Validated clinico-pathologic nomogram in the prediction of HER2 status in gastro-oesophageal cancer

Lorenzo Fornaro1, Caterina Vivaldi1, Annamaria Parmofielo2, Clara Ugolini3, Giuseppe Aprile4, Giovanna De Maglio5, Irene Pecora1, Donatella Iacono2, Francesca Crivelli6, Lorenzo Calvetti4, Alfredo Falcone1, Gianpiero Fasola2, Gabriella Fontanini8 and Donatella Iacono2, Francesca Crivelli6, Silvia Catanese1, Giovanni Gerardo Cardellino2, Monica Lencioni7, Enrico Vasile1, Francesco Montagnani6

BACKGROUND: HER2 is the only validated predictive biomarker in gastro-oesophageal carcinoma (GOC). However, several factors, such as heterogeneity in protein expression, shortage of evaluable tumour tissue and need for quick target assessment, underline the usefulness of a pre-screening tool in order to anticipate HER2 status.

METHODS: Data from 723 consecutive GOC analysed for HER2 at four Italian Institutions were collected. HER2 positivity was defined as 3+ by immunohistochemistry (IHC) or 2+ with gene amplification by in situ hybridisation (ISH). A multivariate logistic regression model was built using data from 413 cases, whereas 310 patients served as validation cohort. C-index, visual inspection of the calibration plot, Brier score and Spiegelhalter z-test were used to assess the performance of the nomogram.

RESULTS: HER2 positive rate was 17.4%. Four variables were retained after adjustment in the final model: grading, Lauren’s histotype, pathologic material analysed (surgical specimen/biopsy) and site of tissue collection (primary tumour/metastases). Visual inspection of the calibration plot revealed a very good overlap between predicted and observed probabilities, with a Brier score of 0.101 and a non-significant Spiegelhalter z-test (P = 0.319). C-index resulted in 0.827 (95%CI 0.741–0.913).

CONCLUSION: A simple nomogram based on always-available pathologic information accurately predicts the probability of HER2 positivity in GOC.

INTRODUCTION
Gastro-oesophageal carcinoma (GOC) represents the third leading cause of cancer death worldwide,1 with a rising incidence of junctional cancers2 as well as tumours developing among younger individuals.3 In this tumour, the identification of new targets for effective treatment represents an unmet need.4 Up to now, the human epidermal growth factor receptor 2 (HER2) is the only available biomarker for personalised treatment in metastatic GOC. Trastuzumab demonstrated a significant overall survival benefit when added to first-line chemotherapy5 and is now approved for patients with HER2 overexpression at immunohistochemistry (IHC) or gene amplification by in situ hybridisation (ISH) in equivocal cases.6

Globally, the rate of HER2 positivity in GOC ranges from less than 10% to almost 50%, on the basis of the samples used for analysis and the laboratory techniques applied.7 Screning data from the ToGA trial has shown that several clinical and pathologic variables are associated with a higher rate of HER2 positivity, such as the location of the primary tumour in the gastro-oesophageal tract and Lauren’s histotype, HER2 positive cases being more prevalent among upper lesions and in cases with intestinal histology.8 Moreover, the authors reported a slightly higher rate of HER2 positive status for biopsies compared with surgical specimens.8 More recently, a meta-analysis reported that male gender, well or moderately differentiated as well as intestinal histology were all factors associated with higher rates of HER2 positivity in GOC.9 Other authors recently reviewed the available literature data partially confirming previous reports, but also underlining that large heterogeneity exists about the association between single clinico-pathologic markers and HER2 status.10

Considering the relevance of determining HER2 status in metastatic GOC for optimal treatment selection, the limitations linked to sample adequacy and methodology (type of test) used and the need for rapid HER2 report, we aimed to develop and validate a nomogram, based on easily accessible clinical or pathologic characteristics, which is able to anticipate the probability of harbouring a HER2 positive disease before direct tumour assessment by IHC and ISH.
MATERIALS AND METHODS

Sample identification and data collection
Consecutive GOC cases, analysed from January 2011 to December 2017 for HER2 status by IHC and ISH at the Department of Surgical, Medical, Molecular Pathology and Critical Area of the University of Pisa (Italy), were used as developing cohort. Findings in this subset were then confirmed in a separate, independent validation cohort of GOC cases analysed at independent institutions in Italy since January 2011.

