Human Pathology (2019) **84**, 26–34

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Histopathologic tumor regression grading in patients with gastric carcinoma submitted to neoadjuvant treatment: results of a Delphi survey ☆

Andrianos Tsekrekos MD [a,b,*], Sönke Detlefsen MD, PhD [c], Robert Riddell MD, PhD [d], James Conner MD, PhD [d], Luca Mastracci MD, PhD [e], Kieran Sheahan MB, FRCPath [f], Jayant Shetye MD, PhD [g], Lars Lundell MD, PhD [b], Michael Vieth MD, PhD [h]

[a]*Department of Upper Abdominal Surgery, Center for Digestive Diseases, Karolinska University Hospital, 141 57 Huddinge, Stockholm, Sweden*
[b]*Division of Surgery, Department of Clinical Science, Intervention and Technology (CLINTEC), Karolinska Institutet, 141 57 Huddinge, Stockholm, Sweden*
[c]*Department of Pathology, Odense University Hospital, 5000 Odense C, Denmark*
[d]*Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Joseph and Wolf Lebovic Health Centre, Toronto, ON M5G 1X5, Canada*
[e]*Division of Anatomic Pathology, Department of Surgical Science and Integrated Diagnostics (DISC), University of Genoa, 16126 Genoa, Italy*
[f]*Department of Pathology, St Vincent's University Hospital & UCD School of Medicine, Dublin, D04 T6F4, Ireland*
[g]*Department of Pathology, Karolinska University Hospital, 141 57 Huddinge, Stockholm, Sweden*
[h]*Institute of Pathology, Klinikum Bayreuth, 95445 Bayreuth, Germany*

Received 27 May 2018; revised 16 August 2018; accepted 21 August 2018

Keywords:
Gastric cancer;
Neoadjuvant treatment;
Histopathologic tumor regression grade;
Delphi;
Consensus

Summary Studies investigating the histopathologic response of gastric carcinoma to neoadjuvant treatment have used a variety of different tumor regression grading systems. The aim of this Delphi survey was to review the available systems and reach consensus on a potential international standard. An international e-mail–based Delphi survey involving 6 expert pathologists was undertaken between January and October 2017. A questionnaire consisting of 72 items was formed after reviewing the 5 available systems. Rating of the items was done on a symmetric 4-point Likert-type scale, and feedback was provided between rounds. A total of 4 rounds were required to reach consensus on 97% of the items covering the topics: (1) specimen processing, (2) gross examination, (3) cross sectioning/method of sampling, (4) staining, (5) immunohistochemistry, (6) assessment of tumor regression in response to neoadjuvant therapy, (7) tumor regression grading, (8) assessment of regression of nodal metastases, and (9) role of histologic tumor type. Through the outcome of this comprehensive Delphi study, a group of experts is proposing a 4-tiered system for the grading of regression of the primary tumor, combined with a 3-tiered system for lymph node metastases. Grade 1 represents complete response, grade 2 contains less than 10% residual tumor (subtotal regression),

☆ Disclosures: The authors declare no conflict of interest. No financial support was received from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author at: Department of Upper Abdominal Surgery, Center for Digestive Diseases Karolinska University Hospital, Hälsovägen 13, 141 57 Huddinge, Stockholm, Sweden.
E-mail address: andrianos.tsekrekos@sll.se (A. Tsekrekos).

https://doi.org/10.1016/j.humpath.2018.08.028

grade 3 contains 10% to 50% residual tumor (partial regression), and grade 4 contains greater than 50% residual tumor (minimal/no regression). The addition of "a", "b", or "c" indicates complete, partial, or no response of lymph node metastases. It is recommended to use this grading system irrespective of histologic subtype.

1. Introduction

Although the incidence of gastric cancer is steadily declining, it still represents the fifth most common malignancy and the third leading cause of cancer-related death worldwide [1]. In Western countries, the prognosis of gastric cancer remains poor; 5-year survival is reported to be in the range of 25% [2], suggesting that nondetectable metastatic disease—either regional or distant—is already present at the time of surgery in a significant proportion of patients intended to have a potentially curative resection. Given that, a variety of avenues have been pursued in attempts to improve survival rates, mainly by offering systemic treatment in the form of perioperative chemotherapy.

The use of neoadjuvant chemotherapy for gastric cancer was first reported by Wilke et al [3] in the context of advanced, nonresectable tumors. Second-look surgery occasionally revealed a chemotherapy-induced "downstaging," or even complete disappearance of all macroscopic disease, making resection possible. Following these landmark observations, several studies were conducted examining the value of perioperative chemotherapy given to patients with locally advanced but resectable tumors, and this strategy seemed to offer survival benefit as demonstrated in 3 phase III randomized controlled trials [4-6]. These findings translated into a widespread use of perioperative chemotherapy, which has become the standard option in Europe [7].

