PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Quality indicators of Clinical Cancer Care (QC_3) in colorectal cancer
AUTHORS	Bianchi, Valentina; Spitale, Alessandra; Ortelli, Laura; Mazzucchelli, Luca; Bordoni, Andrea

VERSION 1 - REVIEW

REVIEWER	Erin M Siegel, PhD MPH Assistant Member Cancer Epidemiology Moffitt Cancer Center USA I have no competing interests. I have published on the quality of care of colorectal cancer in the US.
REVIEW RETURNED	22-Mar-2013

THE STUDY	This is a methods manuscript which did not include patient recruitment or statistical analysis; therefore those review points are not applicable. There was not a CONSORT checklist provided.
REPORTING & ETHICS	There was not a CONSORT checklist provided; however, methods were sound.
GENERAL COMMENTS	This manuscript outlines the process of developing evidenced based quality of care indicators for colorectal cancer. The Investigators followed a comprehensive three step process for identifying, reviewing and finalizing a list of 27 indicators using both a working group of CRC practitioners and an advisory board. The development of these indicators utilized sound and meticulous methodology. To improve this manuscript, it would be helpful if the investigators could include a plan on how to measure these indicators (e.g. what are the specific criteria for each numerator) in general and more specifically at the population level. The benefits to the scientific literature include both a set of evidence based quality indicators for colorectal cancer as well as a consistent definition of how each quality indicator can be accurately measured. It is recommended that they provide feasibility data on measuring these indicators especially within the population cancer registry. Specific comments are listed below: 1) Did you include any experts in health services research or quality assessment on your boards? 2) Please provide more justification for the need to develop a new set of indicators for quality of colorectal cancer since validated quality indicators for colorectal cancer have been published. What unique aspects of cancer care in Switzerland justify the revised quality indicator list? 3) The quality indicators include very detailed clinical information. Are all the data elements needed to assess each quality measure included within cancer registry data? 4) Can this set of indicators be measured accurately and at the

	population level? The discussion references a retrospective pilot study. It is suggested that the details and results of this pilot be included in the manuscript in more detail (including the number of indicators that can be measured and those that cannot) to support the feasibility of assessing these quality indicators. 5) How does this final set of indicators compare to others quality indicators used nationally as well as internationally? Discussion regarding how your data will be comparable to any other quality efforts within your national health care setting as well as with European and American countries will be of great value. 6) The indicators listed in Table 1 are presented as proportions (e.g. numerator/denominator). Additional information on how these indicators will be measured, the criteria for inclusion in the numerator, and the cutpoint for what is deemed good quality of care should be explained in more detail. 7) Some indicators need clarification. For example, what would represent diagnostic quality, a higher proportion of patients diagnosed through screening, symptoms or accidental finding? How would you define the “definitive pathology report”? (e.g. must include all listed measure or a proportion?). 8) Indicators that involve the regimen or dosing of chemotherapy or radiation did not make your final list; however, these have been published before as validated measures. Can you discuss why they were excluded from your list? 9) It is widely accepted that overall survival and disease free survival are outcome measures of colorectal cancer. Can you provide more information on how you will use these outcomes to define quality of cancer care? Will they be related to other QI measures?

REVIEWER	Marcia M. Russell, MD Assistant Professor of Surgery David Geffen School of Medicine at UCLA Staff Surgeon - Colon & Rectal Surgery VA Greater Los Angeles Healthcare System United States of America No competing interests.
REVIEW RETURNED	28-Mar-2013

GENERAL COMMENTS	The authors seek to develop quality indicators for colorectal cancer using a 2-step modified Delphi process. This type of work to measure quality of care is important. The paper could benefit from the following revisions as outlined below: Introduction: 1) Multiple other groups have previously developed quality indicators for colorectal cancer – please expand on why you started from scratch for your project rather than adopted/modified available indicators. Are there issues unique to colorectal cancer care in Switzerland? Methods: 1) Please expand on what occurred at the in-person meeting held at the beginning of the process? Were the indicators announced there? Did this group have any role in selecting the 181 candidate indicators? 2) Please clarify as to whether the Delphi round 2 was in person or via mail. If by mail, please explain why you chose not to have an in-person meeting which is generally what is used for the RAND/UCLA Appropriateness Methodology.

