Neuro
endocrinology

DOI: 10.1159/000509091
Received: 4/9/2020
Accepted: 6/2/2020
Published(online): 6/3/2020

Blood Molecular Genomic analysis predicts the disease course of GEP NET patients: a validation study of the predictive value of the NETest®
van Treijen M.J.C. van der Zee D. Heeres B.C. Staal F.C.R. Vriens M.R. Saveur L.J. Verbeek W.H.M. Korse C.M. Maas M. Valk G.D. Tesselaar M.

ISSN: 0028-3835 (Print), eISSN: 1423-0194 (Online)
https://www.karger.com/NEN
Neuroendocrinology

Neuroendocrinology

Manuscript:	NEN-2020-4-11/R1 RESUBMISSION
Title:	Blood Molecular Genomic analysis predicts the disease course of GEP NET patients: a validation study of the predictive value of the NETest®
Authors(s):	Mark J. C. van Treijen (Corresponding Author), Dennis van der Zee (Co-author), Birthe C Heeres (Co-author), Femke C. R. Staal (Co-author), Menno R. Vriens (Co-author), Lisette J. Saveur (Co-author), Wieke H. M. Verbeek (Co-author), Catharina M Korse (Co-author), Monique Maas (Co-author), Gerlof D. Valk (Co-author), Margot E. T. Tesselaar (Co-author)
Keywords:	Biomarkers, Chromogranin A, Gastroenteropancreatic neuroendocrine tumors, Liquid biopsy, Survival
Type:	Research Article

Blood Molecular Genomic analysis predicts the disease course of GEP NET patients: a validation study of the predictive value of the NETest®

M.J.C. van Treijen[1,2], D. van der Zee[3], B.C. Heeres[2,4], F.C.R. Staal[2,4], M.R. Vriens[2,5], L.J. Saveur[2,6], W.H.M. Verbeek[2,6], C.M. Korse[2,7], M. Maas[2,4], G.D. Valk[1,2*] en M.E.T. Tesselaar[2,8*].

1. Department of Endocrine Oncology, University Medical Center Utrecht, Utrecht, Netherlands

2. Center for Neuroendocrine Tumors, ENETs Center of Excellence, Netherlands Cancer Institute, University Medical Center Utrecht, Utrecht, Netherlands

3. Department of Radiology, Bernhoven Hospital, Uden, the Netherlands

4. Department of Radiology, The Netherlands Cancer Institute, Amsterdam, Netherlands

5. Department of Endocrine Surgical Oncology, University Medical Center Utrecht, Utrecht, Netherlands

6. Department of Gastroenterology, The Netherlands Cancer Institute, Amsterdam, Netherlands

7. Department of Clinical Chemistry, The Netherlands Cancer Institute, Amsterdam, Netherlands

8. Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, Netherlands

** These authors share senior authorship*

Corresponding author:

Mark van Treijen, MD

UMC Utrecht cancer center – department of endocrine oncology

Heidelberglaan 100, 3584CX Utrecht, the Netherlands

M.j.c.vantreijen@umcutrecht.nl

Keywords: Gastroenteropancreatic neuroendocrine tumors, biomarkers, liquid biopsy, survival, chromogranin A.

Word count: 4962

Abstract (250 words)

Reliable prediction of disease status is a major challenge in managing gastroenteropancreatic neuroendocrine tumors (GEP-NET).

The aim of the study was to validate the NETest®, a blood molecular genomic analysis, for predicting the course of disease in individual patients compared to Chromogranin A (CgA).

NETest® (normal ≤20%) and CgA (normal <100 µg/l) were measured in 152 GEP-NETs. Median follow-up was 36 [4-56] months. Progression free survival (PFS) was blindly assessed (RECIST 1.1). Optimal cutoffs (area-under-the-ROC curve (AUC)), odds ratios (OR), negative and positive predictive values (NPV/PPV) were calculated for predicting stable (SD) and progressive disease (PD).

Of the 152 GEP-NETs, 86% were NETest®-positive and 52% CgA-positive. NETest® AUC was 0.78 vs CgA 0.73 (p=NS). The optimal cut-offs for predicting SD/PD were 33% for the NETest® and 140 µg/l for CgA. Multivariate analyses identified NETest® as the strongest predictor for PD (OR: 5.7 [score: 34-79%]; 12.6 [score≥80%]) compared to CgA (OR 3.0), tumor grade (OR 3.1) or liver metastasis (OR 7.7). NETest® NPV for SD was 87% at 12 months. The PPV for PD were 47% and 64% (scores 34-79% and ≥80%, respectively). NETest® metrics were comparable in watchful waiting-, treatment- and no evidence of disease (NED) subgroups. For CgA (>140ng/ml), NPV and PPV were 83% and 52%. CgA could not predict PD in watchful waiting or NED subgroups.

The NETest® reliably predicted SD and was the strongest predictor of PD. CgA had lower utility. The NETest® anticipates RECIST defined disease status up to one year before imaging alterations are apparent.

Introduction

Overall- and progression free survival rates diverge widely between the different sub-types of gastroenteropancreatic neuroendocrine tumors (GEP-NETs) with type of tumor, tumor grade and stage as independent predictors for tumor progression[1–3]. Despite these parameters, it remains very difficult for clinicians to predict the clinical course in an individual patient[4–6]. Clinical management decisions are often driven by combining the features of the tumor such as grade and stage with the course of the disease as assessed by radiologic exams. Therefore, even several years after the initial diagnosis, in many patients, clinical decision making is based on the original pathological examination of a small tissue sample that no longer represents the current biological status of the heterogeneous and polyclonal tumor that has evolved with time and as a consequence of treatment.

In patients with local or locoregional disease, surgery remains the fundamental component of all management strategies. However, even after surgery with curative intent, post-operative surveillance is still necessary for many years to exclude residual or metastatic disease[6], with current techniques confined to imaging and Chromogranin A (CgA) measurement. Nevertheless, a significant proportion of GEP-NETs are metastatic at diagnosis[2,7] and management strategies in these tumors predominantly focus on symptom control and inhibition of tumor growth[8]. In non-functional GEP-NETs several guidelines consider a watchful waiting strategy after diagnosis as appropriate to enable estimation of the propensity for growth. In the event of progressive disease, different therapeutic modalities are available to regain tumor growth control and enable maximal progression free survival (PFS)[9,10].

It has thus become self-evident that continual assessment of the disease status remains the fundamental basis of the management of GEP-NET disease [6]. Up to now, a combination of symptomatology, functional and anatomical imaging and biomarkers is utilized. Despite this multimodal assessment strategy there are well documented limitations for each parameter[11]. Current biomarkers are considered insufficient for providing accurate reproducible information in respect of the aggressive and proliferative capacity of an individual tumor [12]. Nevertheless, CgA is used both as prognostic marker at diagnosis and as marker for disease progression or disease recurrence during surveillance [13]. Although CgA correlates with tumor burden[14] reports on the ability to predict the course of the disease are equivocal[15–19].

Therefore, over recent years, research in GEP-NET disease as in other oncological disciplines has focused on the development of alternative tools that delineate the biological characteristics of this heterogeneous group of tumors[20,21]. In particular, it is now recognized that multianalyte assessment of tumor biology is more effective than monoanalyte evaluation of membrane antigens (PSA or CEA) or secretory products such as serotonin or CgA[12]. Circulating molecular information from GEP-NETs (circulating tumor DNA or -cells, and mRNA) can possibly be used as a liquid biopsy to provide information on individual tumor behavior and prediction of the clinical course. With such real time information the management and treatment of GEP-NET could directly be adapted to the individual patient's needs.

One of the emerging biomarkers in GEP-NETs, is the NETest®. This test is a multianalyte algorithmic analysis intending to provide the biological signature of an individual tumor, quantified by a 'NET activity score'. This score is based on the gene expression of 51 marker

genes and the differential analyses of specific gene clusters (omes) which differentiates stable disease (SD) from progressive disease (PD). Available data on the NETest® different applications and its clinical utility has recently been systematically reviewed and analyzed[22]. In this review it was described that the NETest® is diagnostic and appears to have clinical utility in monitoring therapeutic efficacy. Therefore, the authors concluded that the NETest® has a significant advantage over CgA. Currently, only three previous studies illustrated the utility to predict the natural course of disease in GEP-NETs. These studies all had methodological shortcomings. One study with a long-term follow-up assessed the utility only in a small group of patients (n=34) [23]. The other two studies included different types of NETs and had short median follow up of only 6 and 8 months[24,25]. Moreover, in one of these studies, clinicians could use the NETest at their discretion for clinical management[24]. Although encouraging, these results require validation before the clinical utility for predicting the course of disease in individual patients can be judged.

In order to specifically address the clinical utility of the NETest® and compare it to CgA we investigated the two biomarkers in a well-defined large prospective cohort of patients with well differentiated GEP-NET with long-term follow up. We assessed the effectiveness for prediction of PFS, identification of disease recurrence and all-cause mortality in individual GEP-NET patients.

Methods

Consecutive patients with histologically proven, well-differentiated sporadic GEP-NETs were approached for inclusion between March 2014 and March 2017 at the Netherlands Cancer Institute (NCI) (Amsterdam, The Netherlands), ENETS Center of Excellence. At inclusion, central standardized pathology review was performed for all patients. All NETs were graded according to the World Health Organization (WHO) 2017 grading system[26].

At inclusion, samples (6ml EDTA-collected whole blood) were thoroughly mixed and immediately stored on ice. Samples were stored at -80°C within 2 hours after collection according to standard molecular diagnostics protocols for PCR-based studies[27]. Baseline samples for NETest® assessment were sent in different anonymized batches to Wren Laboratories, Connecticut, USA from October 2015 – October 2018. Samples were always drawn in combination with CgA and radiological imaging studies. Patients were followed in a standardized manner according to the ENETS guidelines. Study design and analysis plan were defined and agreed upon before the start of the study. Ethics committee approval (NCI, Amsterdam) was obtained and all patients signed informed consent.

Biomarkers

Details of the PCR methodology, mathematical analysis, and validation have previously been described comprising a 2-step protocol (RNA isolation/cDNA production and q-PCR)[28–31]. Target transcript levels are subsequently normalized and quantified versus a historical (2014) population control[29]. NETest® outcomes are expressed as an activity index from 0-100%[28]. NETest outcomes are classified in different categories. The upper limit of normal (ULN) has

previously been set at 20%[14], stable disease (SD) is defined as ≤ 40% and PD as an activity score > 40% with 41-79% as intermediate tumor activity and scores ≥ 80% as high tumor activity [23,24].

CgA was measured with B·R·A·H·M·S Chromogranin A, an automated immunofluorescent assay for the quantitative determination of CgA in human serum using the KRYPTOR instrument (BRAHMS GmbH, Hennigsdorf, Germany). The upper limit of normal (ULN) is established as 100 µg/l. CgA levels were determined at the NCI (Clinical Laboratory).