In the ongoing search for more effective multimodal therapeutic protocols in advanced gastric cancer, several phase II trials have been initiated, where one issue is which relevant surrogate markers can predict a significant clinical effect; one is the grade of tumor response to neoadjuvant therapy, where 3 evaluation criteria are currently available. The Response Evaluation Criteria in Solid Tumors (RECIST) is considered the gold standard [8]. However, RECIST requires the presence of a measurable lesion and, given the fact that resectable gastric carcinoma seldom has measurable lesions, its use is not a valid option for clinical research in the neoadjuvant setting. The Japanese Classification of Gastric Carcinoma advocates a response evaluation criterion involving barium x-ray or endoscopic examination for the assessment of such tumors without measurable lesions, but until now, this method has not been met by a widespread acceptance [9]. Finally, tumor regression grade (TRG) can be evaluated histologically in the resected specimen.

In 2013, Kurokawa and colleagues [10] conducted a correlative study to determine the best surrogate end point variable for overall survival in neoadjuvant studies for gastric cancer, stimulated by the finding that in esophageal cancer the histopathologic response rate was accurate and superior to RECIST. The results suggested that histopathologic response is indeed a valid surrogate end point variable relevant for survival outcomes and harbors the potential to be used in neoadjuvant trials [11]. Recent studies have provided evidence to support this, confirming the prognostic value of histopathologic tumor regression in gastroesophageal cancer and the correlation of response with a favorable outcome [12,13]. Such a conclusion has recently been amplified by the results of the FLOT4-AIO trial, which compared 2 different perioperative chemotherapy regimens in resectable gastric and gastroesophageal junction cancer; the primary end point of the phase II part of the trial was the proportion of patients with pathological complete regression, and the analysis revealed a 10% difference in the rate of pathological complete regression between the groups [14]. Subsequently, this was reflected in the results of the phase III part of the trial, demonstrating a corresponding difference in overall survival [6].

In studies investigating the histopathologic response to neoadjuvant therapy in gastric cancer, several different TRG systems have been used [9,15-18]. The common feature of these systems is that they are based on the estimated percentage of either residual viable tumor tissue in relation to the initial tumor bed or the amount of therapy-induced fibrosis in relation to residual tumor. This diversity of the grading systems encountered in the literature complicates the interpretation and generalizability of research data, making cross-trial comparisons redundant. The aim of this Delphi survey was to comprehensively review these TRG systems with the objective to either reach consensus on which one has the potential to become the international standard or alternatively launch a novel system for this purpose.

2. Materials and methods

2.1. Identification and selection of participants—panel size

Candidates were invited because of their special interest, and scientific and clinical competence in the field of gastrointestinal pathology, whereupon 6 experts finally agreed to participate. This approach did not allow for true anonymity, as the participants were known to the researchers and partially even to one another. Nevertheless, the respondents' judgments

Table 1 Outcomes of the survey and recommendations regarding the specimen workup (specimen processing, gross examination, cross sectioning/method of sampling, staining, and immunohistochemistry)

Topic	Recommendations
Specimen processing	• The specimen has to be sent to the pathology department fresh and intact, optimally in a vacuum-assisted, temperature-controlled system. If such a system is not available, conventional sending procedures should be applied, with cold storage at 4°C. • In case of longer distances of transportation, fixation in neutral buffered formalin is mandated.
Gross examination	• After documentation of any serosal tumor involvement, the stomach should be opened by dividing the wall opposite to the tumor (the opening of the specimen needs to be case dependent). • In cases of circumferential tumor growth or no macroscopic residual tumor, the specimen should be opened along the greater curvature. • Photographic documentation of the specimen should be done routinely (ideally fresh). • It is of value to divide the specimens into 3 macroscopic groups, as described by Mandard et al [19]: 1. Obvious residual tumor with ulcerating, fungating, or infiltrative features 2. Apparent tumor regression has occurred, a scar is found instead 3. Doubtful cases with partial response Depending on the gross aspect of the residual tumor, a different sampling method can be considered (see "cross sectioning/method of sampling"). • The adipose tissue located underneath the residual neoplasm should not be dissected to search for lymph nodes before pinning but after fixation and tumor sectioning.
Cross sectioning/method of sampling	• The specimen should be fixed in 10% buffered formaldehyde for >24-48 h before sampling. • In case of "fresh" sampling for biobanking or when fresh tissue for molecular analysis is needed, tissue should be frozen (ideally in OCT media) until final histologic examination of the specimen is reported, in case it needs to be revaluated. • The fixation period should not exceed 72 h because this creates excessive formalin-mediated cross linking, leading to potential problems with immunohistochemistry and other molecular techniques. • Representative sections should include the region of the tumor, surgical margins, and uninvolved portions of the stomach as a reference to assess background mucosal changes/pathology on which the carcinoma arose (adequate sections of antrum, corpus, and fundus). • Cross sections should be captured serially at 3- to 5-mm intervals from the following: 1. The entire macroscopically identifiable tumor 2. Areas with abnormalities (ie, increased consistency or thickness of the wall, ulceration) 3. The area of the stomach with scarring, indicating the site of the previous tumor • Sampling should be extended to include 2 cm of the adjacent tissue around the tumor. The same sampling strategy should be applied in the situation with no macroscopically identifiable tumor because subjecting only ulceration/scar tissue for histology may underestimate tumor response. • In the case of no gross identifiable tumor bed, correlation with pre–therapy endoscopy, previous histology reports, or imaging findings should be considered. • In large tumors that macroscopically seem not to have responded to therapy, it may be sufficient to only submit every second slice for histologic examination.
Staining	• Initial assessment by H&E only is sufficient. • The use of additional stains depends on what is visible on the H&E. • Alcian blue–PAS can be used to assist in distinguishing signet ring cells (or other more immature subtypes) from histiocytes. However, the use of immunohistochemistry for broad-spectrum cytokeratins is recommended for this purpose owing to its superior specificity.
IHC	• There is no need for the upfront use of IHC in the routine assessment of TRG, but it can be added depending on what is visible on the standard stains. • The supplementary use of IHC depends more on the degree of tumor response to therapy rather than its histologic type. • IHC techniques to determine the expression of broad-spectrum cytokeratin-positive cells carry a potential to improve the identification of viable tumor cells within large areas of fibrosis. This may be useful not only in cases of "no residual carcinoma" diagnosed on the H&E staining but also in cases with up to 10% residual tumor. Above all, it may be useful in diffusely infiltrating, poorly cohesive/signet ring carcinoma, where tumor cells can be difficult to recognize (and quantify) on a background of post–chemotherapy inflammation with macrophages and fibrosis.

