Published in final edited form as:

Int J Endocr Oncol. 2016 August ; 3(3): 203–219. doi:10.2217/ije-2016-0006.

The ENETS/WHO grading system for neuroendocrine neoplasms of the gastroenteropancreatic system: a review of the current state, limitations and proposals for modifications

Marcela S Cavalcanti¹, Mithat Gönen², and David S Klimstra^{*,1}

¹Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

²Department of Epidemiology & Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Abstract

The understanding of neuroendocrine neoplasms has evolved significantly since their initial descriptions in the 1800s to early 1900s. In the gastroenteropancreatic system, this group of malignant tumors is subdivided into well and poorly differentiated neuroendocrine neoplasms based on morphologic, proliferative and biologic differences. However, it has become increasingly apparent that well-differentiated neuroendocrine tumors are not a homogeneous group. Attempting to better predict outcome of these tumors has been the motivation behind numerous proposed classification systems, the evolution of which culminated with the currently used system, the ENETS/WHO classification. Herein, we review the genesis of this classification system and some of its shortcomings. In addition, we discuss some of the most recent proposals that suggest modifications to the current system.

Background

Neuroendocrine neoplasms (NENs) originate from the diffuse neuroendocrine cell system. Given the extent of the gastroenteropancreatic (GEP) tract, it harbors most of the neuroendocrine cells in the human body and as such, is the most common origin of NENs. Although uncommon, their incidence has been on the rise in the last four decades, presumably due to enhanced detection [1–3]. Much research has been dedicated to deepening our understanding of this disease, but many gaps remain to be filled. This is largely due to the extreme heterogeneity of this group of tumors. While major morphologic and biologic differences distinguish well-differentiated neuroendocrine tumors (WDNETs) from poorly differentiated neuroendocrine carcinomas (PDNECs), WDNET clearly is not a single entity, despite tumors classified as such sharing many histologic and

* Author for correspondence: Tel.: +1 212 639 5905; klimstrd@mskcc.org.

For reprint orders, please contact: reprints@futuremedicine.com

Financial disclosures/conflicts of interest

This research was funded in part through NIH/NCI Cancer Center Support Grant P30 CA008748. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

immunophenotypic features. Conversely to PDNECs, which are highly aggressive, WDNets are considered indolent, slow-growing neoplasms, but they are indisputably malignant. They demonstrate markedly variable outcomes, with 10-year disease-specific survivals ranging from 40 to 85%, which are difficult to predict [4–7]. With the purpose of better stratifying these tumors, particularly WDNets, a host of terminologies and classification systems has been applied. While each new system may have certain advantages, clear and uniform communication among investigators has been hampered by the array of taxonomy. Therefore, recent proposals have unified the criteria for the most critical prognostic classification parameters, grade and stage, for all GEP NENs. Herein we review the genesis of the currently used classification systems for NENs, discuss their shortcomings and conclude with a few modifying proposals we believe are warranted.

Historical aspects & basic terminology

Tumors with histologic features of what we currently recognize as WDNets were initially described in the 19th century in the form of case reports [8–12] but it was not until 1907 that the term ‘carcinoid’ (i.e., carcinoma-like) was first introduced by Siegfried Oberndorfer, distinguishing them as a separate entity. The term was used to denote their morphologic similarities to the more common intestinal (adeno)carcinomas, but distinct clinical behavior from the latter. In his original publication, Oberndorfer described 6 minute tumors (‘Geschwülstchen’ or tumorlets) of the ileum incidentally found in autopsies, which he believed to be benign due to their small dimensions, circumscribed borders and indolent behavior [13,14]. Nearly 20 years later, after having studied many more similar cases, Oberndorfer himself concluded that these tumors could be malignant and metastasize [13]. Subsequently, their origin from the diffuse neuroendocrine cell system, which in turn was presumed to derive from the neural crest, gave rise to the term ‘neuroendocrine’. While the first assumption may be true, the embryologic origin from the neural crest is not; it was later shown that neuroendocrine cells derive from endoderm instead [12]. This is the reason for considerable debate regarding the most accurate terminology (neuroendocrine vs endocrine) to be used for these tumors. However, the use of ‘neuroendocrine’ is currently preferred because these epithelial cells co-express antigens common to neural cells, namely synaptophysin and neuron-specific enolase, in addition to other neurosubstances, and they may exhibit neurite-like cell processes.