All samples were anonymized and coded, and laboratory investigators at both sites were blinded to the clinical diagnosis, and disease status.

Disease status

Disease status at entry and follow up – the primary outcome – was evaluated at consecutive imaging according to the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1[32]. All imaging studies were re-assessed in a standardized manner by two independent senior radiologists who were blinded to the biomarker results. Both radiologists were equally expert in the different imaging modalities involved in this study. Patients with a minimum follow up of six months and a minimum of two consecutive imaging modalities, appropriate to reliably measure disease status were included. According to protocol, patients underwent anatomical imaging every 3-12 months, alternated with functional imaging once every 1-2 years, depending on their clinical condition and response to treatment. The preferred anatomical imaging for the assessment of the outcome measure was computed tomography (CT) if multiple imaging

modalities were available in the same surveillance period. Magnetic resonance imaging (MRI) or ultrasound was used in some individuals due to patient- or tumor characteristics. Ultrasound was only used in some accessible patients (n=4) who underwent curative surgery as surveillance for recurrence or liver metastasis. Ultrasound was always alternated with MRI and/or functional imaging. Outcomes of functional imaging (^{68}Gallium-DOTATATE PET with low dose CT (DOTA PET CT)) were used in cases where conventional radiological imaging was not available. Since sensitivity of DOTA PET CT is superior compared to conventional imaging modalities, new lesions on the first DOTA PET CT after previous conventional imaging were not taken into consideration in the determination of the disease status. New lesions identified on conventional imaging had to be confirmed as present on consecutive imaging.

Patients were considered to have measurable disease if a tumor lesion was visualized on consecutive imaging modalities. No evidence of disease (NED) was defined as negative consecutive imaging (minimal 2) after surgery with curative intent.

Analysis

Statistical analyses were performed using Statistical package for Social Sciences (SPSS) 25. Statistical significance was defined at a p value ≤ 0.05. To describe clinical characteristics, NETest® scores and CgA levels, the mean ± SD or median with range were calculated in normal distributed and non-normal distributed data respectively (Kolmogorov-Smirnov; K-S).

Only blood samples collected at baseline were used in this study. The utility of both biomarkers to predict PFS according to RECIST 1.1 was the primary outcome of this study. PFS was

calculated as the time between the baseline measurement and the first date patients were considered to have PD. Baseline imaging was compared to previous imaging procedures (if available), to estimate the disease status (SD of PD) at inclusion to accurately estimate time to progression. Predictive values are described by area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the originally described cut-offs of both tests and the optimal cut-offs for both tests. Optimal cut-off for both biomarkers were assessed by using the AUC. The McNemar test was used to compare the NETest with CgA. Kaplan Meier analysis and log-rank test for PFS were performed to estimate differences in PFS between the cut-off points.

Secondary outcomes were all-cause mortality and recurrence of disease after intended curative surgery. Spearman correlation was used to assess the correlation between biomarkers and outcome measures. Univariate analyses were performed to identify predictors for tumor progression within 12 months of follow up. Identified covariates for PD in literature were included[4]. Significant parameters were included stepwise in a multivariate logistic regression analysis.

Results

A total of 152 out of 176 patients was eligible for inclusion in this study. Twelve patients were lost to follow up or referred back to their referral hospital within 6 months. Seven patients had metastasized disease that could not be used for the primary outcome (e.g. peritoneal metastases). One patient was excluded because of a metastasized second malignancy and four

patients were excluded because of curative surgery (3) or PRRT (1) shortly after baseline and therefore were considered to have an "unnatural alteration" in the course of disease. An overview of the population and different subgroups at baseline is illustrated in **Figure 1**.

Baseline characteristics of the included patients are described in **Table 1**. Median follow up was 36 months [4-56 months]. Median NETest® score was 33%. NETest® was positive (> 20%) in 92% of all patients with measurable disease and in 76% with NED. Median NETest in patients with NED was 27% (7-100%) and 33% (13-93%) in patients with measurable disease ($p < 0.01$). Median CgA was 107 µg/l. CgA was positive (> 100 µg/l) in 58% of all patients with measurable disease and in 30% with NED. Median CgA in patients with NED was 71 µg/l (19-798 µg/l) compared to 146 µg/l (12-44150 µg/l) in patients with measurable disease ($p = 0.001$)

The predictive value for progression free survival

Disease progression was identified in 17%, 32%, 38% and 45% of all included patients after 6, 12, 18 and 24 months of follow up, respectively.

Figure 2 shows the distribution of NETest® and CgA in those with- or without progression within the first year after baseline. The highest accuracy for the NETest® to predict disease status was demonstrated at 12 months of follow up. The AUC for predicting disease status (SD vs PD) up to12 months from baseline was 0.78 (95% CI 0.70-0.86) for the NETest® and 0.73 for CgA (95% CI: 0.64-0.83; p = NS; **Figure 3**). Of the 101 patients who were considered to have SD at this time-interval, 74% had a NETest® score ≤40% compared to 57% for CgA (ULN: 100 µg/l) ($p< 0.01$). Of the patients with PD, 68% had an elevated NETest® and 70% had an elevated CgA

outcome (p= NS). Median NETest in SD group was 27% versus 73% in PD group (p< 0.001). Median CgA was 78 µg/L versus 483 µg/L respectively (p< 0.001). The PFS for the previously established NETest® categories (≤40% (low tumor activity), 41-79% (intermediate tumor activity) and ≥80% (high tumor activity)) and CgA (ULN: 100 µg/l) are illustrated in **Supplementary Figure 1**. A significant difference was observed between the low- and high activity NETest® categories (p < 0.001). The PPV for intermediate – and high tumor activity categories was 44% and 64% respectively. The NPV was 83%. The PFS of patients was also significantly different between normal and elevated CgA levels (p = 0.04), with a PPV and NPV of 43% and 80%, respectively (**Table 2)**. No difference in AUC was observed in the subgroups of patients with pancreatic NETs and small intestine NETs.

Optimal cutoff

The highest accuracy for the NETest® to predict PD was demonstrated at 12 months of follow-up with an activity scores >33% as optimal cut-off (combining the optimal sensitivity and specificity). Using a low activity category of 0-33%, an intermediate activity category of 34-79% and a high risk category of ≥ 80%, PD was observed in 13%, 47% and 64% after 12 months of follow up, respectively. For 24 months of follow up, this was 24%, 54% and 79%, respectively. **Figure 4a** demonstrates the course of disease in relation to these NETest® categories. The NETest® categories significantly differentiated in mPFS: 55 months compared to 18 months and 11 months respectively (p <0.001; intermediate - high: p = 0.08).

The recalculated optimal cut-off for CgA to predict progression at +12 months of follow-up was 140 µg/L. **Figure 4b** illustrates the mPFS for this CgA cut-off. An elevated CgA (ULN 140 µg/L) predicted PD in 52 % of patients at + 12 months of follow up compared to 17% of patients with a CgA outcome below 140µg/L. For 24 months of follow up, this was 59% and 31%, respectively. Median PFS was 55 months versus 12 months, respectively ($p < 0.001$).

Metrics (original and optimal cut-offs) for NETest® and CgA to predict SD and PD at 12 months are shown in **Table 2**. The NETest had overall better metrics compared to CgA. Patients with a NETest® outcome > 33% had almost nine times higher chance for PD compared to those with an outcome ≤ 33% (OR 8.6). Patients with an optimized CgA outcome ≥ 140µg/L had a 5.2 times higher chance for progressive disease compared to those with a lower outcome.

Predictors for disease progression

In multivariate analysis, the NETest®, CgA, tumor grade and presence of liver metastases were independent predictors for PD. The model explained 58% (Nagelkerke R^2) of the variance in disease progression and correctly classified 82% of the cases. The NETest® was the strongest predictor. Intermediate scores (34 – 79%) were associated with 5.7 [CI 95%: 1.7-18.5] times increased likelihood for patients to develop tumor progression. High scores (≥80%) increased the risk of tumor progression 12.6 [CI 95%: 3.7-43.1] fold. Tumor progression was 3.0 [CI 95%: 1.3-6.9] times more likely for every 10-fold elevation of CgA. Patients with grade 2 tumors were 3.1 [CI 95%: 1.0-9.5] times more likely to progress within 1 year compared to grade 1. Patients

with liver metastases were 7.7 [CI 95%: 1.6-37.4] times more likely to progress compared to patients with non-liver metastases. There was no predictive association with age or gender.

Combination CgA and NETest®.

Figure 5 demonstrates the cumulative PFS when the outcomes of the NETest® and CgA were combined. When both tests are below the optimal cutoff level (NETest®: ≤33%, CgA: ≤140µg/L), a large proportion of patients remained stable over a long period of time (log-rank test p=0.02). The NPV was 96% [95% CI: 87-99]. The PPV for PD was 69% [95% CI: 56-79].

Watchful waiting strategy versus treatment in patients with measurable disease

Fifty-five patients with measurable disease had no treatment at inclusion, the 'watchful waiting group'. Fifty of those patients (91%) had a positive NETest (>20%) and 33 patients (61%) had a positive CgA (>100µg/l; p = 0.001). Thirty-two percent of patients in the watchful waiting cohort developed PD within one year after inclusion. Of the patients with low NETest® scores (≤33%) a total of only 16% had PD in the first 12 months of follow up, compared to 50% and 54% in the intermediate (34-79%) and high (≥ 80%) activity category (p = 0.02; intermediate - high: NS). This significant difference between the survival curves was sustained during the entire follow up period (**Figure 6**). The AUC of the NETest® in this subgroup was 0.70 [0.55-0.85] and the cut-off combining the optimal sensitivity and specificity was 33%. The NPV and PPV were also calculated for 24 months: 70% had still stable disease in the NETest low activity category at 24

months. Patients with intermediate- and high activity scores showed disease progression in 50% and 74% at two years of follow up.

CgA failed to predict the course of disease in the watchful waiting subgroup. The AUC for CgA was 0.64 [0.47-0.82] and the optimal cut-off was 140µg/L. Progressive disease was observed in 21% with low CgA outcomes (≤ 140 µg/L), compared to 41% with elevated CgA (p=NS). Differences decreased after two year. The proportions of patients with PD were 37% and 51% for low and high CgA outcomes, respectively after two years of follow up and proportions cumulative PFS converged after 28 months (**Supplementary Figure 2a**).