Abbreviations: H&E, hematoxylin and eosin; IHC, immunohistochemistry; OCT, optimal cutting temperature; PAS, periodic acid–Schiff.

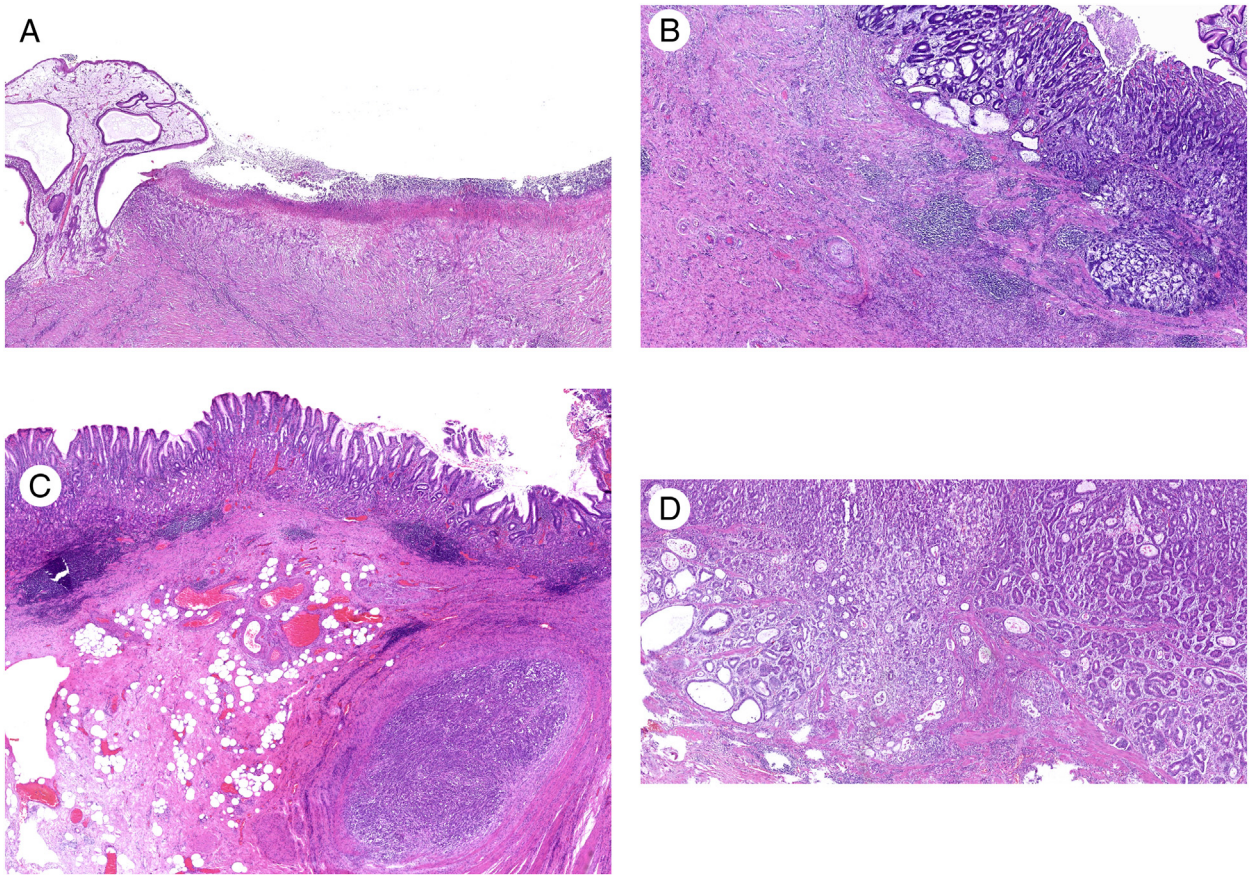

Fig. 1 Representative histologic images of the 4 grades of tumor regression — primary tumor (hematoxylin-eosin staining). A, Primary tumor bed showing no vital tumor cells after chemotherapy. Instead, ulceration, granulation tissue, fibrosis and reactive changes of the nonneoplastic mucosa (at the upper left-hand side of the image) are observed — regression grade 1 (original magnification 25x). B, G2 adenocarcinoma with less than 10% vital tumor cells and around 10% mucinous component. The regressive changes consist of a marked fibrosis with inflammation — regression grade 2 (original magnification 25x). C, G3 adenocarcinoma showing fibrosis after chemotherapy around the remains of the primary tumor — regression grade 3 (original magnification 10x). D, G2 adenocarcinoma showing no regression after chemotherapy — regression grade 4 (original magnification 30x).

remained strictly anonymous to the expert panel throughout the survey. Correspondence was e-mail based, and we obtained a response rate of 100% in all rounds.

2.2. Brief description of the Delphi survey process

The Delphi survey was based on a process consisting of a total of 4 rounds, with interspersed controlled feedback. A search of the relevant literature had identified 5 different systems for the histopathologic assessment and classification of tumor response in gastric cancer [9,15-18]. For the first round, we proceeded with the development of an evaluation form for each of the systems, presenting their respective characteristics. We also added the 2 most commonly used systems for esophageal and gastroesophageal junction cancers [19,20], in order for the participants to reflect on their potential application in gastric cancer as well (Supplementary Appendix A). The experts were asked to evaluate the different elements of each individual system and express their opinion on their pros and cons. In addition, there was ample space for comments and proposals, with the purpose to identify issues to be addressed in subsequent rounds. Based on the collected responses, a questionnaire consisting of 72 statements was formed, covering the following topics: (1) specimen processing, (2) gross examination, (3) cross sectioning/method of sampling, (4) staining, (5) immunohistochemistry, (6) assessment of tumor regression in response to neoadjuvant therapy, (7) tumor regression grading, (8) assessment of regression of nodal metastases, and (9) role of histologic tumor type (Supplementary Appendix B). In round 2, the participants were asked to give their opinion by rating each of the 72 statements on a symmetric 4-point Likert-type scale ranging from 1 (completely disagree) to 4 (completely agree). Consensus was achieved in 53 of 72 items (74%), whereupon a new questionnaire was constructed for round 3, consisting of the items where consensus had not been reached and incorporating feedback in terms of the results of the previous round (Supplementary Appendix C). In this next round, the statements were either slightly rephrased to integrate participants' proposals or remained unchanged, and the panelists were provided with the other experts' original

