IMPORANCE The past 2 decades have witnessed progress in the management of metastatic colorectal cancer (mCRC) with more effective agents and better surgical, medical, and supportive care. While substantial progress has been made, much more must be achieved to prolong the lives of patients.

OBJECTIVE To conduct a systematic review to ascertain what percentage of the life expectancy gain in locally advanced and mCRC over the past 2 decades is due to novel therapies vs improvements in supportive care or secular trends and to thus inform treatment development strategies.

EVIDENCE REVIEW We searched Cochrane Controlled Trials Register, Medline, Embase, CancerLit, and Healthstar electronic databases for trials covering the period 1993 to 2015, scanned reference lists of articles, and searched recent conference abstracts. Ninety-six phase 3 trials and large (>50 patients) phase 2 trials in mCRC were examined. Outcomes evaluated in the experimental arms (EAs) and control arms (CAs) included overall response rate, stable disease, progression-free survival (PFS), and overall survival (OS).

FINDINGS Over the period covered by the studies, the OS in EAs increased at a mean (95% CI) rate of 0.80 (0.67-0.93) mo/y. Importantly, OS in the CAs improved 0.63 (0.51-0.75) mo/y, reflecting in part the use of experimental regimens in subsequent studies. Chemotherapy contributed only partly to the gains in OS, given that (1) mean (95% CI) improvements in PFS were only 0.31 (0.22-0.39) mo/y in the EAs and 0.23 (0.15-0.31) mo/y in CAs; (2) gains in survival not directly attributable to the protocol were greater than gains in PFS (0.46 [0.36-0.57] mo/y in EAs and 0.39 [0.29-0.49] mo/y in CAs; and (3) effects on OS were much lower in second-line trials (median [interquartile range] response rates, 8.6% [0%-11.0%] in EAs and 7.5% [3.8%-12.8%] in CAs) compared with first-line trials (39.5% [24.0%-50.2%] for EAs and 29.4% [16.4%-39.4%] for CAs).

CONCLUSIONS AND RELEVANCE The OS of patients with mCRC has improved gradually over the past 2 decades, with gains from chemotherapy occurring alongside gains from lead-time bias and improved locoregional approaches and supportive care. Gains from first-line therapies have been modest but consistent; however, gains from second-line therapies have been disappointing. We believe that future progress will be greater if emphasis is placed on enrolling patients in experimental trials to explore and develop alternative first-line regimens and better second-line therapies.
Colorectal cancer mortality rates have declined steadily over the last 2 decades, and this is widely believed to be a consequence of improvements in the therapy of metastatic colorectal cancer (mCRC). For this reason, mCRC is often cited as an example of the value of gradual progress in oncology, with gains built on serial successes. Despite enthusiasm for the effects of novel chemotherapies, to our knowledge, there exists no empirical study documenting what percentage of the gain in mCRC life expectancy is due to novel therapies vs improvements in supportive care or secular trends. We sought to examine this issue by looking at phase 3 and large phase 2 mCRC trials published since the early 1990s. Our goals were to assess the extent of progress, to understand what contributed to that progress, and to draw lessons that might inform treatment development strategies.

Methods

Study Selection
We identified phase 3 and large (>50 patients) phase 2 randomized clinical trials comparing chemotherapy regimens for patients with locally advanced (unresectable) or mCRC to quantify survival benefit over time with first-line and subsequent therapies. Patients might also have received previous chemotherapy as adjuvant treatment after surgery or for metastatic disease.

Search Strategy
We searched the following electronic databases for eligible trials covering the period from 1993 to 2015: Cochrane Controlled Trials Register, Medline, Embase, CancerLit, and Healthstar. We also scanned the reference lists of original and review articles and all primary studies identified and searched recent conference abstracts to identify further eligible trials. Key words used during the search included “colorectal cancer” or “colon cancer” or “rectal cancer” and “phase III” or “phase II” or “randomized.” Our search strategy uncovered approximately 400 articles, which were retrieved and scanned by hand. Ninety-six trials that enrolled 26 561 patients were included in this analysis. Six of the trials were designed as noninferiority or equivalence trials.

Review Procedures
Two reviewers independently extracted data on study methods, participants, therapies, and outcomes from all eligible trials; differences in how to classify the experimental agent in a randomized trial were resolved by discussion. The principal outcomes examined in the experimental arms (EAs) and control arms (CAs) included overall response rate (ORR), stable disease, progression-free survival (PFS), overall survival (OS), and survival not directly attributable to the protocol (SNAP). Trials were classified according to whether chemotherapy was administered as first-line, second-line, or subsequent-line treatment.