	3) I am not sure exactly what is meant by the word 'megatrend' in the advisory board review. Please provide more details about how the indicators were decreased 89 to 27 during this portion of your project. Results: 1) Why was your response rate only 80% for the Delphi Round 1 and 2? Which panel members did not complete 1 or both rounds and how did you account for this in your results? 2) Please provide more details about how the 181 candidate indicators were narrowed down to a list of 149? 3) Please provide more details and transparency about how the 89 indicators rated as valid after the Delphi process were narrowed down to 27. Why did you have this last step – it is not part of the formal RAND/UCLA Appropriateness Methodology. What were reasons for not including some of the indicators rated as valid? Discussion: 1) I don't think most readers will know what 'anamnesis' is. 2) Can paragraph 2 be better represented as a table with each of the main steps of the diagnostic-therapeutic process as a column? 3) I am concerned about the lack of indicators in the follow-up domain? If we are going to provide high quality colorectal cancer care aren't we obligated to address issues of surveillance and survivorship? Please expand on the decision not to have any follow-up indicators in the final group – were there any in the 89 indicators rated as valid after the Delphi process? 4) Please discuss limitations and next steps. 5) Please discuss how your work expands on previous quality indicator development in colorectal cancer.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

Erin M Siegel, PhD MPH
Assistant Member
Cancer Epidemiology
Moffitt Cancer Center
USA

I have no competing interests. I have published on the quality of care of colorectal cancer in the US.

This is a methods manuscript which did not include patient recruitment or statistical analysis; therefore those review points are not applicable. There was not a CONSORT checklist provided.

There was not a CONSORT checklist provided; however, methods were sound.

This manuscript outlines the process of developing evidenced based quality of care indicators for colorectal cancer. The Investigators followed a comprehensive three step process for identifying, reviewing and finalizing a list of 27 indicators using both a working group of CRC practitioners and an advisory board. The development of these indicators utilized sound and meticulous methodology. To improve this manuscript, it would be helpful if the investigators could include a plan on how to measure these indicators (e.g. what are the specific criteria for each numerator) in general and more specifically at the population level. The benefits to the scientific literature include both a set of

evidence based quality indicators for colorectal cancer as well as a consistent definition of how each quality indicator can be accurately measured. It is recommended that they provide feasibility data on measuring these indicators especially within the population cancer registry. Specific comments are listed below:

1) Did you include any experts in health services research or quality assessment on your boards?

Among the Panellists of the Advisory Board, we included Prof. J. Faivre, Director or the Registre Bourguignon des Cancers Digestifs, Dijon, France. He is an expert epidemiologist in colorectal cancers involved in several projects concerning pattern of care, patient outcome and quality assurance / assessment of colorectal cancers. Furthermore, Prof. Ph. Quirke collaborated is part of the Advisory Board; he has a broad knowledge and experience in the quality care assessment. In the quality indicators selection process, both of them provided a careful vision, with comments and suggestions, concerning each quality indicator in terms of the feasibility of data collection at the population-based level.

2) Please provide more justification for the need to develop a new set of indicators for quality of colorectal cancer since validated quality indicators for colorectal cancer have been published. What unique aspects of cancer care in Switzerland justify the revised quality indicator list?

Quality indicators (QI) are generally produced to evaluate specialised cancer care setting with the consequence that are less applicable at the population-based and cancer registration level, which represent and observe a whole geographical community. The proposed list of QI includes: a) already known QI, retrieved from the available literature and list of QI already approved in the past, which have been re-evaluated by the Working Group and the Advisory Board with an “eye” on the population-based data collection and evaluation; b) new QI mostly proposed by the Working Group. We believe that any additional Delphi process can increase our knowledge on each QI, through the update of some specific QI, the revision of other QI in terms of definition and reference denominator. Moreover, the Delphi process can per se generally increase the acceptance and participation of clinical practitioners to the study. Thus, the proposed set of QI should be considered more as a promotion for population-based evaluation than as a pure regional Swiss initiative. These considerations have been introduced in the Introduction, Methods, Results and Discussion sections of the manuscript.

3) The quality indicators include very detailed clinical information. Are all the data elements needed to assess each quality measure included within cancer registry data?

