

## Diagnostic criteria and symptom grading for delayed gastric conduit emptying after esophagectomy for cancer: international expert consensus based on a modified Delphi process

M. Konradsson,<sup>1,2</sup> M. I. van Berge Henegouwen<sup>3</sup>, C. Bruns<sup>4</sup>, M. A. Chaudry<sup>5</sup>, E. Cheong<sup>6</sup>, M. A. Cuesta<sup>7</sup>, G. E. Darling<sup>8</sup>, S. S. Gisbertz<sup>3</sup>, S. M. Griffin<sup>9</sup>, C. A. Gutschow, R. van Hillegersberg<sup>11</sup>, W. Hofstetter<sup>12</sup>, A. H. Hölscher<sup>13</sup>, Y. Kitagawa<sup>14</sup>, J. J. B. van Lanschot<sup>15</sup>, M. Lindblad<sup>1,16</sup>, L. E. Ferri<sup>17</sup>, D. E. Low<sup>18</sup>, M. D. P. Luyer<sup>19</sup>, N. Ndegwa<sup>1</sup>, S. Mercer<sup>20</sup>, K. Moorthy<sup>28</sup>, C. R. Morse<sup>22</sup>, P. Naftoux<sup>23,24</sup>, G. A. P. Nieuwehuijzen<sup>19</sup>, P. Pattyn<sup>25</sup>, C. Rosman<sup>26</sup>, J. P. Ruurda<sup>11</sup>, J. Räsänen<sup>27</sup>, P. M. Schneider<sup>28</sup>, W. Schröder<sup>4</sup>, B. Sgromo<sup>29</sup>, H. Van Veer<sup>23,24</sup>, B. P. L. Wijnhoven<sup>15</sup>, M. Nilsson<sup>1,16</sup>

<sup>1</sup>Department of Clinical Science, Intervention and Technology, Karolinska Institute, Stockholm, Sweden, <sup>2</sup>Department of Gastroenterology, Landspítali National University Hospital, Reykjavik, Iceland, <sup>3</sup>Amsterdam UMC, location AMC, University of Amsterdam, Department of Surgery, Cancer Center Amsterdam, <sup>4</sup>Department of General, Visceral and Cancer Surgery, University of Cologne, Cologne, Germany, <sup>5</sup>Department of Surgery, Royal Marsden Hospital, London, UK, <sup>6</sup>Norfolk and Norwich University Hospital, Norwich, UK, <sup>7</sup>Department of Surgery, Amsterdam UMC, location VUmc, Amsterdam, Netherlands, <sup>8</sup>Department of Surgery, Division of Thoracic Surgery, University Health Network, University of Toronto, Toronto, Canada, <sup>9</sup>Northern Oesophagogastric Unit, Royal Victoria Infirmary, Newcastle upon Tyne, UK, <sup>10</sup>Department of Surgery and Transplantation, University Hospital Zurich, Zurich, Switzerland, <sup>11</sup>Universitair Medisch Centrum Utrecht, Utrecht, The Netherlands, <sup>12</sup>Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas, <sup>13</sup>Centre for Esophageal and Gastric Surgery, AGAPLESION Markus Krankenhaus, Frankfurt, Germany, <sup>14</sup>Department of Surgery, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-Ku, Tokyo, 160-8582, Japan, <sup>15</sup>Department of Surgery, Erasmus MC, University Medical Center, Rotterdam, The Netherlands, <sup>16</sup>Department of Upper Abdominal Diseases, Karolinska University Hospital, Stockholm, Sweden, <sup>17</sup>Department of Thoracic and Upper Gastrointestinal Surgery, McGill University Health Centre, Montreal, QC, Canada, <sup>18</sup>Virginia Mason Medical Center, Seattle, WA, USA, <sup>19</sup>Department of Surgery, Catharina Ziekenhuis, Eindhoven, The Netherlands, <sup>20</sup>Queen Alexandra Hospital Portsmouth, United Kingdom, <sup>21</sup>Department of Surgery and Cancer, Imperial College London, London, UK, <sup>22</sup>Division of Thoracic Surgery, Massachusetts General Hospital, Boston, Massachusetts, USA, <sup>23</sup>Department of Thoracic Surgery, University Hospitals Leuven, Leuven, Belgium, <sup>24</sup>Department of Chronic Diseases, Metabolism and Aging, KU Leuven, Belgium, <sup>25</sup>Department of Gastrointestinal Surgery, Ghent University Hospital, Ghent, Belgium, <sup>26</sup>Department of surgery, Radboud university center Nijmegen, The Netherlands, <sup>27</sup>Department of General, Thoracic and Esophageal Surgery, Heart and Lung Center, Helsinki University Hospital and Helsinki University, Helsinki, Finland, <sup>28</sup>The Center for Visceral, Thoracic and Specialized Tumor Surgery, Hirslanden Medical Center, Zurich, Switzerland, and <sup>29</sup>Oxford University Hospitals, Oxford, UK.