Statistical Analysis
Data were analyzed according to the median date for the period during which each trial was conducted and the subsequent date of publication. Since this did not result in substantial differences, we report outcomes by date of publication.

Data were initially analyzed by both median time of trial enrollment and date of trial publication. Figure 1 shows the gradual improvements in OS of patients with mCRC in the past 2 decades. Improvements in OS were observed in both EAs and CAs, but the slopes of the regression lines (95% CIs) demonstrate better rates of improvement in EAs (0.80 [0.66-0.95]) than in CAs (0.63 [0.51-0.75]). Since analysis by median enrollment date did not yield important differences, all subsequent data are reported by trial publication date.

With the hope of eventually discerning how much of the OS gain in the last 2 decades can be ascribed to better chemotherapy and how much to other factors, we turned to the gains in PFS over time. We did this to capture the benefit accrued from the experimental therapy. Prior evidence suggests that while PFS contributes to OS, PFS is a measure of a drug’s effect while administered and is not a surrogate for OS. As was the case with OS, the data demonstrate incremental increases in PFS (Figure 2), indicating that therapies have be-
come gradually more effective in delaying progression. However, the slope (95% CI) of improvement during these 2 decades is more modest than that seen in the OS analysis: 0.31 (0.22-0.39) mo/y and 0.23 (0.15-0.31) mo/y for the EAs and CAs, respectively. Indeed, most of the improvement in OS during this time resulted from survival time not directly attributable to the protocol therapy (SNAP). For the EAs and CAs, rates of improvement (95% CI) in SNAP were 0.46 (0.36-0.57) mo/y and 0.39 (0.29-0.49) mo/y, respectively.

As shown in Figure 3, survival gains were highly correlated with the drugs’ activity as measured by ORR (complete response rate + partial response rate). Better therapies yielded higher ORRs, and these in turn resulted in longer PFS. Importantly, similar gains in the duration of PFS as a function of the ORR were found in both the EAs and CAs. As measured by slopes (95% CIs) representing gains or decreases in fractions of a month of PFS per 1% change in the ORR, the EAs showed an increase of 0.13 (0.12-0.15) PFS months, and the CAs showed an increase of 0.11 (0.09-0.13) PFS months. This finding is not surprising because in most cases, control therapies were former experimental therapies that gained traction in the medical community, bolstered by the results of prior studies. While a positive correlation was also observed between OS and ORR, gains in OS did not positively correlate with the rate of stable disease. In fact, an inverse correlation was observed in both the EAs and CAs with increasing stable disease rate associated with decreasing OS, likely because the higher the stable disease rate, the lower the ORR (Figure 3).

The steeper slopes of improvement over time seen in OS compared with PFS (compare Figure 1 with Figure 2) could be explained if successively better therapies had an effect both while administered—manifested as an improvement in PFS—and after their discontinuation—manifested as an additional effect on OS in addition to its effect on PFS. The continuing effect after discontinuation is unlikely, since within a single study, the gains in PFS, defined as PFS of the EA minus PFS of the CA, paralleled the gains in OS (Figure 4). The slope (95% CI) of the regression line, 1.39 (1.08-1.71), indicates that on average, the gains in OS are similar to the gains in PFS within the study. Thus, the greater gains in OS over time cannot be ascribed to greater gains within a study. These gains that have occurred outside of the study, have been gradual and continuous over the past 2 decades, and indeed have contributed more to the gains in OS than the gains achieved within a study. We saw in Figures 1

Graphs show improvements in overall survival as a function of date of median enrollment (left) or date of publication (right) in both the experimental arms (top) and control arms (bottom) of the clinical trials evaluated in this analysis. The slopes of the regression lines (solid lines) (95% CIs [dotted lines]) demonstrate the rates of improvement. Similar results were observed by time of enrollment and date of trial publication, and so all data in subsequent figures are presented by trial publication date.
and 2 respective gains in OS for the EAs and CAs of 0.80 mo/y and 0.63 to 0.64 mo/y; in PFS of 0.31 mo/y and 0.23 mo/y; and in SNAP of 0.46 mo/y and 0.39 mo/y that represent gains achieved in addition to the gains in PFS.