Sixty-four patients were on treatment at baseline (**Table 1**). In this group, the NETest was positive (>20%) in 56 patients (88%) compared to 36 patients (56%) with positive CgA (>100µg/l; p< 0.001). PD was observed in 45% within 12 months of follow up. Progressive disease at 12 months of follow up was observed in 17% of the patients with NETest® low activity scores (≤33%). This was a significantly lower proportion compared to the intermediate category (PD: 61%; p <0.001) and high tumor activity (PD: 74%, intermediate-high: NS: **Figure 6**). The AUC for the NETest® in the treatment group was 0.83 [0.73-0.93]. At 24 months, 64% of patients with low activity scores exhibited stable disease. The PPV for intermediate- and high tumor activity categories was 69% and 92% at this time-interval.

A significant difference was also observed between normal CgA levels (26% PD) (ULN 140 µg/L) and an elevated CgA (64%PD; p = 0.03; **Supplementary Figure 2b**). AUC for CgA in this subgroup was 0.76 [0.64-0.88]. PPV and NPV for each subgroup are illustrated in **Table 3**.

Patients with no evidence of disease

Thirty-three patients were considered to have no evidence of disease (NED) at baseline. In 88% of those patients this was based on a combination of anatomical and functional imaging and in the remaining 12% on a combination of different types of anatomical imaging. Median follow up in this subgroup of patients was 38 months [12-56 months]. Six patients (18%) developed metastases or recurrence of disease. The median NETest® score in patients who still exhibited NED at follow up was 27% compared with 53% in patients with recurrence (p =0.07). In this patient group, a low NETest® activity score (≤33%) had a high negative predictive value (88%; [95% CI: 76-94%]. A high disease activity score (scores > 33%; n=8) had a PPV of 38% [95% CI: 16-65%]) for disease recurrence. **Figure 7a** illustrates recurrence over time in the 33 patients with NED at baseline, as assessed by NETest® (cut-off 33%; p = 0.032). No patient with a negative NETest® score (≤ 20%) (n=8) had recurrence of disease during follow up.

CgA could not differentiate between recurrence or continued NED. Median CgA outcome was 60 µg/L versus 75 µg/L (p=0.46), respectively. Equivalent proportions: 18% (≤140 µg/L) and 20% (>140 µg/L) exhibited recurrence of disease (p=0.97; **Figure 7b**). An overview of test performances in each subgroup is presented in **Table 3.**

Mortality

Thirty-one patients died during follow up. Fifteen patients died within two years. Patients with elevated NETest® scores (>33) had a minor but significant lower cumulative survival (p=0.02).

Differences in all-cause mortality remained significant between the groups when only the first two years were analyzed with 6% (5 of 86) (NETest score® ≤33%) versus 15% (10 of 66) (NETest score® > 33%) being deceased (p = 0.05). CgA proved to be a stronger predictor. Only 2.4% (2 of 84) of all patients with negative CgA died within two years compared to 19% (13 of 67) in patients with elevated CgA (>140 µg/L; p= 0.01)

Discussion

In this independent and largest prospective cohort study to date, a low NETest® – a multigene-based blood test measuring circulating transcripts – proved to reliably predict long-term stable disease in GEP-NET patients. The NETest predicts RECIST defined disease status up to 1 year before this is apparent on imaging with a predictive accuracy of 78%. Patients with a low NETest® score (≤33%) had an 87% and 75% chance of stable disease at 12 months and even 24 months of follow up, respectively. In addition, there was a clear difference in the course of disease between patients with low- and higher scores even for over two years after baseline. Comparable results were evident in subgroups of patients who were following a watchful waiting strategy (NPV 84%) or were on treatment (NPV 83%). In line with earlier reports, in patients who underwent surgery with curative intent (n= 33), very low NETest® outcomes (≤ 20%) reliably predicted no recurrence of disease in years of follow up [33,34]. We also noted that low activity NETest® scores (≤33%) were associated with a significantly longer time to recurrence compared to NETest® > 33%. These results illustrate that the NETest® can be used as a 'rule-out' biomarker to provide assessment of surgical efficacy. Very low NETest outcomes

(≤20%) could even replace other currently used measures of disease status like CgA and possibly even imaging. However, the subgroup of patients with very low NETest outcomes was too small (n= 8) for drawing firm conclusions.

Furthermore, in multivariate analysis the NETest® was identified as the strongest predictor of disease course, with an almost six and 13 times higher chance of disease progression in patients with an intermediate (34-79%) or high (≥ 80%) NETest® outcome, respectively. Although the AUC of CgA (0.73) was comparable to the NETest® (0.78), CgA was unable to predict the course of disease in the watchful waiting subgroup and could not predict recurrence in the NED subgroup.

The present study was set up in a manner to limit potential bias. All eligible consecutive patients with GEP-NETs were recruited for inclusion and therefore represent the population of interest. Furthermore, all patients were followed according to protocol. The disease status of the patients - primary outcome - was re-assessed (blinded/anonymized fashion) for this study by independent radiologists using a pre-defined protocol. The NETest® was performed in the laboratory without any knowledge of the disease status of patients and clinicians and radiologists were unaware of NETest® results. As a result of this study set up, we created a robust and independent evidence base for the predictive ability of the NETest® for individual patients with GEP-NETs encountered in daily clinical practice. The unique prospective long term follow up leads to new insights into the predictive value even after 24 months.

A recent meta-analysis by Öberg *et al*. reported a median accuracy of the NETest® to reflect disease status to be 85%[22]. This review however focused on actual disease status at the time

of blood draw and not on predicting the course of disease over time which was our goal. Therefore, the outcomes are not comparable to the results of the current study.

We are aware of only three studies that assessed the utility of the NETest to predict the course of the disease in GEP-NET patients[23–25]. The PPV in the high (≥ 80%) NETest® activity outcome group varies between studies. In our study, 64% of all patients with high tumor activity scores were progressive within one year and even 79% at 24 months. Two of the three earlier studies reported comparable PPV's. Pavel *et al* reported a PPV of approximately 70%, 1 year from baseline in 31 patients. Malczewska *et al* calculated a PPV of 70% at a shorter median follow up of 8 months [25]. In contrast, in a previous US-registry based study, also with a shorter period of follow up (median 6 months), PPV was 81% [24]. In this particular study, since it was real-life format, NETest® outcomes could be used at the discretion of clinicians and symptomatology of patients were part of the primary outcome. The variations in outcomes of the individual studies may reflect the different approaches.

The high NPV in our study (87%) is consistent with the calculated values in previous studies[23–25]. A biomarker with a high NPV can be used to alter management strategies, such as imaging frequency or initiation of therapy. However, the predictive value should be well above 90% to ensure that only a very minimal proportion of patients are misclassified. Despite the high NPV found in this study, an individual patient with a low NETest® outcome still had a 13% chance of progressive disease at 12 months of follow up in our population. It is debatable if this is acceptable when changes in management, for example a reduction in imaging frequency, are considered based on low NETest® scores, but it certainly remains an attractive possibility.

In our study, the combination of a low NETest® outcome and negative CgA had an excellent NPV of 96%. Lowering surveillance frequency and refraining from expensive treatment options such as somatostatin analogues, PRRT or Everolimus in this patient group would readily be envisaged to result in lower health-care costs. However, since this was a post-hoc analysis this could be a chance finding. Additionally, the combination of both biomarkers is only useful when both biomarkers are positive or negative. With different analytical performances of CgA assays, these results are difficult to validate and therefore probably have limited clinical application.

In our study CgA performed better when compared to other studies evaluating both biomarkers [23,24]. This might be explained by the standardized work-up and processing of the samples. Since the accuracy of CgA is highly dependent on the used assay, our results on the accuracy of CgA cannot be extrapolated to the general population [35,36]. Furthermore, to evaluate both biomarkers identically, we also calculated the optimal cutoff for CgA. This results in an overestimation of the predictive value compared to the original cutoff and results are therefore not transferable to other CgA assays and institutions. Despite using the optimal cutoff, CgA results were still contradictory. CgA was positive in only 52% of patients using the standard cut-off of 100µg/L. Increasing this to 140µg/L was associated with an even lower positive rate of 44%. False-negative outcomes therefore remain a critical limitation since CgA could not be used in these patients (with measurable disease) as a biomarker that would provide relevant clinical information. Additionally, CgA could not predict recurrence of disease or the disease status in the watchful waiting subgroup. Contrarily, CgA was a stronger predictor for mortality. This ambiguity can probably be explained by the previously supposed correlation between CgA and tumor load[14,37]. CgA is a secretory protein and therefore volumetric marker of disease and is

mostly negative in those with microscopic disease or patients with low tumor burden, while its correlation with hepatic tumor load probably makes it predictive for shorter survival[38]. However, as a result of the limitations, the independent contribution of CgA in daily practice is limited, especially with the ongoing advances in other diagnostics such as imaging that now use multidimensional mathematically calculated tumor volume as outcome measurement. We previously demonstrated that the NETest® does not have a correlation with tumor load [14], which is consistent with observations that it provides a measurement of tumor 'activity'. A predictive biomarker that reflects biological disease activity as opposed to tumor load creates a new method to delineate disease status and has therefore significant clinical utility.

Neuroendocrine tumors – like all malignancies – represent dynamic entities with evolution over time. Consequently, RNA levels and gene expression alter based upon tumor evolution and influencing factors like treatment. Determining the molecular alterations of tumors over time is a fundamental requisite of the NETest®. The reliability and reproducibility of serial NETest® measurement over a long period of follow up is therefore of utmost importance and must be validated. Serial liquid biopsies over years in patients with stable disease on imaging will give more insight in the dynamic behavior of GEP-NETs. Since NETest® gene expression measurement is based on the quantity of circulating transcripts, factors affecting the quantity of these transcripts in NET patients undergoing treatment (tumor degradation, ischemia) or suffering from comorbidities (other malignancies, benign diseases) need to be assessed. To our knowledge, independent validation of the reliability of serial NETest® and the reflection of the disease status over time is currently limited to only a subgroup of patients in one study [24]. Blind validation of multiple NETest samples in a prospective study with sufficient sample size,

intercurrent interventions (e.g. treatment initiation, (radio-)embolization) over a long period of follow up is therefore needed.

In conclusion, this study shows that the NETest® is currently the strongest predictor for disease progression and predicts RECIST defined disease status up to 1 year before this is apparent on imaging. The high negative predictive value, can support a watch-and-wait management in patients with a well differentiated GEP-NET. In head to head comparison, novel genomic analysis proved to provide more value than the monoanalyte marker CgA. It is apparent that with the NETest® personalized medicine in the management of GEP-NETs is one step closer.

Acknowledgements

We would like to thank Wren Laboratories for the performance of the NETest® in all samples, free of charges

Statement of ethics

This study was carried out in accordance with the recommendations of Netherlands Cancer institute (NCI) local ethics committee with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Netherlands Cancer institute (NCI) local ethics committee.

Conflict of interest statement

The authors declare no conflicts of interest

Funding source

The authors received no financial support for the research, authorship, and/or publication of this article.