The following clinical and pathologic data were collected in both cohorts: gender (male versus female); primary tumour location (oesophago-gastric junction versus gastric body versus gastric fundus versus not specified); histologic subtype according to Lauren’s classification (intestinal versus diffuse versus not specified); tumour grading (G1 versus G2 versus G3 versus not specified); site of sampling (primary tumour versus metastasis); adequacy of pathologic material for analysis (inadequate material versus adequate material versus surgical specimen). Adequacy of biopsy for HER2 assessment was defined according to available evidence and recommendations, i.e. a number of at least 6 biopsies was considered as optimal for analysis.\textsuperscript{11–14}

Pathologic features collected in the analysed datasets (tumour grading and Lauren’s subtypes) were evaluated by pathologists with high expertise in GOC. With regard to tumour grading, the WHO 2010 classification has been applied.\textsuperscript{15} According to this scale, G1 (well-differentiated) tumours are composed of well-formed glands, sometimes resembling metaplastic intestinal epithelium, whereas G3 (poorly-differentiated) tumours are composed of glands with highly irregular and atypical features that are recognised with difficulty. G2 (moderately-differentiated) tumours are indeed represented by neoplasms with intermediate features between well- and poorly-differentiated carcinomas.

HER2 status assessment
HER2 status was initially assessed by IHC and in equivocal cases (i.e., 2+ at IHC) by ISH assays: HER2 positivity was then defined as IHC score of 3+ or 2+ with a positive ISH.

IHC was performed on formalin-fixed, paraffin-embedded tumour sections using the commercial antibodies PATHWAY antiHER-2/neu (4B5) Rabbit monoclonal (Roche-Ventana Medical Systems, Tucson, Az, USA). Sections were stained using automated slide stainer (Benchmark ULTRA, Roche Ventana Medical Systems, Tucson, Az, USA). Sections were stained using automated slide stainer (Benchmark ULTRA, Roche Ventana Medical Systems, Tucson, Az, USA). An appropriate scoring system,\textsuperscript{14} also assimilated the College of American Pathologists and regulatory authorities,\textsuperscript{11} exclusive for gastric tumours and accounting for at least 20 cancer cells. The positivity of FISH was considered at a HER2:chromosome 17 ratio of ≥2.0. The entire specimens were screened for amplified regions at an x20 magnification. In borderline amplification cases (ratio within the range 1.8–2.2), 20 additional cells were re-counted.\textsuperscript{11,14,16}

Statistical analyses
Four hundred and thirteen patients in the development cohort were used to build an unconstrained logistic regression model able to predict HER2 positivity. First, all the variables were tested in univariate models. All statistically significant variables were then used to build multivariable models. Variables not significant, but with a strong literature support in favour of an association with HER2 status, were also considered for inclusion. Both backwards and forward method were used. Collinearity was evaluated using Fisher’s test, t-test and ANOVA, depending on the nature of the covariates, and Variance Inflation Factor (VIF). Global fit was evaluated with Nagelkerke’s R², Somer’s D and model log-likelihood ratio chi-square. Final model was selected considering statistical significance of the covariates, the percentage of models in which it remained significant, their clinical plausibility and the global fit. A nomogram was then developed from the final model. Validation and calibration were performed on an external, independent dataset from three different Italian centres. C-index, visual inspection of the calibration plot, Brier score and Spiegelhalter z-test were used to assess the performance of the nomogram. The 95% confidence intervals of C-index were calculated with bootstrap method. A ROC curve was built with data from the validation set to assess sensitivity and specificity of the test at different cut-offs of predicted probabilities.

Package ‘rms’ and ‘pROC’ of R were used for all the analyses.

