	0.58 (0.55-0.61)
12 weeks	3207/4242	0.61 (0.56-0.65)	<.0001	0.58 (0.54-0.61)
EnPD				
6 weeks	1736/2488	0.31 (0.27-0.36)	<.0001	0.56 (0.54-0.57)
8/9 weeks	829/1101	0.24 (0.19-0.30)	<.0001	0.56 (0.54-0.57)
12 weeks	3207/4242	0.35 (0.32-0.39)	<.0001	0.57 (0.55-0.59)
Best Overall Response	4768/6219	0.59 (0.56-0.63)	<.0001	0.58 (0.55-0.61)
Confirmed Response	2926/4038	0.65 (0.59-0.70)	<.0001	0.55 (0.52-0.59)

Figure 3--FINAL

Figure 3

Figure 4 BW--FINAL

Figure 4

Figure 4 color--FINAL

Figure 4