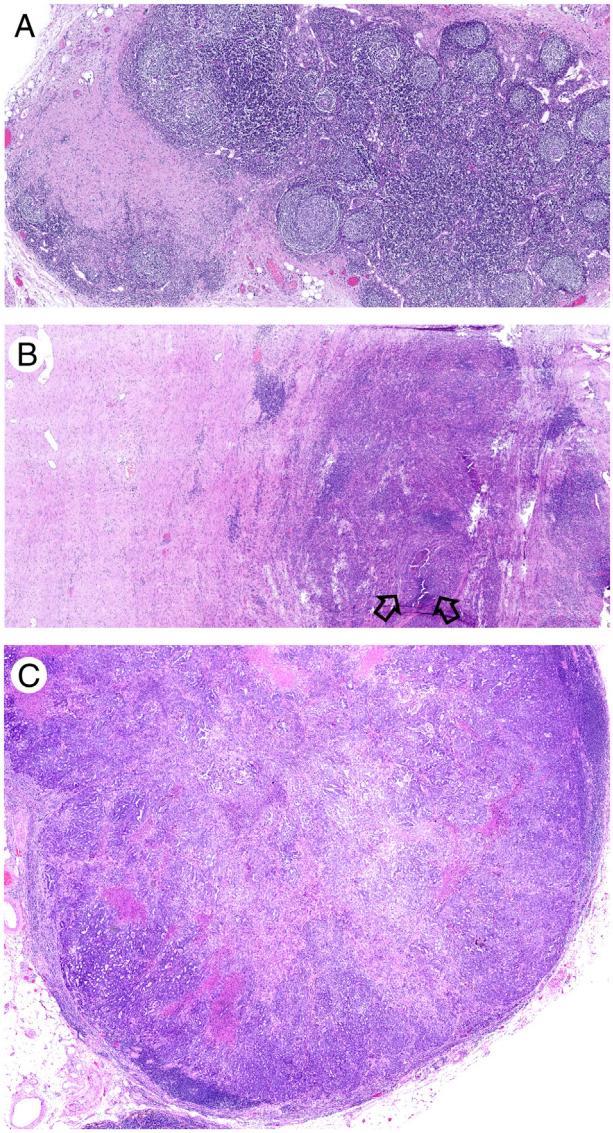

Fig. 2 Representative histologic images of the 3 grades of tumor regression — metastatic lymph nodes (hematoxylin-eosin staining, original magnification 25x). A, Lymph node from a patient with gastric G2 adenocarcinoma, where a former metastasis in its entirety has been replaced by regressive fibrosis after chemotherapy with no viable tumor cells left — regression grade a. B, G2 adenocarcinoma lymph node metastasis showing large amounts of regressive fibrosis after chemotherapy. Only small foci of metastatic adenocarcinoma remain (arrows) — regression grade b. C, G2 adenocarcinoma metastasis to a lymph node showing no regression after chemotherapy — regression grade c (nonresponder).

remarks instead. In a number of items, the choice was restricted to a "yes" or "no" answer. Finally, in round 4, one single question was distributed, with the sole objective to consent on the exact form of the TRG system to be recommended (Supplementary Appendix D). After obtaining consensus in 97% of the statements and with a majority of 4 participants advocating the same system, we decided to terminate the survey and summarize the results. A more detailed description of the survey process is provided in Supplementary Appendix E.

2.3. Data analysis, statistical interpretation, and definition of consensus

To date, no universally accepted criterion exists when it comes to determining consensus on Delphi surveys, and a variety of statistical analysis techniques for interpreting the data are encountered in the literature. We therefore consulted a professional statistician who assisted us in defining consensus as follows: based on the participants' scoring on the previously described 4-point scale, a median value was calculated and consensus was considered to exist if the median was either 3.0 or greater (corresponding to completely or partly agreement from the majority) or 2.0 or less (corresponding to completely or partly disagreement from the majority). On the contrary, consensus was not reached on the statement in question if the median was between 2.1 and 2.9.

3. Results

Herein we present the results of the Delphi survey and recommendations concerning the assessment and grading of tumor regression in the primary tumor and in nodal metastases. The detailed outcomes of the survey regarding the processing and gross examination of the resection specimen, as well as the cross sectioning/method of sampling, staining, and immunohistochemistry are presented in Table 1.

3.1. Assessment of tumor regression in response to neoadjuvant therapy