Authors contribution

MvT performed the analysis and wrote the manuscript with input from all authors. MT, CK, LS, and WV collected all samples. DvdZ and BH re-assessed all imaging studies. GV and MT supervised the project. All authors discussed the results and contributed to the final manuscript.

Literature

1 Klöppel G, Rosa S La. Ki67 labeling index: assessment and prognostic role in gastroenteropancreatic neuroendocrine neoplasms. Virchows Arch. 2017;(472):341–9.

2 Dasari A, Shen C, Halperin D, Zhao B, Zhou S, Xu Y, et al. Trends in the Incidence, Prevalence, and Survival Outcomes in Patients With Neuroendocrine Tumors in the United States. JAMA Oncol. 2017 DOI: 10.1001/jamaoncol.2017.0589

3 Garcia-Carbonero R, Capdevila J, Crespo-Herrero G, Díaz-Pérez JA, Martínez del Prado MP, Alonso Orduña V, et al. Incidence, patterns of care and prognostic factors for outcome of gastroenteropancreatic neuroendocrine tumors (GEP-NETs): Results from the National Cancer Registry of Spain (RGETNE). Ann Oncol. 2010;21(9):1794–803.

4 Carmona-Bayonas A, Jiménez-Fonseca P, Lamarca Á, Barriuso J, Castaño Á, Benavent M, et al. Prediction of progression-free survival in patients with advanced, well-differentiated, neuroendocrine tumors being treated with a somatostatin analog: The Getne-Trasgu study. J Clin Oncol. 2019 DOI: 10.1200/JCO.19.00980

5 Clift AK, Faiz O, Goldin R, Martin J, Wasan H, Liedke M-O, et al. Predicting the survival of patients with small bowel neuroendocrine tumours: comparison of 3 systems. Endocr Connect. 2017 Feb;6(2):71–81.

6 Knigge U, Capdevila J, Bartsch DK, Baudin E, Falkerby J, Kianmanesh R, et al. ENETS Consensus Recommendations for the Standards of Care in Neuroendocrine Neoplasms: Follow-Up and Documentation. Neuroendocrinology. 2017; pp 310–9.

7 Korse CM, Taal BG, van Velthuysen M-LF, Visser O. Incidence and survival of neuroendocrine tumours in the Netherlands according to histological grade: Experience of two decades of cancer registry. Eur J Cancer. 2013;49(8):1975–83.

8 Pavel M, O'Toole D, Costa F, Capdevila J, Gross D, Kianmanesh R, et al. ENETS Consensus Guidelines Update for the Management of Distant Metastatic Disease of Intestinal, Pancreatic, Bronchial Neuroendocrine Neoplasms (NEN) and NEN of Unknown Primary Site. Neuroendocrinology. 2016;103(2):172–85.

9 Strosberg JR, Halfdanarson TR, Bellizzi AM, Chan JA, Dillon JS, Heaney AP, et al. The north American neuroendocrine tumor society consensus guidelines for surveillance and medical management of midgut neuroendocrine tumors. Pancreas. 2017;46(6):707–14.

10 Merola E, Pavel ME, Panzuto F, Capurso G, Cicchese N, Rinke A, et al. Functional Imaging in the Follow-Up of Enteropancreatic Neuroendocrine Tumors: Clinical Usefulness and Indications. J Clin Endocrinol Metab. 2017;102(5):1486–94.

11 Oberg K, Krenning E, Sundin A, Bodei L, Kidd M, Tesselaar M, et al. A Delphic consensus assessment: imaging and biomarkers in gastroenteropancreatic neuroendocrine tumor disease management. Endocr Connect. 2016;5(5):174–87.

12 Oberg K, Modlin IM, De Herder W, Pavel M, Klimstra D, Frilling A, et al. Consensus on biomarkers for neuroendocrine tumour disease. Lancet Oncol. 2015;16(9). DOI: 10.1016/S1470-2045(15)00186-2

13 Oberg K, Couvelard A, Delle Fave G, Gross D, Grossman A, Jensen RT, et al. ENETS Consensus Guidelines for Standard of Care in Neuroendocrine Tumours: Biochemical Markers. Neuroendocrinology. 2017 Apr;105(3):0–0.

14 Van Treijen MJC, Korse CM, Van Leeuwaarde RS, Saveur LJ, Vriens MR, Verbeek WHM, et al. Blood Transcript Profiling for the Detection of Neuroendocrine Tumors: Results of a Large Independent Validation Study. Front Endocrinol | www.frontiersin.org. 2018;9:740.

15 Rogowski W, Wachula E, Lewczuk A, Kolesinska-Cwikla A, Izycka-Swieszewska E, Sulzyc-Bielicka V, et al. Baseline chromogranin A and its dynamics are prognostic markers in gastroenteropancreatic neuroendocrine tumors. Futur Oncol. 2017;13(12):1069–79.

16 Tian T, Gao J, Li N, Li Y, Lu M, Li Z, et al. Circulating Chromogranin A as A Marker for Monitoring Clinical Response in Advanced Gastroenteropancreatic Neuroendocrine

Tumors. PLoS One. 2016;11(5). DOI: 10.1371/journal.pone.0154679

17 Marotta V, Zatelli MC, Sciammarella C, Ambrosio MR, Bondanelli M. Chromogranin A as circulating marker for diagnosis and management of neuroendocrine neoplasms : more flaws than fame. Endocr Relat Cancer. 2018;(Jan;25(1)):R11–29.

18 Jensen KH, Hilsted L, Jensen C, Mynster T, Rehfeld JF, Knigge U. Chromogranin A is a sensitive marker of progression or regression in ileo-cecal neuroendocrine tumors. Scand J Gastroenterol. 2013;48(1):70–7.

19 Chan DL, Clarke SJ, Diakos CI, Roach PJ, Bailey DL, Singh S, et al. Prognostic and predictive biomarkers in neuroendocrine tumours. Crit Rev Oncol Hematol. 2017;113:268–82.

20 Singh S, Asa SL, Dey C, Kennecke H, Laidley D, Law C, et al. Diagnosis and management of gastrointestinal neuroendocrine tumors: An evidence-based Canadian consensus. Cancer Treat Rev. 2016 Jun;47:32–45.

21 Boons G, Vandamme T, Peeters M, Van Camp G, Op de Beeck K. Clinical applications of (epi)genetics in gastroenteropancreatic neuroendocrine neoplasms: Moving towards liquid biopsies. Rev Endocr Metab Disord. 2019;20(3):333–51.

22 Öberg K, Califano A, Strosberg JR, Ma S, Pape U, Bodei L, et al. A meta-analysis of the accuracy of a neuroendocrine tumor mRNA genomic biomarker (NETest) in blood. Ann Oncol. 2020 DOI: 10.1016/j.annonc.2019.11.003

23 Pavel M, Jann H, Prasad V, Drozdov I, Modlin IM, Kidd M. NET Blood Transcript Analysis Defines the Crossing of the Clinical Rubicon: When Stable Disease Becomes Progressive.

Neuroendocrinology. 2017;104(2):170–82.

24 Liu E, Paulson S, Gulati A, Freudman J, Grosh W, Kafer S, et al. Assessment of NETest Clinical Utility in a U.S. Registry-Based Study. Oncologist. 2018;theoncologist.2017-0623.

25 Malczewska A, Witkowska M, Makulik K, Bocian A, Walter A, Pilch-Kowalczyk J, et al. Netest liquid biopsy is diagnostic of small intestine and pancreatic neuroendocrine tumors and correlates with imaging. Endocr Connect. 2019;8(4):442–53.

26 Lloyd R V, Osamura YR, Kloppel G, Rosai J. WHO classification of tumours of endocrine organs. WHO Press. 2017; pp 78–80.

27 Raza A, Ali Z, Irfan J, Murtaza S, Shakeel S. Analytical variables influencing the HCV RNA determination by TaqMan real-time PCR in routine clinical laboratory practice. Mol Biol Rep. 2012 Jul;39(7):7421–7.

28 Kidd M, Drozdov I, Modlin I. Blood and tissue neuroendocrine tumor gene cluster analysis correlate, define hallmarks and predict disease status. Endocr Relat Cancer. 2015;22(4):561–75.

29 Modlin IM, Drozdov I, Kidd M. The Identification of Gut Neuroendocrine Tumor Disease by Multiple Synchronous Transcript Analysis in Blood. PLoS One. 2013 May;8(5):e63364.

30 Modlin IM, Drozdov I, Kidd M. Gut neuroendocrine tumor blood qpcr fingerprint assay: Characteristics and reproducibility. Clin Chem Lab Med. 2014;52(3):419–29.

31 Modlin IM, Drozdov I, Alaimo D, Callahan S, Teixiera N, Bodei L, et al. A multianalyte PCR

blood test outperforms single analyte ELISAs (chromogranin A, pancreastatin, neurokinin A) for neuroendocrine tumor detection. Endocr Relat Cancer. 2014 Aug;21(4):615–28.

32 Eisenhauer E a., Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.

33 Genç CG, Jilesen APJ, Nieveen van Dijkum EJM, Klümpen HJ, van Eijck CHJ, Drozdov I, et al. Measurement of circulating transcript levels (NETest) to detect disease recurrence and improve follow-up after curative surgical resection of well-differentiated pancreatic neuroendocrine tumors. J Surg Oncol. 2018;118(1):37–48.

34 Modlin IM, Frilling A, Salem RR, Alaimo D, Drymousis P, Wasan HS, et al. Blood measurement of neuroendocrine gene transcripts defines the effectiveness of operative resection and ablation strategies. Surgery (United States). 2016. DOI: 10.1016/j.surg.2015.06.056

35 Stridsberg M, Eriksson B, Öberg K, Janson ET. A comparison between three commercial kits for chromogranin A measurements. J Endocrinol. 2003;177(2):337–41.

36 Molina R, Alvarez E, Aniel-Quiroga A, Borque M, Candás B, Leon A, et al. Evaluation of chromogranin A determined by three different procedures in patients with benign diseases, neuroendocrine tumors and other malignancies. Tumor Biol. 2011 Feb;32(1):13–22.

37 Seregni E, Ferrari L, Bajetta E, Martinetti a, Bombardieri E. Clinical signi¢cance of blood

chromogranin A measurement in neuroendocrine tumours. Ann Oncol. 2001;12(2):69–72.

38 Arnold R, Wilke A, Rinke A, Mayer C, Kann PH, Klose K-J, et al. Plasma chromogranin A as marker for survival in patients with metastatic endocrine gastroenteropancreatic tumors. Clin Gastroenterol Hepatol. 2008 Jul;6(7):820–7.