**SUMMARY.** Delayed gastric conduit emptying (DGCE) after esophagectomy for cancer is associated with adverse outcomes and troubling symptoms. Widely accepted diagnostic criteria and a symptom grading tool for DGCE are missing. This hampers the interpretation and comparison of studies. A modified Delphi process, using repeated web-based questionnaires, combined with live interim group discussions was conducted by 33 experts within the field, from Europe, North America, and Asia. DGCE was divided into early DGCE if present within 14 days of surgery and late if present later than 14 days after surgery. The final criteria for early DGCE, accepted by 25 of 27 (93%) experts,

Address correspondence to: Magnus Konradsson, MD, Department of Clinical Science, Investigation and Technology (CLINTEC), Karolinska Institutet, 14186 Stockholm, Sweden. Email: magnus.konradsson@ki.se

Specific author contribution: Design: Magnus Konradsson, Magnus Nilsson, and Mats Lindblad. Conception, administration of Delphi study, data extraction, data analysis, and drafting of article: Magnus Konradsson and Magnus Nilsson. Statistical analysis: Nelson Ndegwa. Chairman of Interim Live Group Discussions: Magnus Nilsson. Participation in Delphi study and Interim Live Group discussions: All authors except Mats Lindblad and Nelson Ndegwa. Revision and final approval of article: All authors.

© The Author(s) 2019. Published by Oxford University Press on behalf of International Society for Diseases of the Esophagus.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

were as follows: >500 mL diurnal nasogastric tube output measured on the morning of postoperative day 5 or later or >100% increased gastric tube width on frontal chest x-ray projection together with the presence of an air–fluid level. The final criteria for late DGCE accepted by 89% of the experts were as follows: the patient should have ‘quite a bit’ or ‘very much’ of at least two of the following symptoms; early satiety/fullness, vomiting, nausea, regurgitation or inability to meet caloric need by oral intake and delayed contrast passage on upper gastrointestinal water-soluble contrast radiogram or on timed barium swallow. A symptom grading tool for late DGCE was constructed grading each symptom as: ‘not at all’, ‘a little’, ‘quite a bit’, or ‘very much’, generating 0, 1, 2, or 3 points, respectively. For the five symptoms retained in the diagnostic criteria for late DGCE, the minimum score would be 0, and the maximum score would be 15. The final symptom grading tool for late DGCE was accepted by 27 of 31 (87%) experts. For the first time, diagnostic criteria for early and late DGCE and a symptom grading tool for late DGCE are available, based on an international expert consensus process.

**KEY WORDS:** consensus, esophagectomy, gastric emptying, malnutrition.

## INTRODUCTION

Multimodal treatment including esophagectomy with gastric conduit reconstruction is the treatment of choice for esophageal cancer. Recent improvements in survival after curative treatment put additional focus on short- and long-term morbidity and functional outcome.<sup>1, 2</sup> Delayed gastric conduit emptying (DGCE) is recognized as one of the most important oesophagectomy problems. DGCE is associated with short-term adverse outcomes including anastomotic leakage, pneumonia, increased ICU, and total hospital stay. DGCE may also lead to nutritional problems and a reduced quality of life.<sup>2–6</sup>

The pathophysiology of DGCE is not fully understood. Proposed contributing causes are pyloric and gastric dysmotility due to vagal and sympathetic denervation, mobilization of the conduit from a positive pressure compartment to a negative pressure compartment, and disruption of the native antireflux mechanisms.<sup>7, 8</sup> Also other factors, such as conduit size or reconstructive route, relate to variations in clinical presentation.<sup>8, 9</sup> No widely accepted diagnostic criteria and tools to evaluate the presence and severity of DGCE are available. This has led to limitations in the assessment of the current literature on the incidence of DGCE and the effect of preventive and therapeutic measures.<sup>9, 10</sup>