To discern if any of the OS advantages could be ascribed to salvage therapies, we examined the results of second- and third-line therapies. These data consisted largely of second-line evaluations because only 3 studies of third-line therapies meeting our search criteria could be found. The data in the upper 4 panels of Figure 5 show that second- and third-line therapies are much less effective than first-line regimens, as evidenced by much lower ORRs. Compared with the results in first-line regimens (Figure 3 and Figure 5) where the median (interquartile range [IQR]) response rates were 39.5% (24.0%-50.2%) for the EAs and 29.4% (16.4%-39.4%) for CAs, the response rates found in second-line therapies were 8.6% (0%-11.0%) in EAs and 7.5% (3.8%-12.8%) in CAs. However, second-line PFS correlations with ORR mimicked those in first-line studies with PFS/ORR slopes (95% CI) of 0.13 (0.03-0.24) and 0.14 (0.05-0.23) for the EAs and CAs, respectively, indicating a 0.13- and 0.14-month increase in PFS for every 1% increase in ORR (Figure 3). Furthermore, as shown in the lower 2 panels of Figure 5, it is difficult to convincingly demonstrate gains in OS from second-line therapies over the past 2 decades.

Finally, examination of ORR by trial arm indicates an apparent loss of efficacy of first-line regimens when given as the CA compared with when given as the EA. As shown in the eFigure in the Supplement, this observation was made of the median (IQR) response rates in study regimens containing both irinotecan (34.0% [32.0%-38.7%] for the CAs and 43.0% [38.6%-47.8%] for the EAs) and oxaliplatin (41.0% [36.0%-43.5%] for the CAs and 48.0% [43.5%-54.1%] for the EAs), with the difference between CA and EA somewhat more pronounced in the irinotecan-containing regimens.

Discussion

The present study provides evidence of the continuous progress made in the therapy of mCRC over the past 2 decades. Focusing on studies that evaluated first-line therapies in mCRC, we noted a gradual but continued improvement in OS. A prolongation of OS of 0.82 mo/y can be documented, so the median OS of patients with mCRC is now about 24 months. The analysis reveals other results as well: (1) survival gains were made in the CAs of the studies during the same period; (2) a substantial fraction of the OS gains can likely be attributed to factors other than the chemotherapy regimens adminis-
We observed continued gradual improvements in OS in both the EAs and the CAs over the past 2 decades, not surprisingly with a better slope of improvement in the EAs than in the CAs. Improvement in CAs can be explained, at least in part, by the fact that once a new experimental therapy has been established as effective, that therapy becomes the control in subsequent trials. In the cases of experimental therapy, greater improvement in OS might be explained by a better therapy that might be administered more aggressively and also subsequent therapies. Regarding more aggressive administration, we found an apparent loss of efficacy when regimens were administered as the CAs compared with their administration as the EAs, with the difference somewhat more pronounced in irinotecan-containing regimens. A possible explanation for this would be that when administering the EA, investigators might be more aggressive in giving these often toxic therapies, and when it moves to the CA, it is less aggressively administered.

Therapeutic gains are highly correlated with a drug's activity as measured by the overall response rate (complete response + partial response) (top 4 graphs). No such correlation or even a negative correlation is seen for stable disease (bottom 2 graphs). In all panels, the slopes (solid lines) (95% CIs [dotted lines]) represent the gains or decreases in fractions of a month of progression-free or overall survival per 1% change in the overall response rate or rate of stable disease.
Although few second-line or subsequent-line trials have been published, the data clearly demonstrate these therapies are much less active than regimens used in first-line treatment. We would note here that differences in performance between first- and subsequent-line therapies are not surprising, since in all cancers one commonly sees the effectiveness of therapies fall as consecutive regimens are given. However, one can thus begin to understand how it is not possible to expect these subsequent lines of therapy to contribute more to the OS than was contributed by the first-line therapies—a conclusion supported by the fact that there has been little or no gain in OS with second- and subsequent-line therapies.

But beyond the gains from the chemotherapy regimens, the data unequivocally establish the important contribution of factors other than chemotherapy in the improvements in OS observed in both the CAs and EAs. These most likely include (1) lead-time bias, if over time patients have enrolled in clinical trials earlier in their disease course; (2) better and more sophisticated patient care; and (3) the ability to support patients for longer at the end of life. The problem of lead-time bias is well recognized as a problem in many cancers and has also been recently raised with regard to mCRC by others who noted that “improved surveillance and imaging techniques may have introduced lead-time bias in newer-era trials, where recurrences were simply detected earlier.”