Figure 1: *Overview of the study population and different subgroups. The proportion of patients that showed progressive disease (PD) within 12 months is given between brackets. NED: No evidence of disease*

Figure 2

Figure 2 *illustrates the distribution of individual NETest scores (**a**) and CgA (**b**) between patients with progression and those without progression during the first twelve months of follow up. Median NETest in SD group was 27% versus 73% in PD group (horizontal bar). Median CgA was 78 µg/L versus 483 µg/L (horizontal bar). The y-axis in fig 2B is logarithmic*

Figure 3.

Figure 3 *shows the AUC for both the NETest and CgA. The NETest accuracy to predict the disease status at 12 months was 0.78 (95% CI 0.70-0.86) compared to 0.73 (95% CI: 0.64-0.83) for CgA (p = NS)*

Figure 4a+4b

Figure 4A *represents the categories with the threshold for low tumor activity decreased to 33%. A significant difference in mPFS between the low- and higher activity categories was observed:*: *55 months compared to 18 months and 11 months respectively (p <0.001; intermediate - high: p = 0.08).*

Figure 4B *shows the Kaplan Meier curve for CgA (ULN 140* µg/L) with significant difference in mPFS between the two curves: 55 months versus 12 months (p< 0.001).

Figure 5

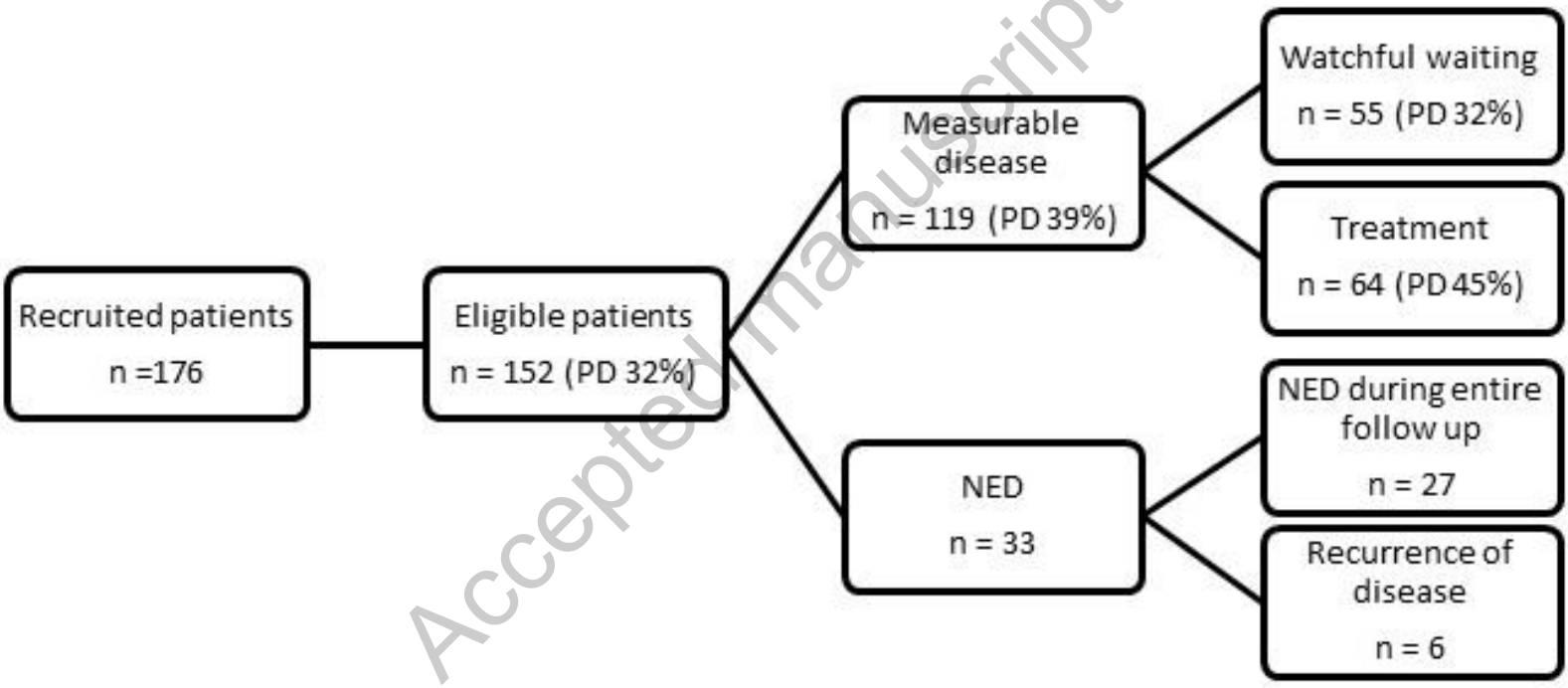
Figure 5 *shows the progression free survival for the combined outcome of CgA and NETest. When both tests are negative(-), a large proportion of patients remained stable over a long period of time.(Log rank p=0.02) Patients with positive results (+) in both tests had a significant lower progression free survival compared to patients with only a positive CgA (CgA+/NETest-: red line; p 0.04), but not compared to NETest+/CgA- (blue line, NS). mPFS was 55 months (both tests negative), 54 months (NETest-/CgA+), 18 months (NETest+/CgA-) and 9 months (both tests positive)*