To establish diagnostic criteria and symptom grading for DGCE, we conducted an international expert consensus process, based on modified Delphi methodology. This was combined with live interim group discussions. The Delphi process is a questionnaire-based method that enables experts to express their opinion independently, avoiding the risk of dominant speakers influencing consensus work in open sessions.<sup>11, 12</sup> The Delphi process methodology is well established and has previously been used to achieve expert consensus within various medical fields.<sup>13, 14</sup>

The aims of this expert consensus process were (i) to reach international agreement regarding diagnostic criteria for DGCE after esophagectomy with gastric conduit reconstruction, both in the immediate postoperative phase and in a more

long-term perspective and (ii) to reach international agreement regarding a symptom grading tool for DGCE after esophagectomy with gastric conduit reconstruction.

## MATERIALS AND METHODS

### Modified Delphi expert consensus process

#### *Expert group*

This international expert consensus process was a collaboration between experts in Europe, North America, and Asia, using repeated online questionnaires and interim live group discussions. The consortium of 33 experts included a previously assembled collaborative group of surgeons in Europe within the field of minimally invasive esophagectomy, the European Minimally Invasive Esophagectomy Think Tank, supplemented by a number of international leading esophageal surgeons listed in [Table 1](#). A modified Delphi process with live interim group discussions was used to gradually achieve consensus on the exact combination of symptoms, clinical findings, and diagnostic modalities required for the diagnosis of DGCE. The same method was used to reach expert consensus on symptoms and methods used for severity grading of DGCE.

#### *Delphi study*

A literature search was performed in April 2017 to scan the published literature for possibly relevant symptoms and diagnostic modalities suggested to represent DGCE. The search was performed in PubMed using the following MESH terms: esophagectomy, gastric emptying, and gastric outlet obstruction; additionally, we complemented the search with the following terms: delayed gastric emptying and gastric conduit dysfunction. The Delphi round 1 questionnaire included multiple choice questions with options graded on a 5-point Likert scale evaluating the experts’ opinion on both frequency of each symptom (symptom present in 0–20%, 21–40%,

**Table 1** Participating experts

Name	Country	Institution
Mark I van Berge Henegouwen	The Netherlands	Amsterdam UMC, location AMC, University of Amsterdam, Cancer Center Amsterdam
Christiane Bruns	Germany	University Hospital Cologne
Asif Chaudry	UK	The Royal Marsden Cancer Centre, London
Edward Cheong	UK	Norfolk & Norwich University Hospital
Miguel Cuesta	The Netherlands	Free University Medical Center Amsterdam
Gail E. Darling	Canada	University Health Network
Suzanne S Gisberz	The Netherlands	Amsterdam UMC, location AMC, University of Amsterdam, Cancer Center Amsterdam
Michael Griffin	UK	Royal Victoria Infirmary, Newcastle
Christian Gutschow	Switzerland	University Hospital Zurich
Richard van Hillegersberg	The Netherlands	University Medical Center Utrecht
Wayne Hofstetter	USA	Johns Hopkins Cancer Center, Houston
Arnulf Hölscher	Germany	Agaplesion Markus Hospital Frankfurt
Yuko Kitagawa	Japan	Keio University Hospital, Tokyo
Magnus Konradsson	Sweden/Iceland	Landspítali National University Hospital
Jan JB van Lanschot	The Netherlands	Erasmus Medical Center, Rotterdam
Lorenzo Ferri	Canada	Montreal General Hospital
Donald Low	USA	Virginia Mason Medical Center, Seattle
Misha D Luyer	The Netherlands	Catharina Hospital, Eindhoven
Stuart Mercer	UK	Queen Alexandra Hospital Portsmouth
Krishna Moorthy	UK	St. Mary's Hospital, London
Christopher Morse	USA	Massachusetts General Hospital, Boston
Philippe Naftoux	Belgium	Leuven University Hospital
Grard A Nieuwenhuijzen	The Netherlands	Catharina Hospital, Eindhoven
Magnus Nilsson	Sweden	Karolinska Institutet, Stockholm
Piet Pattyn	Belgium	Ghent University
Camiel Rosman	The Netherlands	Radboud University Medical Center
Jelle P Ruurda	The Netherlands	Utrecht Medical Center
Jari Räsänen	Finland	Helsinki University Hospital
Paul M Schneider	Switzerland	Hirslanden Medical Center, Zürich
Wolfgang Schröder	Germany	University of Cologne
Bruno Sgromo	UK	Oxford University Hospitals
Hans Van Veer	Belgium	Leuven University Hospital
Bas PL Wijnhoven	The Netherlands	Erasmus Medical Center, Rotterdam