That this is true in mCRC is supported by studies showing that improved imaging can find smaller metastatic disease burden earlier. These patients, who would have been considered disease free before more refined imaging technology was available, are now diagnosed with small foci of metastatic disease and are enrolled in clinical trials at an earlier point in the natural history of the disease. This would apply both to patients whose disease recurs after an original resection and to those with de novo metastatic disease in whom detection of any tumor has improved. In addition, patients are more quickly enrolled in randomized clinical trials as a result of the greater efficacy of modern clinical trial operations. That better and more sophisticated care is offered to patients with metastatic disease can be seen in the multidisciplinary recommendations of cancer societies and the services offered by cancer centers. Consistent with this, patients with mCRC are increasingly candidates for metastasectomy such that an increasing number of patients with oligometastatic disease are undergoing surgery with a lower mortality. A recent study looking at the 2002-2011 period found that mCRC was the most common indication for metastasectomy, with 87,407 cases identified. The average annual percentage change was also highest for mCRC, with the rate of metastasectomies increasing 6.83% per year. Liver and lung resections rose the fastest at respective rates of 10.19% and 9.43% per year. Furthermore, despite an increase in the number of comorbidities in patients undergoing metastasectomy, inpatient mortality rates after metastasectomy fell most significantly for mCRC, at a rate of 5.49% per year. In 2011 alone, 11,587 procedures were performed, suggesting that as many as 25% of patients with metastatic disease may be undergoing such surgeries.

As for the ability to better support patients at the end of life, one can look to the increased use of biliary stents and recent studies highlighting the effect of good supportive care for patients with CRC.

**Figure 4. Concordant Single-Study Gains in Progression-Free Survival (PFS) and Overall Survival (OS)**

Illustrated are gains in PFS (delta PFS), defined as PFS of the experimental arm minus PFS of the control arm, and gains in OS (delta OS), similarly defined. The slope of the line demonstrates that within a single study, the gains in OS are similar to the gains in PFS.
While we believe that the literature supports our approach,4,22-25 our present analysis has several limitations. First, we relied on the published literature regarding mCRC therapies. The published literature may not fully reflect the totality of clinical trials conducted; instead, trials with positive and statistically significant results are more likely to appear in the published literature and are published with shorter delay. However, as we selected large phase 2 and phase 3 studies, we believe that publication bias is likely to be smaller than that of smaller, uncontrolled trials. The publication of a large randomized study is provocative, regardless of result, and the conduct of such trials is widely known. Nevertheless, this is an assumption and should be treated as such.

In addition, our analytic method of plotting median OS and PFS by trial arm over time is a novel approach to delineate gains ascribed to treatment from secular gains. Some may dispute this method but we believe it is sound. Specifically, prior work from our group has established a robust

The upper panels compare the overall response rates in the experimental arms (left) and control arms (right) in studies conducted of first- and second-/third-line therapies. The response rates are significantly lower in second- and third-line regimens in both study arms. The middle pair of panels show the gains in progression-free survival as a function of overall response in second- and third-line regimens; the slopes (solid lines) (95% CIs [dotted lines]) are similar to those found in first-line regimens (Figure 3). The bottom panels graph the change in overall survival in studies of second- and third-line regimens over the past 2 decades, where such values have been reported; the negative slopes of the regression lines (solid lines) (95% CIs [dotted lines]) demonstrate the lack of meaningful improvement.
quantitative correlation between PFS and OS. In other words, across a large spectrum of trials of targeted and cytotoxic therapies, if a novel drug delayed progression 3 months, that drug nearly invariably improved survival by 3 months. Extreme deviations from this rule have not been replicated. It should be noted that it is possible that novel therapeutic strategies, such as vaccine or immune checkpoint inhibitors, may exert benefits long after therapy is complete, and the correlation we have reported may not apply. However, for the years and trials we considered, there were no examples of immunotherapy.

One might contend that as physicians gain experience in administering a therapy, improvements in progression grow larger and would be ascribed to control regimens by our analytic method. While we believe that this is a potential concern, we also believe that accounting for it would not alter our results greatly because most of the therapies examined were cytotoxic drugs, given a few days per cycle, with toxic effects readily encountered by physicians. In contrast to targeted agents with unique toxic properties, we believe that the learning curve was not as steep for these agents, and as such their efficacy was less likely to have improved dramatically with physician experience.

Finally some may argue that while indeed more metastasectomies are being performed, they have been made possible by better therapies and that this benefit should be ascribed to the therapies. However, the rapid increase in the number of metastasectomies in the last decade far exceeded the gains in ORR, making it unlikely that novel therapies could account for much of the increase in metastasectomies. For these reasons, and given the important fact that gains in PFS drive OS, we believe that our analysis is a logical extension of that fact. Nevertheless, it is possible that others may disagree with our methods.