Cancer registration is developing more and more in Europe and worldwide, reaching “high level” of registration, where detailed information of the diagnostic and treatment procedure are routinely collected. In fact, some cancer registries (including Ticino Cancer Registry, South of Switzerland) have access to the complete pathological, radiological and clinical reports with the aim to increase the quality of cancer registration (with detailed and precise data collection) and to allow the analysis of such quality indicators.
We, therefore, added some details for each quality indicators both in the Results section of the manuscript and in Table 1.

4) Can this set of indicators be measured accurately and at the population level? The discussion references a retrospective pilot study. It is suggested that the details and results of this pilot be included in the manuscript in more detail (including the number of indicators that can be measured and those that cannot) to support the feasibility of assessing these quality indicators.

The Delphi process was performed to warrant the accurate measurement of quality indicators at the

population-based level. We therefore completely share the Reviewer's suggestion. In the present study we reported the list of those quality indicators which have been considered of scientific and clinical value by a sufficient number of experts (≥70% of the agreement), when non doubtful comments concerning the data collection at the population-based level was mentioned. Therefore, those quality indicators evaluated at least by one of the experts "not completely feasible and difficult to be collected at the population-based level", were definitely excluded.
We aknowledge that this point was not enough clear in the manuscript; therefore, we added a sentence in the Method and Discussion sections of the manuscript.

Some pilot results have been presented to different international congresses (such as the annual meetings of the International Association of Cancer Registries, the European Network of Cancer Registries, the Group of Registry and Epidemiology of Cancer in Latin Speaking Countries). We added the related citations in the Discussion of the manuscript. Our intent for the present study was to focus the readers' attention on the indicator production (numerator, denominator, rationale, medical documents necessary for the data collection) in a population-based setting, avoiding a possible dispersion of the attention to some results, that would have only a preliminary nature. In fact, we plan to submit results for further publications, with the chance to provide in depth comments and international comparisons for each quality indicator.

5) How does this final set of indicators compare to others quality indicators used nationally as well as internationally? Discussion regarding how your data will be comparable to any other quality efforts within your national health care setting as well as with European and American countries will be of great value.

Please, see considerations reported in the answer to question 2.

6) The indicators listed in Table 1 are presented as proportions (e.g. numerator/denominator). Additional information on how these indicators will be measured, the criteria for inclusion in the numerator, and the cutpoint for what is deemed good quality of care should be explained in more detail.

We fully agree with the Reviewer. In order to better show how the indicators should be calculated, we carefully revised Table 1, adding a column with the numerator, one with the rationale and one concerning the necessary medical documentation for the extrapolation of data needed for the indicators measurement.

The concept of standards of care for each quality indicator, in terms of minimum and target requirements (="cutpoint"), is mentioned in the Introduction section of the manuscript as "second aim of the overall QC3 project", but it does not represent the primary goal of the study. In fact, the intent for the present study was to focus the readers' attention on the indicator production (numerator, denominator, rational, medical documents necessary for the data collection) in a population-based setting. Cutpoints will be surely developed in the second phase of the project on the basis of definitive results discussed with the Working Group and compared with the available literature.

We thank the Reviewer for his note, since we realized that cutpoints were erroneously mentioned in the Abstract Summary / Article focus (second point). We have now modified it accordingly.

7) Some indicators need clarification. For example, what would represent diagnostic quality, a higher proportion of patients diagnosed through screening, symptoms or accidental finding? How would you define the "definitive pathology report"?

The first indicator of the “diagnosis” group is important to understand what is happening in a territory where there is not an organized screening programme for colorectal cancers, but only an opportunistic screening strategy. If the tumour is detected because the physician submit the patients older than 50 years old to a colonoscopy control or if a patient, being aware of the possible risk, ask his family doctor to undergo a faecal occult blood test (FOBT) or colonoscopy, is an interesting data to be evaluated, also in the hypothesis of colorectal cancer screening programme implementation. We, therefore, believe that a higher proportion of patients diagnosed through screening (FOBT and/or colonoscopy in asymptomatic patients) would represent a higher diagnostic quality, since the therapeutic approach and, consequently, the patients outcome (in terms of recurrence and survival) would be more favourable, as reported in the literature (Siegel EM et al. J Oncol Pract. 2012 Jul;8(4):239-45; Schoen RE et al. NEJM 2012:366:2345-57; Wilkins T et al. Am Fam Physician 2008;78(12):1385-92; Sikka V, Ornato JP. Am J Emerg Med 2011; Levin B et al. Gastroenterology 2008;134(5):1570-95; Winawer S et al. Gastroenterology 2003;124(2):544-60; Majumdar SR et al. Am J Gastroenterol 1999;94(10):3039-45)