Microscopic tumor regression should be determined semiquantitatively (no need to calculate a precise percentage value), based on an estimation of the amount of residual viable neoplastic tissue in relation to the total assessable carcinoma area, including areas with chemotherapy-induced tissue injury (fibrosis or fibroinflammation).

Viable tumor cells are defined as those cells that are judged to be capable of proliferation. Nevertheless, this is highly subjective, and it may sometimes be impossible to conclude with absolute certainty whether a given tumor cell is viable or not. Likewise, differentiating between desmoplastic stroma, which is also neoplastic, and regressive stroma, which is an expression of response to therapy (treatment related), is also challenging. In assessing the relative amount of residual tumor/regressive changes, not only the cancer cells but also the desmoplastic stroma in their immediate vicinity should be interpreted as "cancer bed", unless obvious regressive fibrosis is seen. In these latter cases, the stroma in the immediate vicinity of the cancer cells should be interpreted as a sign of regression. "Dirty" necrosis, that is, necrosis consisting of nuclear debris, admixed and bordered by viable tumor cells, should not be interpreted as a sign of response. Only "infarct-like" ischemic necrosis, consisting of eosinophilic cytoplasmic remnants but lacking viable tumor cells, should be interpreted as a sign of response.

Table 2 Proposed system for the grading of histopathologic tumor regression in gastric adenocarcinoma submitted to neoadjuvant therapy in the primary tumor and metastatic LNs

TRG	Criteria	Classification of response
Primary tumor	% Viable tumor	
Grade 1	No residual tumor	Complete regression
Grade 2	<10% residual tumor	Subtotal regression
Grade 3	10%-50% residual tumor	Partial regression
Grade 4	>50% residual tumor	Minimal/no regression
Metastatic LNs	Microscopic findings	
Grade a	Only fibrosis or mucin lakes without viable cells	Complete regression
Grade b	Viable neoplastic cells together with regressive changes	Partial regression
Grade c	No signs of tumor response	No regression

Abbreviation: LNs, lymph nodes.

3.2. Grading of tumor regression

During the initial evaluation of the TRG systems in round 2, only 1 participant advocated the use of a 5-tiered system, while the remaining 5 recommended either a 3- or a 4-tiered system. In the third round, most experts (4/6) supported the use of a 4-tiered system. Subsequently, at the final round, a 4-tiered system based on the modified Becker grading system [15] was proposed by the majority:

- Grade 1: no residual tumor (complete response)
- Grade 2: less than 10% residual tumor (subtotal regression)
- Grade 3: 10% to 50% residual tumor (partial regression)
- Grade 4: greater than 50% residual tumor (minimal/no regression)

Examples of representative histologic images are provided in Fig. 1.