Conclusions

To our knowledge, this is the first to attempt delineate survival gains as a result of novel pharmacologic advances from gains in supportive care, adjunct measures, and secular trends in medicine. We reach the conclusion that novel therapies only account for a portion of survival gains in mCRC, contrary to the widespread narrative that improved therapies are primarily responsible for advances. However, in our opinion, the most important conclusion to be drawn from this analysis is the indiscussable value of enrolling patients in clinical trials. The progress made in mCRC in the past 2 decades has been earned through the hard work of rigorous clinical trials conducted by countless dedicated investigators. And while patients with mCRC can now expect on average to live more than 2 years after the diagnosis, oncologists should not consider this acceptable. In mCRC as in other metastatic solid tumors, gains made have been driven primarily by gains in the initial therapy of metastatic disease, and in mCRC there is much room for improvement. For this reason, oncologists should not consider that clinical trials have established optimal first-line therapies and expect future gains to come from more effective second- and third-line therapies. Eligible patients should continue to be referred to and enrolled in clinical trials, especially those evaluating first-line therapies as well as the second- and third-line options that will inform the design of subsequent first-line therapies. We would argue that substantive gains in the survival of patients with mCRC will only be achieved with a radical change in our first-line paradigm and not by strategies that attempt to enhance existing first-line irinotecan- and oxaliplatin-based regimens. Only with aggressive referral to and enrollment in clinical trials will we see in the next decade a continuation of the progress of the past 2 decades.

REFERENCES


For decades, effective treatment for metastatic colorectal cancer (mCRC) was an almost impenetrable obstacle, with a single therapeutic option—fluorouracil. Outcomes were highly consistent across time and across trials, with a median overall survival (OS) of approximately 1 year. The last 20 years have seen highly clinically relevant improvements in not only OS but also progression-free survival (PFS) and response rates. In the most recently reported large first-line randomized trials, median OS has reached or exceeded 30 months.

Randomized trials, the approach through which we evaluate new therapies, have a fundamental, self-imposed limitation of addressing an isolated hypothesis. While trials test 1 element of a patient’s treatment, they are conducted in an environment that is in constant flux (thus the never-ceasing cautions regarding drawing conclusions from cross-trial comparisons). In mCRC, the last 20 years has brought advances in surgical techniques and approaches (increased use of metastasectomy), increased rates of screening (detecting disease earlier—not all stage IV disease is the same), improvements in imaging (earlier detection of recurrence after resection and adjuvant therapy), and various improvements in end-of-life care. These concurrent advances pose a considerable challenge when we try to identify the respective contributions of each to the greater than doubling in median survival that has been observed.

In this issue, Jawed et al attempt to estimate what fraction of the improvement observed in mCRC outcomes in the last 2 decades may be attributed to novel treatments. They performed a literature-based systematic review of 96 published phase 2 and 3 trials, comparing improvements in OS with improvements in PFS and response rate (among many other analyses). The approach hinges on a key assumption, which is that improvements in PFS, more or less, should translate directly and predictably into improvements in OS—we consider the veracity of this assumption below. Their primary conclusion, that chemotherapy has only contributed partly to the overall gains in survival, is highly plausible and is consistent with other findings from observational data.

That being said, as with any provocative analysis, a number of issues are present that may introduce bias into the findings. First, as recently shown by Shi et al, in modern mCRC trials, at the trial level, between-arm differences in PFS are not reliable predictors of differences in OS. The association between PFS and OS, as measured by the R coefficient, is almost identical (approximately 0.50) in the Shi et al and Jawed et al analyses. Interestingly, Jawed et al consider this correlation sufficient to justify an accurate prediction, but Shi et al conclude that the correlation is too low to allow PFS to be considered as a reliable surrogate end point. Second, subtleties in protocol-specific definitions of progression cannot be accounted for in a literature-based analysis, complicating interpretation. In addition, in practice (ie, outside of the trial setting), patients may be treated through RECIST progression with first-, second-, or later-line regimens, further challenging the relationship between RECIST-based PFS as measured in a study and the ultimate effect of an agent on OS.

Most critically, in their analysis, Jawed et al discount the possibility that a novel agent could improve OS in settings other than first-line therapies, stating that gains in second- and later-line outcomes have been unimpressive (based on a cross-trial comparison). Multiple real-world observations challenge this conclusion. Because many patients discontinue first-line treat-