RESULTS

Sample characteristics
Characteristics of the samples included in the two datasets are listed in Table 1. A total of 723 cases were included in the study,

| Table 1. Study cohort’s characteristics | Development cohort \( (N=413) \) | Validation cohort \( (N=310) \) | \( P \)-value |
|----------------------------------------|--------------------------|--------------------------|--------------|
| HER2 status                            |                          |                          |              |
| Negative                               | 343 (83.1)               | 254 (81.9)               |              |
| Positive                               | 70 (16.9)                | 56 (18.1)                |              |
| Gender                                 |                          |                          |              |
| Male                                   | 262 (63.4)               | 210 (67.7)               |              |
| Female                                 | 151 (36.6)               | 100 (32.3)               |              |
| Type of material                       |                          |                          | \(<0.001\)^* |
| <6 biopsies                            | 99 (24)                  | 77 (24.8)                |              |
| ≥6 biopsies                            | 24 (5.8)                 | 82 (26.5)                |              |
| Surgical sample                        | 287 (69.5)               | 151 (48.7)               |              |
| Not specified                          | 3 (0.7)                  | 0 (0)                    |              |
| Site of sampling                       |                          |                          | \(0.077\)    |
| Primary tumour                         | 365 (88.4)               | 287 (92.6)               |              |
| Metastases                             | 48 (11.6)                | 23 (7.4)                 |              |
| Primary tumour location                |                          |                          | \(<0.001\)^* |
| O-G junction                           | 118 (28.6)               | 73 (23.5)                |              |
| Body                                   | 261 (63.2)               | 160 (51.6)               |              |
| Fundus                                 | 21 (5.1)                 | 52 (16.8)                |              |
| Not specified                          | 13 (3.1)                 | 25 (8.1)                 |              |
| Lauren’s histotype                     |                          |                          | \(0.559^*\)  |
| Intestinal                             | 193 (46.7)               | 103 (33.2)               |              |
| Diffuse                                | 183 (44.3)               | 108 (34.8)               |              |
| Not specified                          | 37 (9)                   | 99 (32)                  |              |
| Tumour grading                         |                          |                          | \(0.343^*\)  |
| G1                                     | 10 (2.4)                 | 11 (3.5)                 |              |
| G2                                     | 100 (24.2)               | 56 (18.1)                |              |
| G3                                     | 289 (70)                 | 178 (57.4)               |              |
| Not specified                          | 14 (3.2)                 | 65 (21)                  |              |

\(N\) number, O-G junction oesophago-gastric junction

\(^*\)Not specified cases were excluded from the comparison
with 413 cases in the development cohort and 310 cases in the validation cohort. There were 70 (16.9%) and 56 (18.1%) HER2 positive cases in the development and validation cohorts, respectively. We observed significant differences between the two cohorts regarding the pathologic material used for analysis (i.e., higher number of biopsies compared to surgical specimens in the validation cohort) and primary tumour location (i.e., higher percentage of cases located in the gastric fundus in the validation cohort), whereas no differences were present in respect to all other collected characteristics. Since mixed tumours represented a minority (<3% of all cases) in both cohorts, tumour samples were reviewed and labelled as either intestinal or diffuse type according to the prevalent histotype in each case.

Association of investigated variables with HER2 status and nomogram development

When tested for association with HER2 positive status, the following features were confirmed significant at multivariate analysis in the development cohort: tumour grading (analysed as continuous variable), histotype (diffuse versus intestinal) and site of sampling (primary versus metastases). Type of pathologic material was not significant but, given the amount of data about its correlation with HER2 status, we retained this parameter in the multivariable model. A detailed list of univariate and multivariate analyses is given in Table 2. Site of primary cancer was also significant at univariate analysis but was excluded from the final model due to an excessive amount of collinearity with grading and histologic subtype. Exclusion of this variable did not affect the global fit, which remained good: \( R^2 = 0.19 \), Somer’s D was 0.5, C-index was 0.75 and likelihood ratio chi-squared was 41.12 (\( P < 0.0001 \)). Notably, we detected collinearity between grade and Lauren’s subtype. However, the VIF