Much has evolved in the understanding of these tumors since their initial descriptions, albeit at a slow pace. For a long time, malignancy in NENs was only recognized by the late events of invasion of adjacent organs and distant metastases. Today, it is well established that all NENs are malignant and have a variable potential to metastasize, which cannot be accurately predicted based on morphology alone, particularly when dealing with WDNets. It is this variable clinical behavior that has led to a plethora of studies attempting to establish criteria to better predict behavior and guide management. Several classifications have been proposed and modified in the last five decades (Table 1). It was in the mid 20th century that investigators began to realize the heterogeneity of these tumors, despite their having similar morphologic features. Separations by origin, as done by Williams and Sandler in 1963 [15], and based on architecture and/or silver impregnation properties, such as that proposed by Soga and Tazawa in 1971 [16] and the WHO in 1980 [17,18], were the initial attempts to

account for that heterogeneity, but they proved of limited value for predicting clinical outcome.

In 1995, the first comprehensive classification of NETs of the lung, pancreas and gut was put forward by Capella and colleagues. Tumors were distinguished by site of origin and then subdivided into tumors with benign behavior, uncertain behavior, low-grade malignancies and high-grade malignancies based on histologic differentiation, tumor size, invasion of adjacent organs, angioinvasion, presence of metastases and clinically significant hormonal functionality [17]. It was also suggested to substitute the broader term ‘neuroendocrine tumor’ for ‘carcinoid tumor’ which was subsequently endorsed by the WHO in 2000. This was the first proposal to correlate tumoral features with clinical outcome and it proved to be effective to some extent [19]. In the following years, other factors, namely proliferative activity (as evaluated by diverse techniques), were shown to be of prognostic value in several studies, which would eventually become the mainstay of future classifications [20–24].

Capella’s classification in conjunction with the growing importance of tumor proliferative rate served as the basis for WHO’s NEN classification of the tubular GI tract (in 2000) and of the pancreas (in 2004), a system that combined staging and grading parameters into a single prognostic stratification [25–27]. Tumors from different locations in the GI tract were classified separately in a similar fashion to the 1995 classification, although the anatomic separation differed slightly from that originally proposed. A significant improvement in this classification was the inclusion of tumor differentiation as a category-defining feature. Although prior classifications alluded to the dichotomization of neuroendocrine neoplasms into well and poorly differentiated categories, it was in 2000 that histologic differentiation was incorporated into the classification. Subsequent classification systems are largely based on that characteristic due to its strong prognostic and predictive value [1,28], and as such, differentiation is currently a major basis for treatment decision-making. Histologic differentiation, as is evaluated in other tumors and organs, concerns to the degree a tumor resembles its normal cell counterpart (Table 2). Thus, well-differentiated NENs are characterized by monomorphic cells with granular cytoplasm and stippled (‘salt and pepper’) chromatin arranged in nested/organoid, trabecular, gyriform and/or tubular/glandular patterns (Figure 1). On the other hand, PDNECs are high-grade neoplasms composed of sheets, large nests and/or thick trabeculae of cells with high nucleus:cytoplasm ratio, abundant mitoses and apoptotic bodies, as well as extensive necrosis; they can be of small-cell or large-cell types (Figure 2). However, NENs do not progress in a continuum from well-differentiated tumors to poorly differentiated carcinomas, and therefore, there is currently no concept of a moderately differentiated category (which is not synonymous with intermediate grade or grade 2 – see below), due to its implication of a neoplastic progression model. In the 2000 classification, well-differentiated NENs were further subdivided into neuroendocrine tumor (WDNET) and neuroendocrine carcinoma (WDNEC), based on clinicopathological factors similar to those used in the Capella classification added to proliferative activity. The latter, evaluated by Ki-67 immunohistochemistry, was still a shy adjunct to the other factors. WDNETs could fall into the ‘benign behavior’ or ‘uncertain malignant potential’ categories, while WDNEC were considered low-grade malignancies [25–27]. Although several studies later validated the prognostic value of this classification, there were some limitations. First, its applicability in routine practice was quite challenging