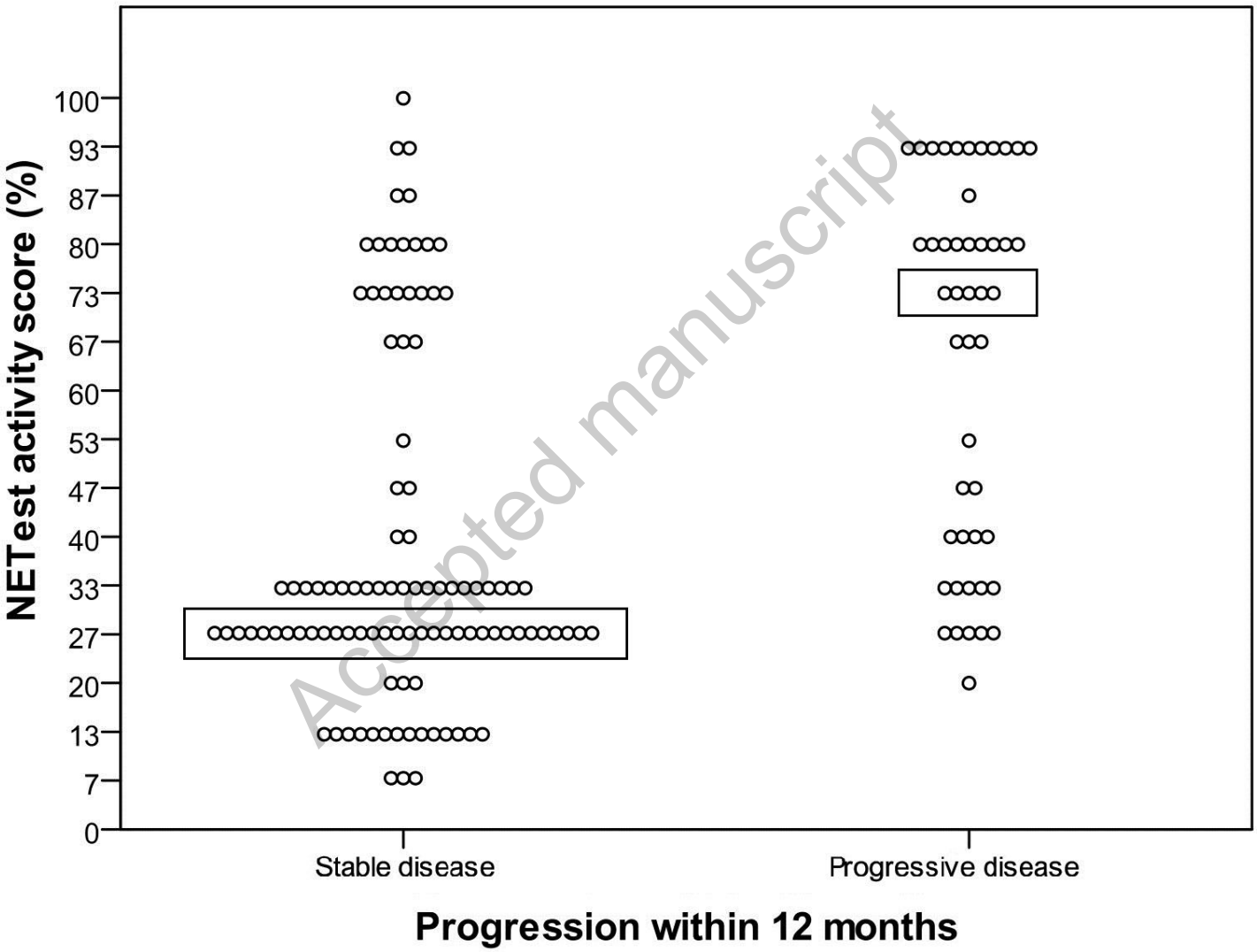
Figure 6

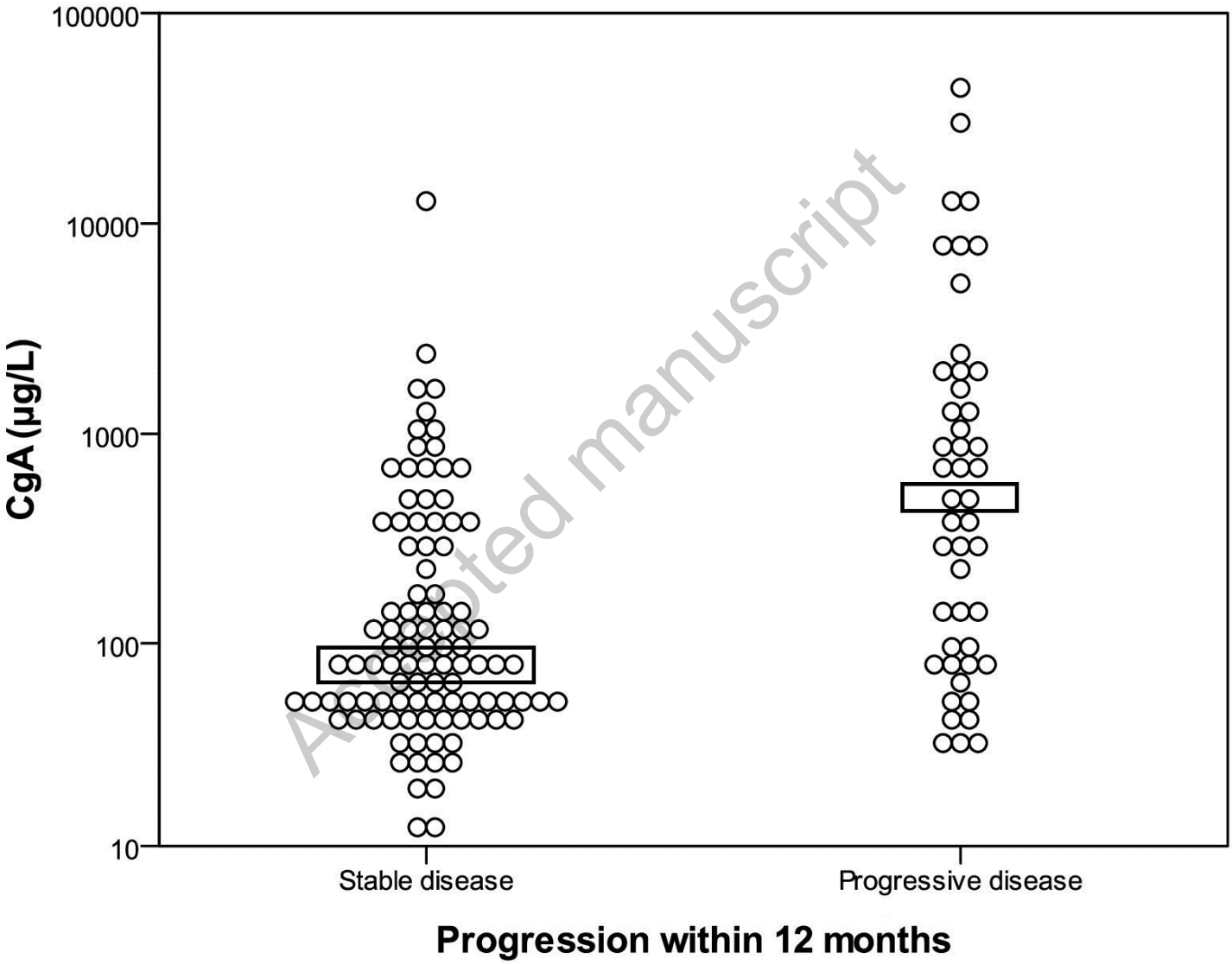
Cumulative PFS in each of the three NETest categories in the watch-and-wait subgroup (left) and in patients on treatment at baseline (right). In the watch and wait subgroup mPFS for the low activity (NETest ≤ 33%) group was 54 months compared to 12 months in intermediate activity group (34-79%; p = 0.015) and 12 months in high activity group (≥ 80%; intermediate-high: NS). In the treatment group, mPFS was not reached for the low activity group, compared to 9 and 11 months for the intermediate- and high activity categories (p<0.001; intermediate-high: NS)

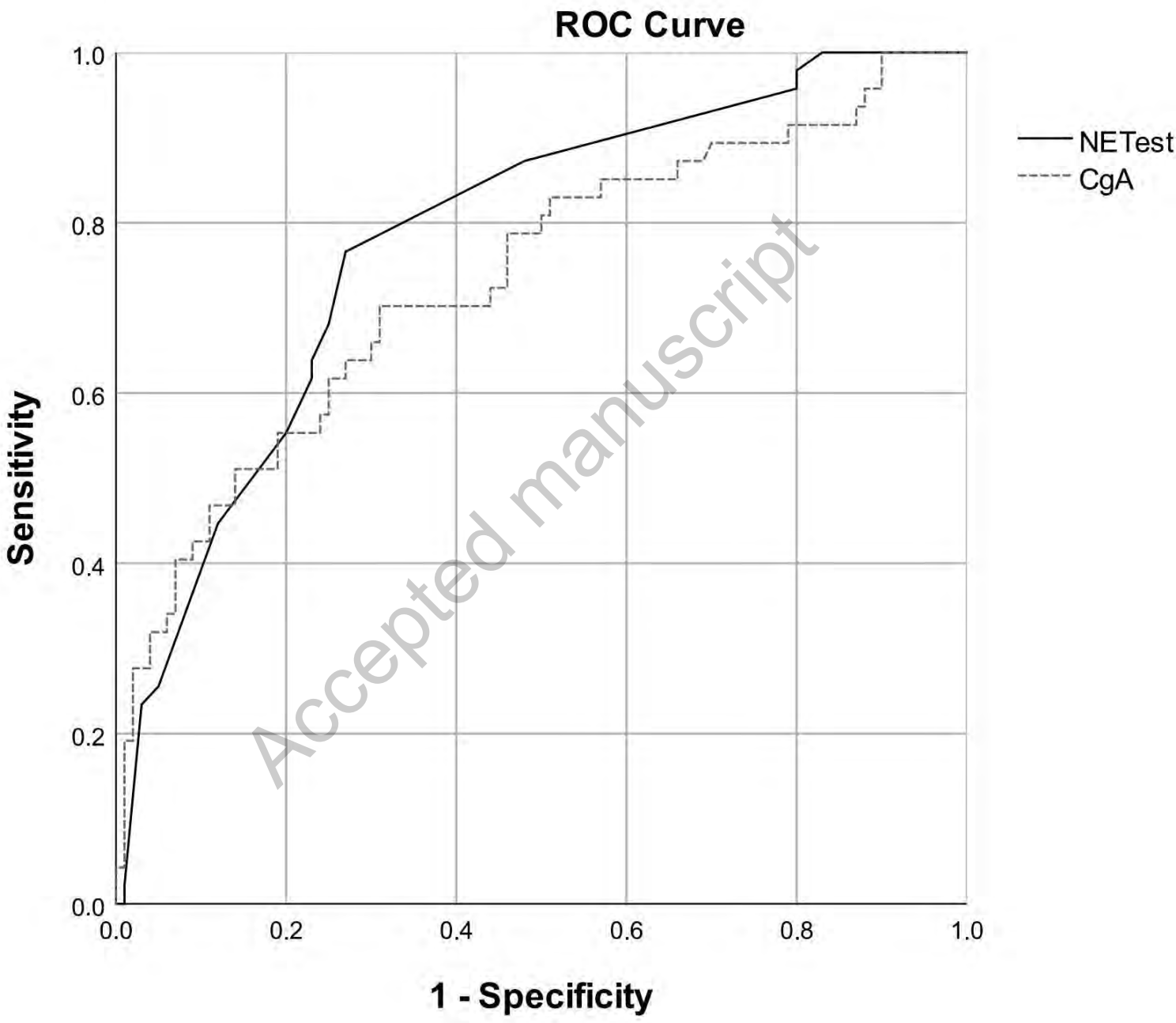
Figure 7a *illustrates the proportion of recurrent disease between patients with low NETest activity scores and those with high scores. All patients (n=33) had no evidence of disease at baseline.*

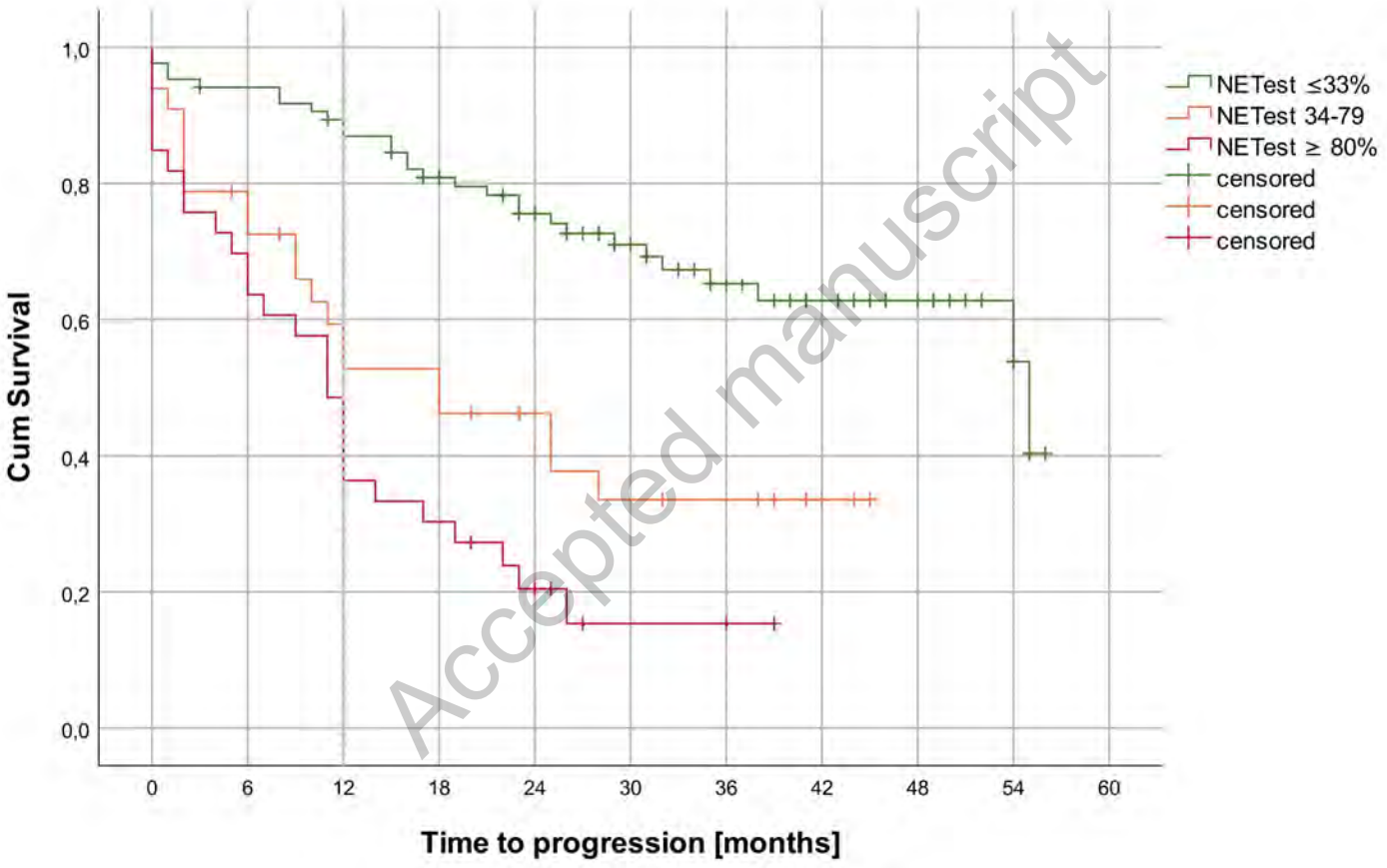
Figure 7b *represents the proportion of patients with recurrent disease for normal- and elevated CgA outcomes. There was no significant difference between the groups.*