### Table 2. Association of clinical and pathologic factors with HER2 status: univariate and multivariate analyses

| Development cohort | Univariate OR (95%CI) | P-value | Multivariate OR (95%CI) | P-value |
|---------------------|-----------------------|---------|-------------------------|---------|
| Gender              |                       |         |                        |         |
| Female              | 1                     | –       | –                       | –       |
| Male                | 0.60 (0.34–1.06)      | 0.076   | –                       | –       |
| Type of material    |                       |         |                        |         |
| Surgical sample     |                       |         |                        |         |
| <6 biopsies         | 1.71 (0.96–3.05)      | 0.063   | 1.93 (0.91–4.11)        | 0.069   |
| ≥6 biopsies         | 2.47 (0.97–6.32)      | 0.058   | 2.29 (0.64–8.22)        | 0.201   |
| Site of sampling    |                       |         |                        |         |
| Primary tumour      | 1                     | –       | –                       | –       |
| Metastases          | 1.80 (1.03–4.15)      | 0.042   | 4.12 (1.11–15.2)        | 0.034   |
| Tumour location     |                       |         |                        |         |
| Body                | 1                     | –       | –                       | –       |
| O–G junction        | 1.88 (1.08–3.26)      | 0.023   | –                       | –       |
| Fundus              | 1.01 (0.28–3.59)      | 0.658   | –                       | –       |
| Histotype           |                       |         |                        |         |
| Intestinal          | 1                     | –       | –                       | –       |
| Diffuse             | 0.23 (0.12–0.44)      | <0.0001 | 0.41 (0.17–0.98)        | 0.045   |
| Tumour grading      |                       |         |                        |         |
| G1 versus G2        | 0.05 (0.02–0.16)      | <0.0001 | 0.09 (0.02–0.41)        | 0.0018  |

| OR (95%CI) | odds ratio (95% confidence interval), O–G junction oesophago-gastric junction |

HER2 nomogram: external validation

Probabilities predicted by the nomogram were tested against those observed in the validation set. The nomogram discriminative ability was very good, with a C-index of 0.827 (95%CI 0.741–0.913), Brier score was 0.101 and the Spiegelhalter Z-test was not significant (\( P = 0.319 \)). Visual inspection of the calibration plot showed a very good overlap between predicted and observed probabilities, with no relevant overestimations or underestimations (Fig. 2).

We developed a ROC curve on the validation cohort to calculate specificity and sensitivity of the test to detect the presence of HER2 positivity at different probabilities, as given by the model (Fig. 3). At a cut-off probability of 0.2, the model had 82% sensitivity and 74% specificity to detect the presence of HER2 positivity.

DISCUSSION

Defining HER2 status is crucial in the management of advanced GOC patients, as HER2 positive cases by IHC (and ISH, when required) may benefit from trastuzumab in combination with first-line chemotherapy.\(^5\),\(^6\),\(^7\) Missing HER2 positive tumours could then result in unexploited treatment opportunities, considering also that enrolment in clinical trials with novel anti-HER2 agents may be a suitable option. Therefore, in view of the relevance of this target (currently the only available validated predictive biomarker in GOC) and the challenges related to laboratory assessment of tumour samples in everyday practice,\(^18\) we developed and validated an easy-to-use and cheap nomogram, which predicts with high accuracy the chance of HER2 positivity before approved molecular diagnostics such as IHC and ISH are performed.\(^19\) Using a cut-off probability as given by the nomogram of 0.2, the test has a sensitivity of 0.82 and a specificity of 0.74 to detect the presence of HER2 positivity.

The variables included in the nomogram had been all associated with HER2 positive status, as it was previously found in available literature: lower tumour grading,\(^19\) intestinal histology,\(^19\),\(^20\) adequacy of analysed samples,\(^21\) and site of collection. Regarding the site of sampling (primary tumour versus metastases) larger heterogeneity is reported in different studies,\(^22\)–\(^25\) ultimately confirming that specific clinico-pathologic features are not able to predict HER2 status when considered separately.\(^16\) In our series, tumour grading was the strongest predictor of HER2 status, in both cohorts. Both pathologic material and site of sampling were included in the model, moving from the evidence that biopsies are associated with higher probability of HER2 positive status compared with surgical specimens in our series. Similar findings were also reported by the ToGA trial investigators,\(^3\) and may be justified by the different cut-off values used to define HER2 positivity. Another plausible explanation is that biopsies could be a surrogate variable, linked to deeper biological alterations associated with HER2 positivity. Again, there could be a potential overlap between biopsies and metastatic lesions, which are more likely to be HER2 positive. We indeed found in the developing set that a greater percentage of metastases are diagnosed by biopsies, introducing a slight collinearity in the model. However, the degree of such collinearity is small, unlikely to fully explain these findings. Our analyses also revealed a partial overlap between grading and Lauren’s subtype. However, the VIF
for each parameter was always lower than 3, strongly suggesting that the model is not affected by relevant collinearity.