due to the number of clinicopathologic variables needed for accurate classification, many of which were only evaluable at the time of resection. Also, this system was a hybrid of grading and staging, parameters which convey different messages about tumor behavior (inherent tumor aggressiveness vs tumor burden at diagnosis, respectively) (Table 2). Thirdly, the ‘uncertain malignant potential’ category brought much uncertainty to clinicians, pathologists and patients for obvious reasons, and in some anatomic sites such as pancreas, many cases in this category proved to have malignant behavior after longer follow-up [29,30]. A final point is that in the 2000/2004 WHO system, all metastatic WDNETs were classified the same (as WDNEC), precluding further prognostic stratification in stage IV patients, which still represent a very heterogeneous group in terms of disease-specific survival. Subsequent studies eventually showed there was no difference in outcome between the benign and uncertain malignant behavior categories of WDNETs, making clear the need for adjustments [29–32].

Genesis of the ENETS/WHO classifications & validation studies

In order to address the issues delineated above and further refine the available classification systems, the European Neuroendocrine Tumor Society (ENETS) proposed a tumor-node-metastasis (TNM) staging system and a separate grading system for foregut NENs in 2006 and for midgut and hindgut tumors in 2007 [33,34]. While staging criteria were site-specific, grading parameters were uniform throughout the entire GI tract. The grading scheme split NENs into low, intermediate or high grade based on mitotic rate and Ki-67 proliferative index (see Table 3 for definitions) based on agreed cut-points, although specific data to support the classification were not well established at the time it was first published. The ENETS grading scheme was later adopted by the WHO in 2010 due to the emergence of supportive evidence of its predictive power and has hence been designated the ENETS/WHO system. Several studies indicated that the proposed cut-points used to separate NENs into the three grades were prognostically significant, showing decreased survival as grade increased (Table 4) [4–6,35]. With regards to the staging proposal, while stage IV was consistently shown by subsequent studies to have a worse prognosis, there seemed to be no difference between stages I and II [4–6,36–40]. Despite that, a recent study comparing ENETS TNM staging system to other prognostic models (WHO 2000, WHO 2010/ENETS grading) showed that when two different systems were combined, the combinations that included ENETS TNM stage had the best prognostic performances. These results underscore the notion that grade and stage are complementary and both independently add important information [5,35].

In 2010, the 7th edition of the American Joint Committee on Cancer/Union for International Cancer Control (AJCC/UICC) TNM (AJCC/UICC TNM, 7th Edition) staging system for malignant tumors for the first time specifically addressed GEP NETs [41]. In this publication, a slightly different TNM staging scheme was proposed, particularly with regards to tumors of the pancreas and appendix. This has generated much confusion in the medical community because both ENETS and AJCC/UICC systems use the same ‘TNM’ terminology to refer to different extents of disease. In addition, no data were presented to justify the introduction of these modifications. When compared head-to-head, the system that performs better as a prognostic tool is still a matter of debate. A few studies, one of

which included 1072 patients with pancreatic NENs, have suggested a slight superiority of the ENETS proposal [30,42–43]. Others, however, have shown that both systems are largely comparable in their prognostication ability [40,44]. Regardless of the system used, data clearly indicate the need for further refinements in staging of these tumors, particularly for lower stage tumors. Currently, because both systems have become widely used (ENETS in Europe and AJCC/UICC in the USA), it is important to clearly state which specific staging system is being used in pathology reports. It has also been suggested that the stage defining features be addressed individually in the pathology report (tumor size and extent of invasion) in order to allow translation between both systems [45–47].