3.3. Assessment of regression of nodal metastases

There was total agreement among the participants that, apart from the conventional ypN stage (ie, total number of lymph nodes identified, number of lymph nodes with signs of metastatic involvement), the pathologist should also record the presence of any signs of tumor regression in the involved lymph nodes (ie, acellular mucin lakes, fibrosis, or aggregates of foamy macrophages). That is, assessment of regression is needed for lymph node metastases as well, but in contrast to the primary tumor, only qualitative assessment and description suffice and can be simplified using 3 categories, summarizing the findings in all the examined lymph nodes as follows:

- Grade a: complete response (only fibrosis or mucin lakes without viable cells)
- Grade b: partial response (viable neoplastic cells together with regressive changes)
- Grade c: no signs of tumor response

Examples of representative histologic images are provided in Fig. 2.

3.4. The role of histologic tumor type

Consensus was reached on the question whether the selection of procedures and the method of evaluation should depend on the histologic type of the tumor. A unanimous recommendation is that these should be the same irrespective of histologic type (Laurén, World Health Organization) or the presence of any possibly high-risk subtypes [21-23].

4. Discussion

The Delphi method is frequently used in clinical research areas that are dominated by significant diversities in opinions, substantial amount of information in the available literature which is mostly of limited scientific grade, and where a corresponding process is intended to form the platform for the launch of a subsequent research plan aiming to offer data of higher scientific grade. It has to be recognized that there are methodological issues in general confined to the application of the Delphi method. Among others, there seems to be little agreement regarding the size of the panel required, how to identify and select appropriate participants, or how to manage the data that the process generates [24-26]. Nevertheless, it is considered the best way to gain the most reliable consensus of opinion of a group of experts.

The aim of this Delphi study was to, taking base from the existing TRG systems in gastric cancer, generate a novel system avoiding the draw backs and limitations of some and adopting the perceived advantages of others. Another objective was to develop and specify the details of a protocol that would secure adequate and standardized histopathologic assessment, from the specimen processing and gross examination, to the method of sampling, the standard staining, and possible supplementary use of immunohistochemistry, to the way of estimating the tumor regression in the histologic slides. This need for international agreement is emphasized in the latest European Society for Medical Oncology clinical practice guidelines on gastric cancer, which state that "There should be national or preferably international guidelines for dissection and reporting" [7]. Mainly,

we aimed to reach consensus on the optimal way to stratify TRG in both the primary tumor and the nodal metastases. This methodology differs from alternative-complementary approaches, as represented by a recent critical review of the relevant literature by Langer and Becker [27]. A total of 4 rounds were required to reach consensus on most items (97%), with some areas deserving further discussion.

The issue of whether one should proceed with additional sectioning in cases of presumed complete response to confirm this finding was proven to be a matter of debate and one of the few topics where consensus was not reached. Half of the experts advocated that it is necessary to confirm this finding by additional sectioning in the area of interest and, if necessary, the whole stomach should be submitted to histologic examination as to allow for a definitive conclusion. According to others, this is not necessary as long as the recommendations regarding the method of sampling are followed correctly (see Table 1, Cross sectioning/method of sampling).

Immunohistochemistry for broad-spectrum cytokeratins (CK IHC) carries a potential role as an adjunct, if standard stains are inconclusive; otherwise, its use should not be uncritical, and it can currently not be recommended as part of the standard procedure. This section of the survey, covering various aspects on the use of IHC, was the one that generated most comments, and although all statements reached consensus in the final round, responses were accompanied by a plethora of remarks and in some cases with certain reservations expressed by the participants. One issue was whether CK IHC should be an essential part of the next step in cases of initial failure to detect residual cancer cells. Taking into account the considerable cost of an overuse of IHC, a compromise was proposed as an alternative, that is, to proceed with IHC on a number of selected sections where epithelial cells are not identifiable, but where the pathologist clearly notes the presence of signs of treatment response (fibrosis, accumulation of macrophages, mucin pools, etc) on the initial review of hematoxylin and eosin slides. To minimize interobserver variability, it may be wise to clearly define in advance in which situations CK IHC should be used, for example, cases with poorly cohesive carcinoma and overt inflammation. Hence, in such clearly defined cases, where the response is complete or near-complete, a standard CK IHC protocol on a few representative slides should be in place. Future research has to concentrate on these issues to offer robust and cost-effective guidelines.