41–60%, 61–80%, or 81–100% of cases of DGCE) and specificity of each symptom (not at all specific, slightly specific, moderately specific, specific, or very specific) for DGCE. For diagnostic modalities, experts were asked to grade the strength of support for the diagnosis of DGCE that each diagnostic modality would provide on a 5-point Likert scale. Opportunities were given in each Delphi round to suggest additional items, and fields for additional comments were available throughout the questionnaires. Repeated Delphi rounds were planned until consensus would be achieved on which symptoms and diagnostic modalities should be kept to form diagnostic criteria, based on successive exclusions during the Delphi process. Repeated Delphi rounds were planned until consensus was reached on diagnostic criteria and a symptom grading tool for DGCE. After each Delphi round, a live interim group discussion was planned, until the last Delphi round would test the final product of the consensus process. In the last Delphi round, basic demography of participants was gathered. The Delphi process was performed using an online survey-system (SurveyMonkey, Palo Alto, CA). The complete Delphi survey questionnaires are provided in the [Supplementary Appendix 1](#).

*Live interim group discussions*

After analysis of the results of each Delphi round, live interim group discussions were planned as a part of the consensus process, until the last Delphi round. The group discussions were planned to provide opportunities to raise safety concerns, discuss practical issues, and provide suggestions for further developments to the diagnostic criteria and symptom severity tool to be considered in following Delphi rounds on the basis of the results of previous rounds.

**Statistics and consensus algorithms**

Consensus to exclude any symptom from the diagnostic criteria was reached when at least 50% of the experts considered frequency or specificity of the symptom to be in the lower two levels of a 5-level Likert scale. Consensus to exclude any diagnostic modality was reached when at least 50% of the experts considered the strength of the diagnostic modality supporting the diagnosis to be of the two lowest levels of a 5-level Likert scale.

In Delphi round 2, consensus was considered to be achieved if 50% agreed on 1–2 of the peripheral levels of a 5-level Likert's scale, and in binary

**Table 2** Modified Delphi process overview

Delphi questionnaire rounds	Response rate	Completion rate	Time period
Delphi round 1	28/33 (85%)	100%	August 18, 2017–November 20, 2017
Delphi round 2	32/33 (97%)	97%	March 12, 2018–May 13, 2018
Delphi round 3	27/33 (82%)	100%	October 2, 2018–October 24, 2018
Live interim group discussions	Participants	Place	Date
After Delphi 1	25/33 (76%)	Utrecht	December 8, 2017
After Delphi 2	23/33 (70%)	Vienna	September 20, 2018

questions or multiple choice questions not compatible with a Likert's scale, if 70% of experts agreed on one answer. Regarding diagnostic criteria and symptom grading, consensus was considered to be achieved if 80% agreed upon one option.

We classified questions regarding diagnostic criteria and symptom score for DGCE into two domains depending on whether they occurred during (domain 1) or after (domain 2) the first 14 days postesophagectomy. We calculated Cronbach alpha for each of the domains to validate internal consistency.

Cronbach alpha was calculated using Stata 14 (StataCorp. 2015. Stata Statistical Software: Release 14; StataCorp LP, College Station, TX).

## RESULTS

The Delphi questionnaires were sent to all 33 experts for completion. [Table 1](#) lists all participating experts, and [Table 2](#) shows an overview of the modified Delphi expert consensus process.

### Expert demographics and study overview

The mean (range) number of esophagectomies performed by the responding 27 of 33 experts' institutions was 81 (30–180) per year. The mean time the experts had been actively engaged in the surgical treatment of patients with esophageal cancer was 18 (6–35) years.

### Relevance of time point of DGCE diagnosis

The Delphi round 1 questionnaire addressed whether a proposed early–late DGCE dichotomy should be established. Consensus was reached that DGCE should be classified as early if diagnosed within the first 14 days of surgery and late if diagnosed later.