All efforts described thus far culminated with the publication of the revised edition of the WHO Classification of Tumors of the Digestive System in 2010, in which a chapter was dedicated to review the nomenclature and classification of NENs of the digestive system. As proposed by ENETS earlier that decade, the three-tiered grading system for NENs of the GEP system was adopted. The cut-points used were the same (Table 3). The Ki-67 index should be determined based on counting proliferative rate ‘hot spots’ identified by scanning the slide at low power, and whichever parameter (mitotic rate or Ki-67 index) pointed to the higher grade in a given tumor should be used for final determination of grade. Additionally, terminology was revised to better reflect the biology of these neoplasms (i.e., all NENs are malignant, regardless of grade). Thus, ‘neuroendocrine neoplasm’ was the preferred term to refer to the group as a whole, irrespective of grade. Grade 1 and 2 tumors comprised the (well-differentiated) ‘neuroendocrine tumor’ category, while grade 3 was referred to as ‘neuroendocrine carcinoma’ and could be of either small-cell or large-cell type. Many groups have validated the power of the current ENETS/WHO 2010 Classification; however, the inherent disadvantages of retrospective studies, such as selection bias and overfitting of data, suggest the need for prospective studies, which have not been conducted thus far [7,30,48]. This scheme proved especially useful in the not uncommon scenario of metastatic NENs of unknown primary, in which assessment of prognosis (and treatment) could be refined by grade despite the lack of a defined primary site [49]. For staging, the WHO recommended the site-specific AJCC/UICC system.

Six years after the WHO 2010 was published, its use has become widespread. However, as is often the case, time has brought some of its limitations to light (Table 5), as discussed below.

Issues regarding Ki-67 assessment

The number of cells expressing Ki-67 is higher than the mitotic count due to the functional properties of Ki-67: the protein (pKi-67) is expressed in all stages of the cell cycle, except G₀ (resting phase) [50]. Therefore, the Ki-67 labeling index is a more reliable measure of proliferation in tumors with a low proliferative rate, such as NETs. Given its critical role as a prognostic and predictive indicator, precise assessment of Ki-67 has become increasingly important. The ENETS/WHO grading scheme recommends counting 500–2000 cells in the areas of highest nuclear labeling (‘hotspots’), recommendations originally driven by common sense rather than data [51]. A subsequent study, however, has verified that the hotspot Ki-67 index is a more accurate determinant of prognosis than the average Ki-67, supporting the ENETS/WHO grading provision [52]. However, there is no clear definition of

what constitutes a hotspot; use of a 1.0-mm diameter region (one 20× field on many microscopes) has been suggested. In addition, differential intensity of staining is commonly encountered with several antibody clones and it is not entirely clear if all positive cells, regardless of intensity of staining, ought to be included in the final count. Given the knowledge that expression of pKi-67 varies depending on the cell cycle phase, it would seem logical to consider any evidence of staining as positive [50,53–54]. Preanalytical factors also play an important role in Ki-67 results. From fixation times and fixative type to section thickness and specific immunohistochemistry protocols [55,56], several steps influence the reaction and hamper full standardization of the process. A recent study revealed a distressing rate of variation for single cases stained in multiple different laboratories [57,58].

Several recent studies have attempted to establish the best method for evaluating Ki-67. Visual approximations, also referred to as ‘eyeballed’ estimates, although quick and simple, are inaccurate. ‘Eyeballing’ cannot distinguish the subtle differences between G1 and G2 (in up to 40% of cases in some series), has a high degree of interobserver variation, and is thus not recommended [59,60]. Another commonly used technique in general practice is visual counting on the microscope (‘real time’ counting of cells in a given microscopic field, in which one relies on memory to keep track of the cells already counted). This method was found to have relatively high accuracy but was the least reproducible among four methods tested [61]. Evidence indicates that manual cell counting, either utilizing printed images or on a computer screen where counted cells can be differentially marked, is a low cost, highly accurate method that requires about 8 min per case [59–61].