The expert panel did not identify a need to evolve a novel system; instead, the 4-tiered grading system proposed by Becker and colleagues, with a minor modification, was favored by the majority (Table 2). The desirability of developing a 3-tiered system emerged, which makes sense as one of the goals in general is to use as few subdivisions as possible. At the same time, complexity must be balanced with maintaining the classification system's capacity to discriminate between patient groups with different outcomes. In another study based on data from a large number of patients (n = 480), multivariate analysis showed that the Becker grading system using 3 grades instead remained an independent prognostic factor for survival. Despite this finding, the authors suggested retaining the subdivision into grade 1a (complete response) and 1b (subtotal response) to preserve the information on the presence of complete tumor regression [28]. Similarly, the Delphi panel reached consensus on a 4-tiered system, as participants argued that (1) it is important to maintain a distinction between complete response and almost complete response (<10% of tumor remaining), and (2) grouping together patients with scarce or rare residual tumor cells, that is, nearly complete response, with those who have up to almost 50% residual tumor is not justified, as such subdivision would probably mask important clinical outcome differences.

The proposed grading system is different from that advocated by Becker et al [15] because it incorporates grading of tumor response in the nodal metastases. Although it is undisputable that the lymph node status offers major prognostic information in gastric cancer, previous studies on histopathologic tumor regression focused almost exclusively on the primary tumor. Nodal regression per se has rarely been studied, and therefore, its clinical implication remains unclear but emerges as an important target for future studies. Becker and colleagues [29], recognizing the relevance of lymph node status, developed a multifactorial histopathologic prognostic score, combining the classical ypT and ypN parameters with TRG. Ultimately, this system could, based on retrospective data analysis, identify 3 patient groups with significantly different prognosis. Nevertheless, the authors grouped together patients with complete or subtotal regression (<10% residual tumor); our conclusion is that these should remain 2 distinct subsets of patients. In addition, we argue that incorporating TRG in metastatic lymph nodes (as an adjunct to conventional ypN status) has the potential to offer a relevant complementary prognostic tool. In contrast to the primary tumor site, we suggest a qualitative assessment and simplified classification into 3 categories, that is, (*a*) complete response, (*b*) partial response, and (*c*) no signs of tumor response (Table 2). Regarding the interobserver and intraobserver variability in the assessment of TRG in general and in lymph nodes in particular, this has to be separately studied within the framework of dedicated prospective protocols. The same applies to the correlation of the proposed system with survival outcomes, as corresponding information is not available in the literature.

Another issue addressed in this survey was whether the assessment protocol should be tailored to the histologic subtype (Laurén, World Health Organization). This approach, although interesting, did not gain the panel's support, as most gastric carcinomas are mixed type and a tailored procedure would complicate an already difficult task.

There are a number of limitations in this study, as well as methodological issues in general, when applying the Delphi method; some unintentional "leading" by the investigators inevitably occurs because of the selection of the information communicated during the process of feedback. The lack of precise definition on what constitutes consensus is another limitation, and it should be emphasized that, even if consensus

is achieved, the conclusions represent at best an opinion and should be interpreted as such. Indeed, during the survey, the participating experts repeatedly stated that, in a number of issues, there is simply not enough evidence to support a conclusion and that more studies are needed. Finally, gastric cancer is common especially in East Asia, where the assessment of TRG follows the guidelines of the Japanese Classification of Gastric Carcinoma [9]. To minimize the complexity of the survey, we selected only experts from the Western hemisphere. Obviously, for a widespread dissemination of results like ours, it is vital that a corresponding survey is conducted in Asia as well and that the proposed system is subjected to the necessary validation in the respective clinical settings.

5. Conclusions

Standardization of the evaluation and reporting of histopathologic TRG after neoadjuvant treatment of gastric carcinoma is warranted and has to be implemented in routine clinical practice, as well as in research projects. Through the outcome of a comprehensive Delphi survey, an international group of experts proposes a 4-tiered system for the grading of regression in the primary tumor, combined with a 3-tiered system for the metastatic lymph nodes. In this system, grade 1 contains no residual tumor (complete response); grade 2, less than 10% residual tumor (subtotal regression); grade 3, 10% to 50% residual tumor (partial regression); and grade 4, greater than 50% residual tumor (minimal/no regression). The addition of "a", "b", or "c" indicates complete, partial, or no response in the metastatic lymph nodes. It is recommended that this tumor regression grading system be used irrespective of the histologic subtype of gastric carcinoma.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humpath.2018.08.028.

Acknowledgments

Donal Barrett, statistician at the Department of Clinical Science, Intervention and Technology (CLINTEC), Karolinska Institutet, is acknowledged for providing advice on the analysis and interpretation of data, as well as Dr Balint Melcher, Institute of Pathology, Klinikum Bayreuth, for providing representative histologic images.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-86.

[2] De Angelis R, Sant M, Coleman MP, et al. Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5—a population-based study. Lancet Oncol 2014;15:23-34.

[3] Wilke H, Preusser P, Fink U, et al. Preoperative chemotherapy in locally advanced and nonresectable gastric cancer: a phase II study with etoposide, doxorubicin, and cisplatin. J Clin Oncol Off J Am Soc Clin Oncol 1989;7:1318-26.

[4] Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med 2006;355:11-20.