In order to clarify this concept to the reader, we added a sentence in the Discussion of the manuscript, according to the Reviewer’s suggestion.

The “definitive pathological report” (now renamed as “surgical pathology report”) derives from the surgical curative intervention and it should be as complete as possible to be useful for future decisions about patient’s treatment. Our intent is to calculate this indicator for all listed items considered together, but also for each item individually analyzed: e.g. proportion of patients with colorectal cancer and a surgical pathology report including the surgical intervention description; proportion of patients with colorectal cancer and a surgical pathology report including the tumour size; proportion of patients with colorectal cancer and a surgical pathology report including assessment of the resection margins; proportion of patients with colorectal cancer and a surgical pathology report including the pathological staging (AJCC pTNM); etc…
We therefore modified the Table 1 and added a sentence in the Discussion of the manuscript accordingly.

8) Indicators that involve the regimen or dosing of chemotherapy or radiation did not make your final list; however, these have been published before as validated measures. Can you discuss why they were excluded from your list?

Initially, these quality indicators (QI) have been selected by the colorectal cancer Working Group, but actually they haven’t passed the Advisory Board’s selection, because there was not the consensus about them; the discussants’ point was that some regimens can differ among countries. For instance, the following QI concerning RT regimen were excluded: “Proportion of patients with rectal cancer undergoing neo-adjuvant RT with a total dose of 25 Gy and a hypofractionation regimen of 5 Gy daily and five fractions weekly (short-course RT schedule)” and “Proportion of patients with rectal cancer undergoing neo-adjuvant RT with a total dose of 44-50.4 Gy and a standard fractionation of 1.8–2.0 Gy daily and five fractions weekly (long-course RT schedule)”. The QI analyzing the chemotherapy regimens more in details have not passed the selection of both the colorectal cancer Working Group and the Advisory Board, due to the fact that, being a prospective study, the chemotherapy regimens could change in itinere.

9) It is widely accepted that overall survival and disease free survival are outcome measures of colorectal cancer. Can you provide more information on how you will use these outcomes to define quality of cancer care? Will they be related to other QI measures?

We fully agree with the Reviewer's suggestion. Overall and disease-free survival, for which it is necessary to wait for a certain follow-up period (i.e. 1, 3, 5 to 10 years from diagnosis for the calculation of overall survival, and from curative treatment for the calculation of disease-free survival) will represent the overall resume of the diagnostic and treatment quality of colorectal cancer patients. We plan to analyse survival and disease-free survival according to some of the proposed QI (such as QI concerning the pathological characteristics of the tumours, QI of the adjuvant chemotherapy in patients with colon cancer and AJCC TNM stage II high-risk or III, QI of colorectal patients operated on with free margins, QI of locally advanced rectal cancer patients undergoing neo-adjuvant radio ± chemotherapy, etc…). We will finally compare our results with other national and international studies (particularly population-based ones), favouring the interpretation of each single QI.
We added a sentence in the Discussion of the manuscript to give more emphasis to this important issue.

****

Reviewer 2:

Marcia M. Russell, MD
Assistant Professor of Surgery
David Geffen School of Medicine at UCLA
Staff Surgeon - Colon & Rectal Surgery
VA Greater Los Angeles Healthcare System
United States of America

No competing interests.

The authors seek to develop quality indicators for colorectal cancer using a 2-step modified Delphi process. This type of work to measure quality of care is important. The paper could benefit from the following revisions as outlined below:

Introduction:

1) Multiple other groups have previously developed quality indicators for colorectal cancer – please expand on why you started from scratch for your project rather than adopted/modified available indicators. Are there issues unique to colorectal cancer care in Switzerland?