Adequate pathologic material for analyses is the key element of any quality-controlled laboratory procedure. Considering the heterogeneity in HER2 expression across tumour cells in GOC, a minimum of 5 biopsies are required according to literature data, with 6 to 8 specimens considered as the optimal threshold for adequate and reliable HER2 assessment in GOC by available recommendations. So, we set 6 as the reference number to categorise biopsy samples adequacy. Unfortunately, this issue has not been fully implemented in routine practice and our study confirms that a significant percentage of biopsies used for HER2 status assessment should have been considered indeed inadequate for appropriate evaluation. In our opinion, this nomogram could prompt clinicians to perform tumour re-biopsy in initially HER2 negative cases at first assessment, in order to immediately evaluate different treatment options or study proposals in patients with higher probability of HER2 positive disease. This could also allow not delaying treatment initiation in patients with very low pre-test probability of harbouring a HER2 positive tumour. These patients could benefit from alternative chemotherapy regimens compared to the cisplatin plus fluoropyrimidine combination used in ToGA: reasonable options are represented by a docetaxel-containing triplet, in order to increase activity and potentially efficacy, or oxaliplatin-based doublets, in order to improve safety.
To conclude, HER2 status might be accurately and rapidly predicted by a simple nomogram based on four validated clinico-pathologic parameters. This tool could be easily implemented in clinical practice during the first assessment of the patient to add information for prompt case management.

AUTHOR CONTRIBUTIONS
Study concept and design: L.F. and C.V.; acquisition of data: all authors; analysis and interpretation of data: L.F., C.V., C.U. and F.M.; drafting of the manuscript: L.F., C.V., C.U., G.A., I.P., E.V. and F.M.; revision the manuscript: all authors; approval the manuscript: all authors.

ADDITIONAL INFORMATION
Competing interests: The authors declare no competing interests.

Ethics approval and consent to participate: The analyses included in this study were performed in accordance with the Declaration of Helsinki and were approved by the Ethics Committee of the Pisa University Hospital. Written informed consent from the patients for research use of data was obtained before the investigation.

Data availability: All data generated or analysed during this study are included in this published article.