Automated counting methods using digital images are emerging as a means to overcome the intra- and inter-observer variabilities inherent in human counting and to save time [59,62]. However, there are shortcomings that prevent its widespread implementation. Fully automated systems are costly, not widely available, and require diligent calibration of the software and training of the technician prior to implementation. Furthermore, the actual counting time taken by the machine is about 5 min per slide – not a huge improvement over manual counting of printed photos [61]. Partially automated systems have comparable reproducibility but are more accessible (there are many free image analysis softwares available); however, they also rely on an experienced observer to select the best area for counting and assess the precision of the final count, excluding any nontumoral cells that the software may have included [63]. Therefore, for now, it seems that the method that best allies accuracy, practicality and reproducibility is manual counting from a printed image. This method is relatively quick and easily performed, both in specialty centers and in the community setting, and the results are much more reproducible than ‘eyeballing’.

Heterogeneity of Ki-67 labeling can occur, giving rise to the already mentioned hotspots. There may be heterogeneity within a primary NET, between the primary and metastasis, or even between different metastatic sites. Tumor heterogeneity further adds to the inconsistency of Ki-67 evaluation, especially based on biopsies that may miss the Ki-67 hotspots present elsewhere in the tumor. Using ‘virtual biopsies’ of resected liver metastases, Yang *et al.* found Ki-67 staining of core biopsies will undergrade the tumor when the whole section reveals G2 regions. If three core biopsies were taken, only 48% of G2 NETs were

correctly graded; the rate dropped to 35% when only one core was obtained [52]. Nonetheless, the grade obtained on core biopsies accurately stratified patient outcome – at least based on the entire cohort being studied – which argues that tumors with more abundant G2 regions (more likely to be sampled on randomly oriented biopsies) have a worse outcome than tumors with limited G2 regions (in which a random biopsy would more likely reveal the predominant G1 foci). Proliferation differences among different metastases have also been investigated by Couvelard *et al.* [64]. The authors of this study interpreted their results to show no significant variation of Ki-67 between different metastases or within a single metastatic site; however, review of their data reveals that in 16 of 29 patients, there was sufficient variation among sites to change the grade, including three cases with variation between G1 and G3 foci in different sites of disease [64]. Grade progression over time is a well-established phenomenon in NETs (ranging from 15 to 65% of cases), and therefore, biopsying new metastases may be important to determine subsequent steps in treatment [65–67].

A number of alternate proliferation markers, namely phosphohistone H3 (PHH3), have recently been proposed. PHH3, a core histone protein, is expressed only during the mitotic phase of a cell and an immunohistochemical assay has been recently investigated as a potential substitute for mitotic figure counting based on H&E. However, only two studies have explored its use specifically in GEP NENs [68,69]. Both studies indicate good correlation between PHH3 and Ki-67 and the prognostic value of PHH3 counts. Given the current paucity of data on the subject however, further studies are needed to redefine or validate the cutoffs and counting methodology proposed by Voss *et al.* [69]. It is fair to assume that a more sensitive technique to identify mitotic figures (such as immunohistochemistry) will yield different values for the optimal prognostic cut-points, and establishing these values will require the same level of diligence that went into defining mitotic rate and Ki-67 index cut-points.

Heterogeneity of the WHO G3 (high grade) category

Although studies have shown a strong correlation between mitotic rate and Ki-67 proliferation index [49], there are cases in which the two parameters are discordant, each placing the tumor in a different WHO grade category. When this occurs, it is typically the Ki-67 index that points to the higher grade. As mentioned, this issue was anticipated in the 2010 WHO classification, which suggested that the higher of the two parameters be used to assign the final grade [70]. Such practice was later supported by data demonstrating mitotic G1/Ki-67 G2 tumors behaved more like mitotic G2/Ki-67 G2 ones [71]. Although this situation may be encountered between any two consecutive grades, it becomes a more important issue when tumors with a mitotic count in the G2 range, still considered WDNET, fall into the G3 category based on a Ki-67 >20%, in which case they are considered high grade NECs in the 2010 WHO system. High grade NECs have a far worse prognosis than their G1 and G2 counterparts, presenting with metastases in 50–70% of patients, and most patients succumbing within the first 2–3 years of diagnosis [32,72]. More importantly, as of the current classification, G3 NECs are regarded as poorly differentiated by definition, being small-cell or large-cell neuroendocrine carcinomas – entities with significant biological, genetic and treatment differences compared with WDNETs. In general, first-line therapy for