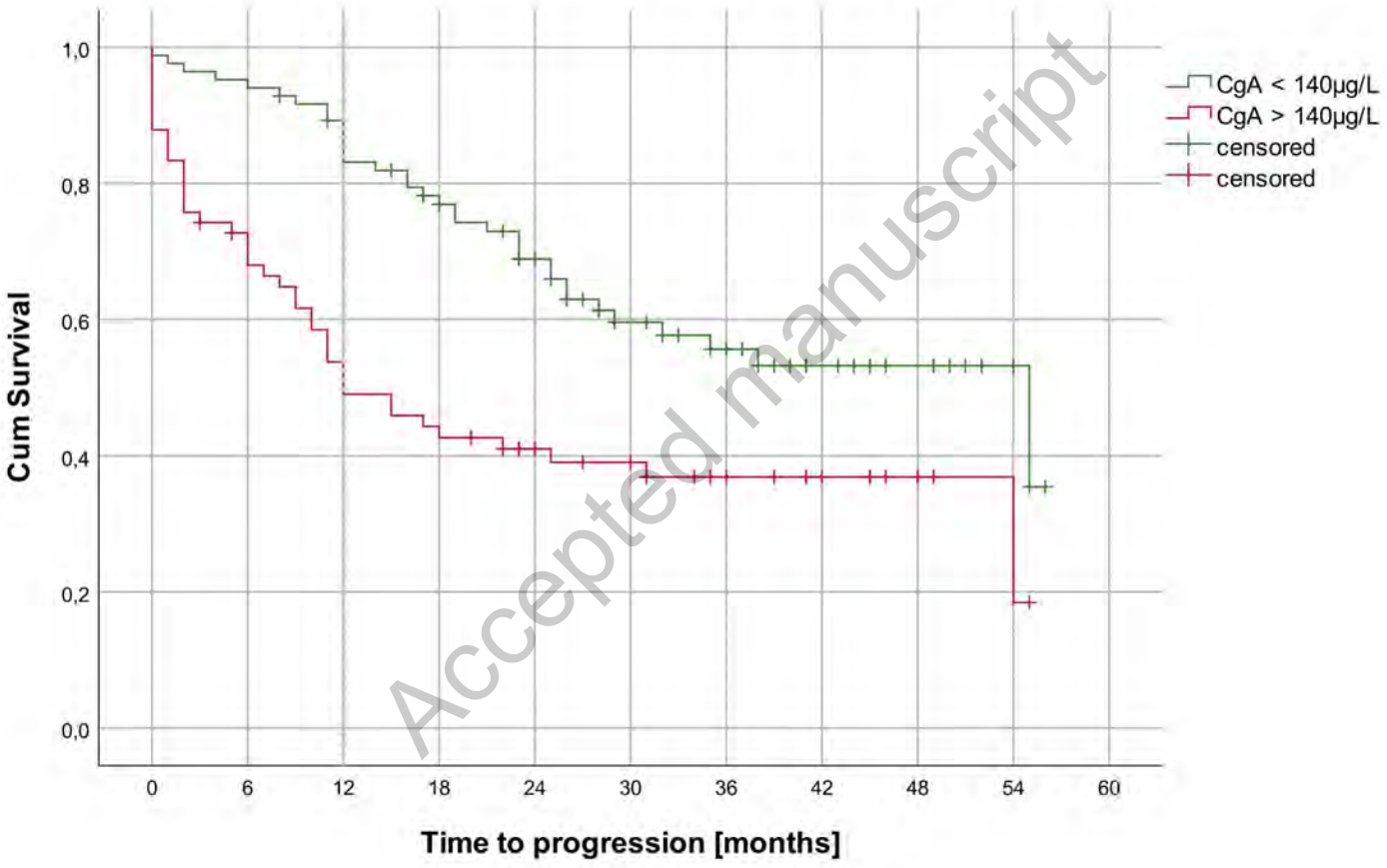


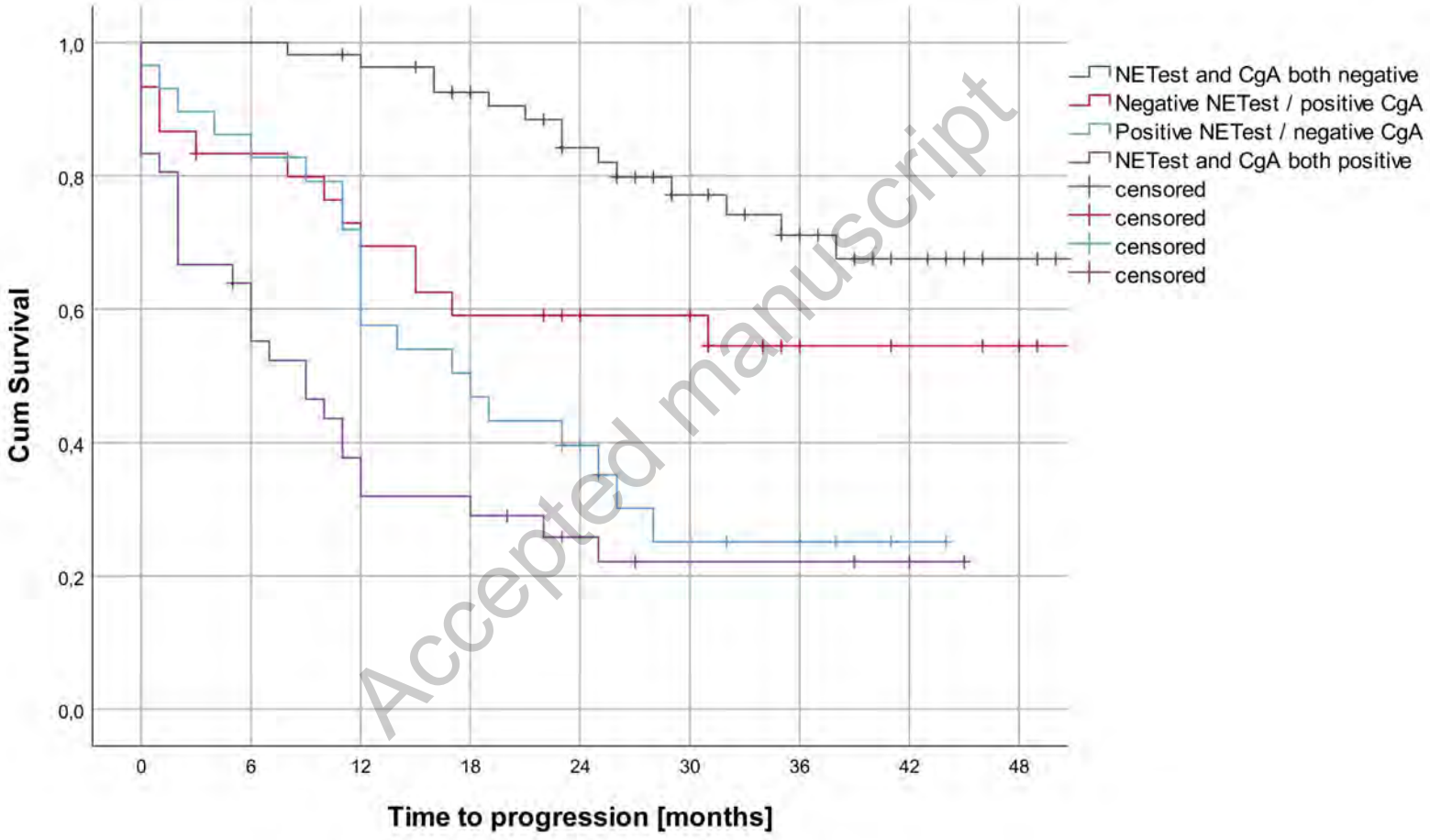


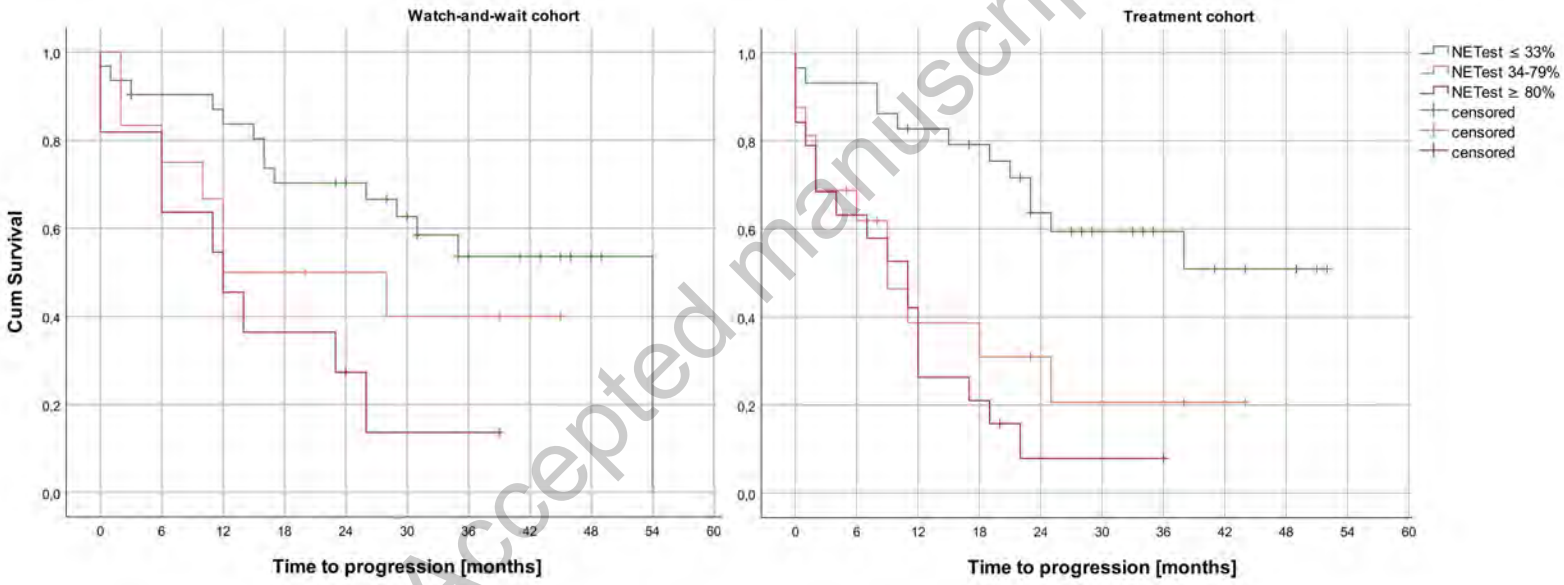


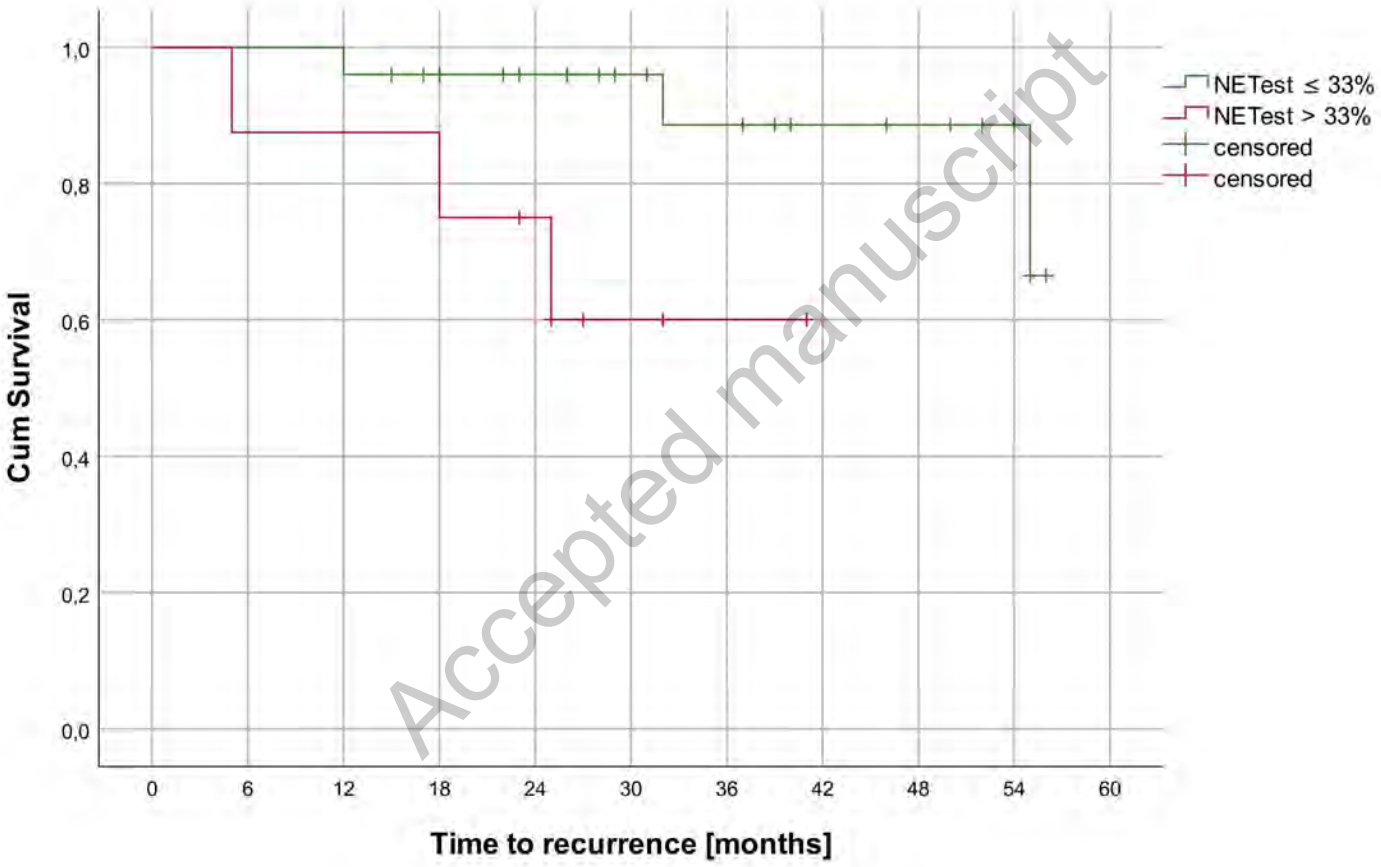












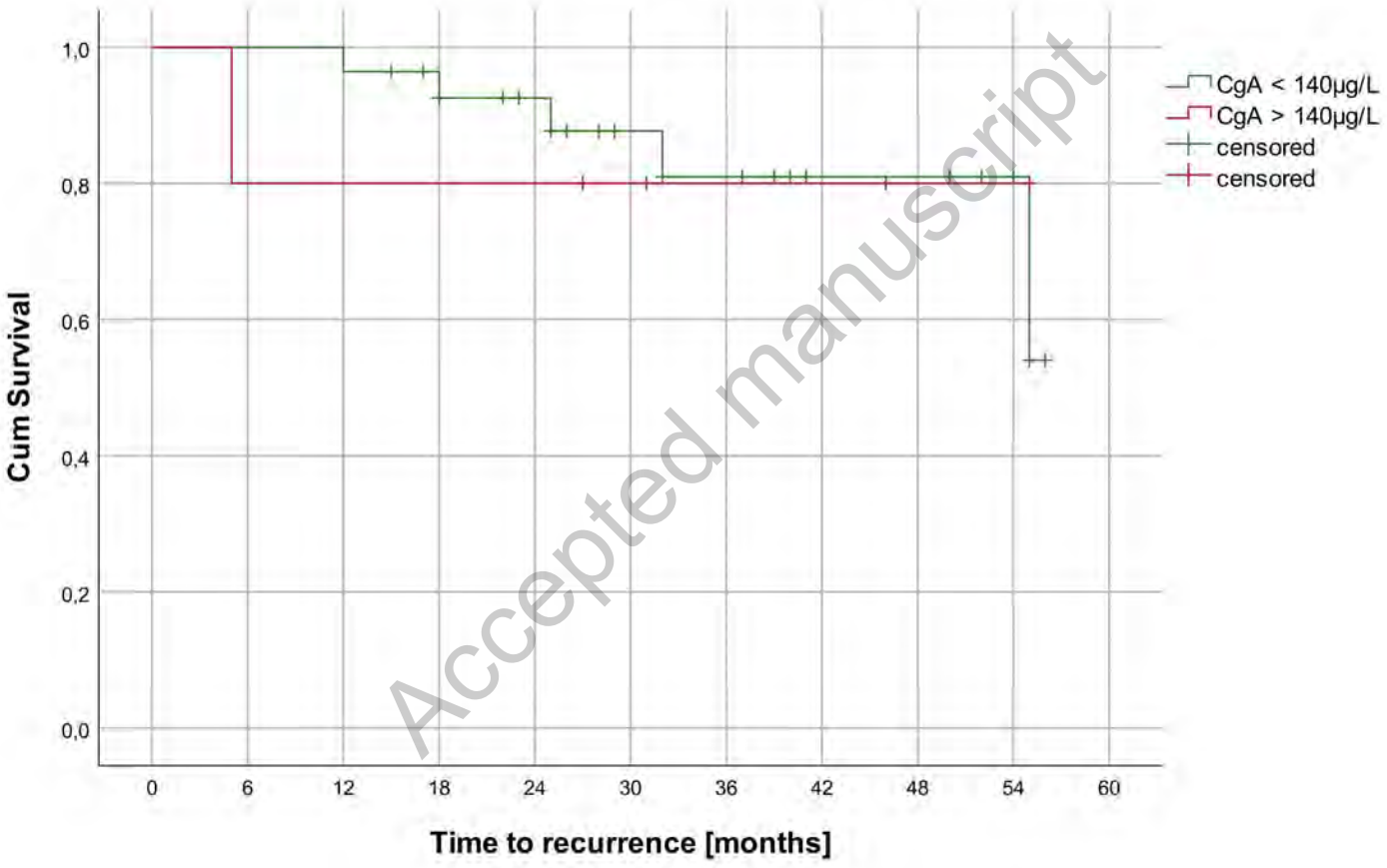


Table 1: baseline characteristics

Number of patients	152
Age in years, median (range)	63 (25-81)
Gender	
Male	82
Female	70
Primary tumor	
Small intestine	104
Pancreas	25
Gastric/duodenal	5
Appendiceal	5
Colon/Rectal	5
Unknown	8
Grade	
Grade 1	105
Grade 2	44
Grade 3	2
Missing	1
Disease stage	
No evidence of disease	33
Loco regional	1
Distant metastases	118
Current treatment	
None	88
SSA	60
Everolimus	3
CAPTEM	1
NETest median	33 (7-100)
Negative (%)	12 (8)
Low scores (%)	93 (61)
Intermediate scores (%)	26 (17)
High scores (%)	33 (22)
CgA median	107 (12-44150)
Normal (%)	72 (47)
Elevated (%)	79 (52)
Missing	1

Table 2: *Overview of metrics*

Test (cutoff)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
NETest (33%)	**77** [CI: 62-88]	**72** [CI: 62-81]	**56** [CI: 47-65]	**87** [CI: 80-92]
NETest (40%)	**68** [CI: 53-81]	**74** [CI: 65-82]	**55** [CI: 46-64]	**83** [CI: 76-89]
NETest (80%)	**45** [CI: 30-60]	**86** [CI: 80-94]	**64** [CI: 49-76]	**77** [CI: 72-82]
CgA (140µg/L)	**70** [CI: 55-83]	**69** [CI: 59-78]	**52** [CI: 43-60]	**83** [CI: 76-89]
CgA (100µg/L)	**70** [CI: 55-83]	**57** [CI: 47-67]	**43** [CI: 36-51]	**80** [CI: 72-87]

Table 2 – *Metrics for predictive ability for stable (NPV) and progressive (PPV) disease. The NETest has three categories and therefore the upper limit for low tumor activity and the lower limit for high tumor activity are presented. For both biomarkers, the original cutoff and the optimal cut off are demonstrated. CI = 95% confidence interval.*