### Early DGCE

#### *Frequency and specificity of symptoms typical for early DGCE*

Based on the literature search 17 symptoms were evaluated. In Delphi round 1, 14 of the 17 symptoms for early DGCE were excluded according to the criteria of the consensus algorithm (see [Table 3](#)). At the

first live interim group discussion held on December 8, 2017, concern was raised that even the symptoms retained after the first Delphi session would be too unspecific and rather represents general postoperative symptoms. Consensus to abort both further attempts to incorporate symptoms in the early DGCE diagnostic criteria and to design a symptom grading scale for early DGCE was reached in Delphi round 2 by 29 of 32 (91%) and 26 of 31 (84%) experts, respectively.

#### *Diagnostic criteria for early DGCE*

Six diagnostic modalities were included based on the literature search. After Delphi round 1, all diagnostic modalities (chest x-ray, upper gastrointestinal [GI] water-soluble contrast radiogram, upper GI endoscopy, gastric scintigram, computerized tomography with oral contrast, and timed barium swallow) remained as candidates for further evaluation ([Table 4](#)). At the first live interim group discussion after Delphi round 1, a unanimous suggestion regarding early DGCE diagnostic criteria emerged as follows: large amount of nasogastric tube drainage fluid or dilated gastric conduit on frontal chest x-ray and upper GI water-soluble contrast radiogram showing delayed contrast passage to the duodenum.

These diagnostic criteria were subsequently tested in Delphi round 2 and were accepted by 29 of 32 (91%) experts, in effect achieving formal consensus. However, at the second live interim discussion held on September 20, 2018, several participating experts questioned the lack of standardization in the performance and interpretation of upper GI water-soluble contrast radiogram in the diagnosis of early DGCE. In addition, some experts expressed concern regarding risk of aspiration associated with this diagnostic procedure in the immediate postoperative phase. This eventually led to a revised suggestion, excluding upper GI water-soluble contrast radiogram, which was unanimously supported by the live meeting participants. The revised diagnostic criteria for early DGCE were subsequently tested in Delphi round 3. The final early DGCE criteria, accepted by 25 of 27 (93%) experts in Delphi round 3 were as follows: >500-mL diurnal nasogastric tube output measured



**Table 3** Suggested symptoms of early and late DGCE and results of Delphi round 1

Suggested symptoms of DGCE	Early DGCE			Late DGCE		
	Frequency of occurrence considered 40% or lower <sup>‡</sup>	Specificity considered none or very low <sup>§</sup>	Action after Delphi round 1	Frequency of occurrence considered 40% or lower <sup>‡</sup>	Specificity considered none or very low <sup>§</sup>	Action after Delphi round 1
Vomiting	53.6%	21.4%	Excluded	25.0%	17.9%	FE
Nausea	42.9%	53.6%	Excluded	32.1%	42.9%	FE
Dysphagia to solids	53.6%	60.7%	Excluded	53.6%	57.1%	Excluded
Dysphagia to liquidized/soft food	64.3%	53.6%	Excluded	71.4%	57.1%	Excluded
Dysphagia to liquids	82.1%	53.6%	Excluded	78.6%	67.9%	Excluded
Oral intake intolerance <sup>†</sup>	50.0%	42.9%	Excluded	42.9%	28.6%	FE
Early satiety/fullness	10.7%	28.6%	FE	10.7%	25.0%	FE
Regurgitation	17.9%	32.1%	FE	25.0%	25.0%	FE
Heart burn (cervical)	60.7%	82.1%	Excluded	60.7%	64.3%	Excluded
Bloating	53.6%	67.9%	Excluded	60.7%	64.3%	Excluded
Chest pressure	60.7%	78.6%	Excluded	53.6%	67.9%	Excluded
Pain	82.1%	96.4%	Excluded	85.7%	89.3%	Excluded
Coughing	35.7%	57.1%	Excluded	32.1%	50.0%	Excluded
Recurring pneumonia	78.6%	46.4%	Excluded	67.9%	42.9%	Excluded
Loss of appetite	39.3%	67.9%	Excluded	35.7%	67.9%	Excluded
Inability to meet caloric need by oral intake	17.9%	57.1%	Excluded	21.4%	35.7%	FE
Large amount of gastric tube drainage fluid (only suggested in early DGCE)	39.3%	14.3%	FE	NA	NA	NA

<sup>†</sup> Excluded in Delphi 2 due to redundancy.