[5] Ychou M, Boige V, Pignon JP, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol Off J Am Soc Clin Oncol 2011;29:1715-21.

[6] Al-Batran S-E, Homann N, Schmalenberg H, et al. Perioperative chemotherapy with docetaxel, oxaliplatin, and fluorouracil/leucovorin (FLOT) versus epirubicin, cisplatin, and fluorouracil or capecitabine (ECF/ECX) for resectable gastric or gastroesophageal junction (GEJ) adenocarcinoma (FLOT4-AIO): a multicenter, randomized phase 3 trial. J Clin Oncol 2017;35:4004.

[7] Smyth EC, Verheij M, Allum W, Cunningham D, Cervantes A, Arnold D. Gastric cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2016;27:v38-49.

[8] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009:228-47.

[9] Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma: 3rd English edition. Gastric Cancer 2011;14:101-12.

[10] Kurokawa Y, Shibata T, Ando N, Seki S, Mukaida H, Fukuda H. Which is the optimal response criteria for evaluating preoperative treatment in esophageal cancer: RECIST or histology? Ann Surg Oncol 2013;20:3009-14.

[11] Kurokawa Y, Shibata T, Sasako M, et al. Validity of response assessment criteria in neoadjuvant chemotherapy for gastric cancer (JCOG0507-A). Gastric Cancer 2014;17:514-21.

[12] Tomasello G, Petrelli F, Ghidini M, et al. Tumor regression grade and survival after neoadjuvant treatment in gastro-esophageal cancer: a meta-analysis of 17 published studies. Eur J Surg Oncol 2017;43:1607-16.

[13] Spoerl S, Novotny A, Al-Batran SE, et al. Histopathological regression predicts treatment outcome in locally advanced esophagogastric adenocarcinoma. Eur J Cancer 2018:26-33.

[14] Al-Batran SE, Hofheinz RD, Pauligk C, et al. Histopathological regression after neoadjuvant docetaxel, oxaliplatin, fluorouracil, and leucovorin versus epirubicin, cisplatin, and fluorouracil or capecitabine in patients with resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4-AIO): results from the phase 2 part of a multicentre, open-label, randomised phase 2/3 trial. Lancet Oncol 2016;17:1697-708.

[15] Becker K, Mueller JD, Schulmacher C, et al. Histomorphology and grading of regression in gastric carcinoma treated with neoadjuvant chemotherapy. Cancer 2003;98:1521-30.

[16] Kiyabu M, Leichman L, Chandrasoma P. Effects of preoperative chemotherapy on gastric adenocarcinomas. A morphologic study of 25 cases. Cancer 1992;70:2239-45.

[17] Mansour JC, Tang L, Shah M, et al. Does graded histologic response after neoadjuvant chemotherapy predict survival for completely resected gastric cancer? Ann Surg Oncol 2007;14:3412-8.

[18] Charalampakis N, Nogueras Gonzalez GM, Elimova E, et al. The proportion of signet ring cell component in patients with localized gastric adenocarcinoma correlates with the degree of response to pre-operative chemoradiation. Oncology 2016;90:239-47.

[19] Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. Cancer 1994;73:2680-6.

[20] Chirieac LR, Swisher SG, Ajani JA, et al. Posttherapy pathologic stage predicts survival in patients with esophageal carcinoma receiving preoperative chemoradiation. Cancer 2005;103:1347-55.

[21] Messager M, Lefevre JH, Pichot-Delahaye V, Souadka A, Piessen G, Mariette C. The impact of perioperative chemotherapy on survival in

patients with gastric signet ring cell adenocarcinoma: a multicenter comparative study. Ann Surg 2011;254:684-93.
[22] Heger U, Blank S, Wiecha C, et al. Is preoperative chemotherapy followed by surgery the appropriate treatment for signet ring cell containing adenocarcinomas of the esophagogastric junction and stomach? Ann Surg Oncol 2014;21:1739-48.
[23] Jimenez Fonseca P, Carmona-Bayonas A, Hernandez R, et al. Lauren subtypes of advanced gastric cancer influence survival and response to chemotherapy: real-world data from the AGAMENON National Cancer Registry. Br J Cancer 2017;117:775-82.
[24] Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000;32:1008-15.
[25] Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003; 41:376-82.
[26] Holey EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. BMC Med Res Methodol 2007;7:52.
[27] Langer R, Becker K. Tumor regression grading of gastrointestinal cancers after neoadjuvant therapy. Virchows Arch 2018;472: 175-86.
[28] Becker K, Langer R, Reim D, et al. Significance of histopathological tumor regression after neoadjuvant chemotherapy in gastric adenocarcinomas: a summary of 480 cases. Ann Surg 2011;253: 934-9.
[29] Becker K, Reim D, Novotny A, et al. Proposal for a multifactorial prognostic score that accurately classifies 3 groups of gastric carcinoma patients with different outcomes after neoadjuvant chemotherapy and surgery. Ann Surg 2012;256:1002-7.