Table 3: predictive value in different subgroups

Population	Disease status at 12 months		Disease status at 24 months	
	NPV (%)	PPV (%)	NPV (%)	PPV (%)
Total population (n=152)				
NETest	87	47/64	76	54/79
CgA	83	52	69	59
Watch and wait (n=55)				
NETest	84	50/54	70	50/74
CgA	79	41	63	49
Treatment (n=64)				
NETest	83	61/74	64	69/92
CgA	74	64	53	74
No evidence of disease (n=33)				
NETest	96	13	96	25
CgA	96	20	93	20

Table 3 *illustrates the positive predictive value (PPV; for the NETest, intermediate/high activity category are given) and negative predictive value (NPV) for the NETest (ULN 33%) and chromogranin A (*ULN 140µg/L*) in our total population and various subgroups.*

Supplementary Figure *1a* *shows the Kaplan Meier survival curve for each of the original NETest activity score categories. mPFS for the low activity (NETest ≤ 40%) group was 55 months compared to 18 months in intermediate activity group (41-79%; p < 0.001) and 11 months in high activity group (≥ 80%; p < 0.001; intermediate – high: p 0.055).*

Supplementary Figure *1B* *illustrates the survival curve for CgA (ULN 100 µg/L). mPFS was 38 months versus 22 months with a significant difference between both categories (p = 0.04)*

Supplementary Figure *2a* *Cumulative PFS in patients with normal- and elevated CgA in the watchful waiting subgroup. There was no significant difference between the groups.*

Supplementary Figure *2b:* *Cumulative survival for CgA outcomes in the subgroup of patients who were on treatment at baseline with a difference between normal- and elevated CgA (p=0.03)*