<sup>‡</sup> Percentage of experts considering symptom occurring in 40% of DGCE cases or less in early and late DGCE, respectively.

<sup>§</sup> Percentage of experts considering symptom specificity very low or none in early and late DGCE, respectively.

FE, further evaluation in the consensus process.

**Table 4** Suggested diagnostic modalities for early and late DGCE and results of Delphi round 1

Diagnostic modality	Early DGCE		Late DGCE	
	Support of diagnosis considered slight or none <sup>†</sup>	Action after Delphi round 1	Support of diagnosis considered slight or none <sup>†</sup>	Action after Delphi round 1
Delayed contrast passage to the duodenum on upper gastrointestinal water-soluble contrast radiogram	3.6%	FE	10.7%	FE
Wide gastric tube with minimal contrast passage below the pylorus on CT thorax–abdomen with on table oral contrast	3.6%	FE	3.6%	FE
Dilated gastric tube on chest x-ray	3.6%	FE	25.0%	FE
Delayed esophagogastric emptying on timed barium swallow	10.7%	FE	10.7%	FE
50% gastric emptying time (T50) > 180 minutes on gastric scintigraphy	21.4%	FE	17.9%	FE
Marked retention of food on upper gastrointestinal endoscopy despite >4 hours fasting	14.3%	FE	10.7%	FE

<sup>†</sup> Percentage of experts considering the strength of support to the diagnosis slight or none, in early and late DGCE, respectively.

on the morning of postoperative day 5 or later (but within 14 days of surgery) or > 100% increased gastric tube width on frontal chest x-ray projection (in comparison to baseline chest x-ray taken on the day

of surgery) together with the presence of an air–fluid level within 14 days of surgery (Fig. 1).

The Cronbach alpha level of internal consistency for domain 1 with 47 items was 0.89.



patient should have a specified number of DGCE symptoms (selected in Delphi round 1) and DGCE on functional imaging. The framework design for diagnostic criteria of late DGCE was accepted by 30 of 32 (94%) experts in Delphi round 2. Opinion was divided in Delphi round 2 regarding the number of symptoms required for the diagnosis of late DGCE. Some 14 of 31 experts considered 2 symptoms appropriate and 14 of 31 three symptoms. The second criterion within the framework design was that diagnosis of DGCE should be supported by functional imaging. The preferred functional imaging was water-soluble contrast swallow for 48.4%, barium swallow for 32.3%, and scintigraphy for 19.3% of participants. At the second live interim group discussion, a first definition of the diagnostic criteria for late DGCE was proposed, unanimously suggesting a minimum number of symptoms of two and additionally a minimum severity of each of those symptoms of 2 points (presence of the symptom graded by the patient as 'quite a bit' or 'very much') on a 0–3 points scale. Furthermore, the functional imaging was unanimously suggested to include both the option of barium swallow and the option of upper GI water-soluble contrast radiogram. The revised diagnostic criteria for late DGCE were subsequently tested in Delphi round 3. The final late DGCE criteria accepted by 24 of 27 (89%) experts were as follows: the patient should have 'quite a bit' or 'very much' of at least two of the following symptoms: early satiety/fullness, vomiting, nausea, regurgitation or inability to meet caloric need by oral intake and delayed contrast passage on upper GI water-soluble contrast radiogram or on timed barium swallow (until precise evaluation criteria are available, relying on the verdict 'delayed contrast passage' by an expert radiologist) (Fig. 1).

#### *Symptom grading tool for late DGCE*

The design of a symptom grading tool based on the symptom grading system used in the EORTC Health related Quality of Life Questionnaires was accepted by 26 of 28 (93%) experts in Delphi round 1.<sup>15</sup> Consequently, when reporting symptom severity, the presence and severity of each symptom should be graded as follows: 'not at all', 'a little', 'quite a bit', or 'very much', generating 0, 1, 2, or 3 points, respectively. For the five symptoms retained after Delphi round 2 as presented above, the minimum score would be 0, and the maximum score would be 15. The final symptom grading tool for late DGCE presented in Figure 1 was accepted by 27 of 31 (87%) experts in Delphi round 2. A questionnaire for symptom grading of DGCE after esophagectomy is presented in Figure 2.

The Cronbach alpha level of internal consistency for domain 2 with 45 items was 0.85.