Please refer to the answer to question 2 of Reviewer 1.

Methods:

1) Please expand on what occurred at the in-person meeting held at the beginning of the process? Were the indicators announced there? Did this group have any role in selecting the 181 candidate indicators?

"181" was the number of the selected literature (guidelines, peer-review articles), used to create the list of 149 quality indicators (QI). These QI were then presented in the in-person meeting of the colorectal cancer Working Group, who had a defined time period to review the literature, add additional QI or to propose modification / elimination of QI not considered suitable for the study. Therefore, the role of the panellists was essential in the definition / approbation of the QI list before

the beginning of the Delphi process. In fact, the initial list of 149 QI was shortened to 104; and these 104 were the “content” of the Delphi process validation (please see Figure 1).

2) Please clarify as to whether the Delphi round 2 was in person or via mail. If by mail, please explain why you chose not to have an in-person meeting which is generally what is used for the RAND/UCLA Appropriateness Methodology.

The modified Delphi technique enables a large group of experts to be contacted cheaply, usually by mail with a self-administered questionnaire and sometimes it can include a round in which the participants meet to discuss the process and resolve doubts or ambiguities in the questionnaire. Since the Delphi process was performed anonymously, we sent the questionnaire by mail (Jones J, Hunter D. BMJ. 1995 Aug 5;311(7001):376-80. Review. Gagliardi AR et al. Can J Surg. 2005 Dec;48(6):441-52). After the 2nd round of Delphi we sent the final results of Delphi to the WG by e-mail asking for a general consensus, which was more practical if compared to a further in-person meeting.
We had, however, a face-to-fate meeting with the experts of the colorectal Working Group before the beginning of the 2-steps Delphi process in order to solve in advance possible problems and ambiguities concerning the whole technique.

3) I am not sure exactly what is meant by the word ‘megatrend’ in the advisory board review. Please provide more details about how the indicators were decreased 89 to 27 during this portion of your project.

The word “megatrends” refers to one of the ranking methods which can be used in a survey to figure out a personal opinion of the answering participant cohort (person could agree or not to a specific topic) (Blind K et al. Technological Forecasting & Social Change 2001;68:131-49). We have asked the Advisory Board’s panellists to express their acceptance about the Working Group’s selected list by a vote “YES” or “NO” and to express their opinion and comments about their choice. This step was conducted by e-mail, because having a wide panel composed by experts in almost every specialty concerning colorectal cancers, we want to know to whom the comments and the preference belong, so to correct and eliminate the QI which were not suitable for the final list.

The reasons why the number of QI decreased from 89 to 27 are essentially the following:

1. The number of QI identified by the QC3 colorectal cancers Working Group and then submitted to the Advisory Board was 89 because the 3rd pathology indicator entitled “Proportion of patients with colorectal cancer and a surgical pathology report including the following characteristics:…” was evaluated for every single item composing it, for a total of 14 single indicators. The intent was to have a pathology indicator which would have encompassed all the necessary aspects of a high quality surgical pathology report (please see also the answer to the question 7 of the Reviewer 1). If we split it in many indicators we will finally have a total of 41 QI, which would be concordant with the 89 QI after the Working Group’s revision. Practically, we have reported it as a whole QI in the manuscript, Table1 and Figure1, but our intent is to calculate this indicator for all listed items considered together, but also for each item individually analyzed. We have added a sentence in the Discussion section of the manuscript concerning the calculation of this QI.

2. the Delphi process involving the colorectal Working Group and the megatrends of the Advisory Board were performed in order to warrant the accurate measurement of quality indicators at the population-based level. In the present study we reported the list of those quality indicators which have been considered of scientific and clinical value by a sufficient number of experts (≥70% of the agreement), when non doubtful comments concerning the data collection at the population-based

level was mentioned. Therefore, those quality indicators evaluated at least by one of the experts “not completely feasible and difficult to be collected at the population-based level”, were definitely excluded. In order to better clarify this concept to the reader, we added a sentence in the method section of the manuscript.

Results:

1) Why was your response rate only 80% for the Delphi Round 1 and 2? Which panel members did not complete 1 or both rounds and how did you account for this in your results?