Note: This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution 4.0 International (CC BY 4.0).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES
1. Ferlay, J. et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int. J. Cancer 136, E359–E386 (2015).
2. Buas, M. F. & Vaughan, T. L. Epidemiology and risk factors for gastroesophageal junction tumors: understanding the rising incidence of this disease. Semin. Radiat. Oncol. 23, 3–9 (2013).
3. Anderson, W. F. et al. Age-specific trends in incidence of noncardia gastric cancer in US adults. JAMA 303, 1723–1728 (2010).
4. Lordick, F., Shiitara, K. & Janjigian, Y. Y. New agents on the horizon in gastric cancer. Curr. Opin. Oncol. 35, 446–464 (2017).
5. Chua, T. C. & Merritt, N. D. Clinopathologic factors associated with HER2-positive gastric cancer and its impact on survival outcomes—an systematic review. Int. J. Cancer 130, 2845–2856 (2012).
6. Bartley, A. N. et al. HER2 testing and clinical decision making in gastroesophageal adenocarcinoma: guideline of the College of American Pathologists, American Society for Clinical Pathology, and the American Society of Clinical Oncology. J. Clin. Oncol. 35, 446–464 (2017).
7. Gravalos, C. & Jimeno, A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. Ann. Oncol. 19, 1523–1529 (2008).
8. Van Cutsem, E. et al. HER2 screening data from ToGA: targeting HER2 in gastric and gastroesophageal junction cancer. Gastric Cancer 18, 476–484 (2015).
9. Wang, H. B., Liao, X. F. & Zhang. J. Clinopathological factors associated with HER2-positive gastric cancer: a meta-analysis. Medicine 96, e6437 (2017).
10. Chua, T. C. & Merritt, N. D. Clinopathologic factors associated with HER2-positive gastric cancer and its impact on survival outcomes—a systematic review. Int. J. Cancer 130, 2845–2856 (2012).
11. Bartley, A. N. et al. HER2 testing and clinical decision making in gastroesophageal adenocarcinoma: guideline of the College of American Pathologists, American Society for Clinical Pathology, and the American Society of Clinical Oncology. J. Clin. Oncol. 35, 446–464 (2017).
12. Gullo, I. et al. Minimum biopsy set for HER2 evaluation in gastric and gastroesophageal junction cancer. Endosc. Int. Open 3, E165–E170 (2015).
13. Lordick, F. et al. HER2 testing in gastric cancer: results of a German expert meeting. J. Cancer Res. Clin. Oncol. 143, 835–841 (2017).
14. Hofmann, M. et al. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. Histopathology 52, 797–805 (2008).
15. Bokan, F. T., Carneiro, F., Hruban, R. H., These, N. D. (eds.). WHO Classification of Tumours of The Digestive System—WHO Classification of Tumours 4th edn, Vol. 3 (World Health Organization, Geneva, Switzerland, 2010).
16. Ruschoff, J. et al. HER2 testing in gastric cancer: a practical approach. Mod. Pathol. 25, 637–650 (2012).
17. Smyth, E. C. et al. Gastric cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann. Oncol. 27, v38–v49 (2016).
18. Salto-Tellez, M., James, J. A. & Hamilton, P. W. Molecular pathology—the value of an integrative approach. Mol. Oncol. 8, 1163–1168 (2014).
19. Shah, L., Ying, J. & Lu, N. HER2 expression and relevant clinopathological features in gastric and gastroesophageal junction adenocarcinoma in a Chinese population. Diagn. Pathol. 8, 76 (2013).
20. Matsuoka, S. et al. Clinopathological factors associated with HER2 status in gastric cancer: results from a prospective multicenter observational cohort study in a Japanese population (JFMC44-1101). Gastric Cancer 19, 839–851 (2016).
21. Ahn, S. et al. Ideal number of biopsy tumor fragments for predicting HER2 status in gastric carcinoma resection specimens. Oncotarget 6, 38372–38380 (2015).
22. Amato, M. et al. HER2 status in gastric cancer: comparison between primary and distant metastatic disease. Pathol. Oncol. Res. 23, 55–61 (2017).
23. Salto, T. et al. Preferential HER2 expression in liver metastases and EGFR expression in peritoneal metastases in patients with advanced gastric cancer. Gastric Cancer 18, 711–719 (2015).
24. Boffetti, C. et al. Comparison of HER2 status in primary and paired metastatic sites of gastric carcinoma. Br. J. Cancer 104, 1372–1376 (2011).
25. Wang, D. D., Kumarasinghe, M. P., Platten, M. A. & de Boer, W. B. Concordance of HER2 expression in paired primary and metastatic sites of gastric and gastroesophageal junction cancers. Pathology 47, 641–646 (2015).
26. Warmke, V. S. et al. Her2/neu testing in gastric cancer: evaluating the risk of sampling errors. Ann. Oncol. 24, 725–733 (2013).
27. Al-Batran, S. E. et al. Effect of neoadjuvant chemotherapy followed by surgical resection on survival in patients with limited metastatic gastric or gastroesophageal junction cancer: the AIO-FLOT3 trial. JAMA Oncol. 3, 1237–1244 (2017).
28. Ter Veer, E. et al. The efficacy and safety of first-line chemotherapy in advanced esophagogastric cancer: a network meta-analysis. J. Natl. Cancer Inst. 108, djw166 (2016). https://doi.org/10.1093/jnci/djw166