## DISCUSSION

To date, the studies focusing on DGCE have relied on local practice or diagnostic criteria created for the purpose of each individual study, making valid comparisons between studies and summaries of results difficult.<sup>10</sup> The need for widely accepted diagnostic criteria has been mentioned in the conclusion of several trials, particularly in studies attempting to compare or summarize the results of multiple studies.<sup>9, 10, 16</sup> In this study, a modified Delphi process was used based on the opinion of an international expert panel, to establish the diagnostic criteria for early and late DGCE, and a symptom severity grading tool for late DGCE.

These results can be used to compare the outcomes of DGCE in various studies determining post-esophagectomy morbidity and quality of life. In addition, the symptom grading tool for late DGCE provides a common symptom grading system.

A strength of this consensus process is the combination of the Delphi method and the live group discussions. The Delphi part allowed all participants to independently express their opinion without any peer pressure or influence of dominant speakers.<sup>11</sup> Participants were prompted to comment specific parts of each Delphi round, and free opportunity for additional comments or proposals was given. The interim group discussions allowed for an open, structured discussion on relevant topics regarding safety, clinical relevance, and feasibility of specific parts of the diagnostic criteria and symptom score. Any suggestion, addition, or revision of former results provided in the interim group discussions was put to test in a subsequent Delphi round. A further strength of this study is the international participation of experts from three continents, indicating that the principal consensus statement on DGCE may be adapted in various countries despite local practice differences.

The live interim group discussions unavoidably carry a risk of strong opinion affecting the consensus process, and approximately  $\frac{1}{4}$  of the expert group were not present at those meetings (Table 2). A majority of the experts did, however, participate in the live interim group discussions, and care was taken that any suggestions formed at those meetings, such as the final statement, would be tested in a subsequent survey. A weakness of this study is that available scientific evidence is limited, and the statement is thus based only on expert opinion. Further studies are mandatory to validate the diagnostic criteria and symptom score against diagnostic modalities. Despite the lack of validation, it may be argued that the diagnostic criteria and symptom score carry at least the same validity as criteria singularly created for the purpose of solitary studies and furthermore carries the strength of a

Please answer these questions by circling the answer below (0-3) that best describes your situation during the past week.

Name: \_\_\_\_\_

Birth date: \_\_\_\_\_

Today's date: (day, month, year) \_\_\_\_\_

### During the past week:

	Not at all	A little	Quite a bit	Very much
1. Have you felt full up too quickly while having your meal?	0	1	2	3
2. Have you vomited?	0	1	2	3
3. Have you felt nausea?	0	1	2	3
4. Have you had acid, bile or food coming up into your throat or mouth?	0	1	2	3
5. Have you been unable to eat or drink enough to meet your daily need for energy?	0	1	2	3

**Fig. 2** DGCE questionnaire after esophagectomy with gastric conduit reconstruction.

consensus reached within an international group of experts.

This study sought to provide diagnostic criteria for early and late DGCE and a symptom severity score for DGCE regardless of whether a whole stomach or a gastric tube was used as a conduit. To avoid disparity in the consensus process regarding the use of chest x-ray in the early DGCE diagnostic criteria, all items concerning functional radiology in the questionnaires and during interim group discussions were focused on the use of a gastric tube conduit. Whether the diagnostic criteria for early DGCE are less applicable in the presence of a whole stomach conduit remains to be evaluated.

In conclusion, this modified Delphi consensus process provides diagnostic criteria for early and late DGCE and a symptom grading tool supported by a group of international experts. This will be an important tool that can be used for future studies and allows defining a benchmark on DGCE.

### SUPPLEMENTARY DATA

Supplementary data are available at *DOTESO* online.

### ACKNOWLEDGMENTS

The European Minimally Invasive Esophagectomy (MIO) Think Tank group is sponsored by Medtronic. The following authors are members of MIO Think Tank group: Mark I. van Berge Henegouwen, Christiane Bruns, Asif Chaudry, Edward Cheong, Miguel A. Cuesta, Suzanne S. Gisbertz, Christian Gutschow, Richard van Hillegersberg, Arnulf Hölscher, Misha Luyer, Stuart Mercer, Krishna Moorthy, Philippe Nafteux, Grard Nieuwenhuijzen, Piet Pattyn, Camiel Rosman, Jelle P. Ruurda, Jari Räsänen, Paul M. Schneider, Wolfgang Schröder, Bruno Sgromo, Hans Van Veer, and Magnus Nilsson.