Being the Delphi process anonymous, we cannot know who did not answer to the questionnaire. The data analysis has been done on the available data (available case analysis).

2) Please provide more details about how the 181 candidate indicators were narrowed down to a list of 149?

Maybe there was a misunderstanding: 181 was the number of the selected literature (guidelines, peer-review articles), used to create the list of 149 quality indicators (QI). Please see the answer to question 1 of the Methods section.

3) Please provide more details and transparency about how the 89 indicators rated as valid after the Delphi process were narrowed down to 27. Why did you have this last step – it is not part of the formal RAND/UCLA Appropriateness Methodology. What were reasons for not including some of the indicators rated as valid?

Please consider the answer to question 3 of the Methods section.

Discussion:

1) I don’t think most readers will know what ‘anamnesis’ is.

We have changed the word “anamnesis” with “clinical history”.

2) Can paragraph 2 be better represented as a table with each of the main steps of the diagnostic-therapeutic process as a column?

We fully agree that the diagnostic-therapeutic process should be as clear as possible. In this sense, Table 1 represents the whole process divided in the following chapters: diagnosis, pathology, treatment (including surgery, chemotherapy and radiotherapy) and outcome. Each quality indicator was included in the corresponding chapter. Table 1 was enriched including a column with the description of the diagnostic-therapeutic process (named “clinical domain”) and other essential information, such as the numerator, the rationale and the medical documentation necessary for the data collection of each quality indicator at the population-based level.

3) I am concerned about the lack of indicators in the follow-up domain? If we are going to provide high

quality colorectal cancer care aren't we obligated to address issues of surveillance and survivorship? Please expand on the decision not to have any follow-up indicators in the final group – were there any in the 89 indicators rated as valid after the Delphi process?

Unfortunately, the follow-up quality indicators (QI), which were initially 6, were not finally approved due to the fact that the surveillance of colorectal cancers is not yet standardized; this was the most common reason for exclusion given by the involved experts. This concept is reported in the Discussion section of the manuscript:
"Concerning the QI about follow-up, AB did not finally include any of them. Indeed, although the follow-up procedures are suggested by several international guidelines, they are based on level II-III evidence and controversies remain regarding selection of optimal strategies for following up patients after potentially curative CRC surgery"

4) Please discuss limitations and next steps.

Considering limitations of the study, we have listed in the Discussion section of the manuscript some critical aspects; in particular:

- the level of evidence found in the literature: for some indicators, strong evidence of their validity was not available from randomized clinical trials (RCT); but this limit was solved using the expert panel methodology for the selection of indicators
- in the literature selection, we could have missed some relevant articles; however, members of the Working Group were likely to be very familiar with the literature, and had the opportunity to suggest other indicators based on their experience and literature search; in this way, we believe to have integrated the best research evidence with clinical expertise, as reported by Sackett et al.
- a further limit could be the feasibility of measuring indicators in terms of data collection and calculation. However, both the Working Group and the Advisory Board were concerned about the feasibility, validity and reliability of clinical data collection, necessary for the calculation of each single indicator at the population-based level. In fact, in order to warrant an accurate measurement, those indicators reaching more than 70% of the agreement, confirming their scientific and clinical value, but evaluated at least by one of the experts not feasible and difficult to be collected at the population-based level, were definitely excluded. In this way, we have overcome the feasibility limit.

Next step will be to produce definitive results on the selected quality indicators. Our intent for the present study was to focus the readers' attention on the indicator production (numerator, denominator, rational, medical documents necessary for the data collection) in a population-based setting, avoiding a possible dispersion of the attention to some results, that would have only a preliminary nature. In fact, we plan to submit results for further publications, with the chance to provide in depth comments and international comparisons for each quality indicator.

5) Please discuss how your work expands on previous quality indicator development in colorectal cancer.

Please refer to the answer to question 2 of Reviewer 1.

VERSION 2 – REVIEW

REVIEWER	Marcia M. Russell, MD Assistant Professor, David Geffen School of Medicine at UCLA Staff Surgeon, VA Greater Los Angeles Healthcare System
REVIEW RETURNED	12-Jun-2013

GENERAL COMMENTS	I thought the revised manuscript was much improved and the authors adequately responded to the comments from the reviewers.