G1 and G2 WDNETs includes surgery and – depending on tumor burden, primary origin and symptomatology – somatostatin analogs, chemotherapy with nonplatinum agents, targeted therapy, or local ablation (bland or chemoembolization). In contrast, platinum-based chemotherapy and/or radiotherapy with or without surgery is the advocated treatment for PDNECs [73–75]. Therefore, the distinction between well-differentiated and poorly differentiated neoplasms is paramount for patient management. While the 2010 WHO classification uses the terms ‘high grade’ and ‘poorly differentiated’ interchangeably for tumors in the G3 category, recent studies have challenged that paradigm by demonstrating that the G3 category is in fact heterogeneous, comprised of two different subgroups: (1) well-differentiated NETs with a high proliferative rate (grade-discordant WDNETs or G3 WDNETs) and (2) true poorly differentiated NECs (small-cell or large-cell types) [76–78].

Clinical evidence to support the G3 heterogeneity was first put forth by Sorbye and colleagues in a study that included 305 patients with high grade NEC of the GI tract. Their findings indicated a lower response rate after platinum-based systemic chemotherapy (15 vs 42%, respectively; $p < 0.001$), but a longer median overall survival (14 vs 10, respectively; $p < 0.05$) among patients with tumors with a Ki-67 $< 55\%$, compared with those having a higher Ki-67 index [72]. Other groups have reached similar conclusions with smaller case cohorts [76,79].

More recently, Tang *et al.* reported on 31 patients with WDNETs (by morphology and proliferative indices) with a separate, morphologically distinct high-grade component, characterized by confluent growth, areas of necrosis, increased nuclear atypia and higher proliferation, surpassing the 20% Ki-67 threshold, but still regarded as well-differentiated G3 components rather than true PDNECs. Genotyping of some of these tumors showed *DAXX*, *ATRX* or *MEN1* mutations – molecular alterations that have been described exclusively in WDNETs – in both the lower grade and high-grade components in three of four cases tested. Conversely, *RB1* mutations and alterations in p53 expression, which are exclusively found in PDNECs, were not detected in any of the tested cases. This study also emphasizes that G3 WDNET may occur as either morphologically classic well-differentiated NETs in which their high-grade nature is only revealed by Ki-67 staining, or as a well-differentiated NET with areas displaying atypical cytoarchitectural features suggestive of progression to a high-grade neoplasm (Figure 3) [80,81].

In summary, while all PDNECs are high grade, most often displaying a Ki-67 well beyond the 20% threshold (average within the 60–75% range) [79], not all high grade NENs as defined by Ki-67 are poorly differentiated based on morphology. Substantial evidence supports that there is a subset of high grade NENs that displays well-differentiated cytoarchitectural features. Tumors belonging to this subgroup seem to occur more frequently in the pancreas [76–77,80]. They also tend to have a Ki-67 in the lower end of the G3 category (usually between 20–50%), fare better than true PDNECs but somewhat worse than G2 NETs, and do not respond well to platinum-based chemotherapy. Furthermore, their molecular abnormalities seem to parallel those of WDNETs (inactivating mutations of *DAXX* and *ATRX*, mutations in *MEN1*), supporting that these G3 WDNETs are indeed distinct from PDNECs [80,82]. Therefore, there is a need for modifications in current classifications and treatment protocols in order to incorporate these novel findings.

Tumor location issues

Although all WDNETs arise from the diffuse neuroendocrine system, accounting for their largely shared phenotype, there are many epidemiologic and biologic singularities depending on their site of origin, partly owing to the 14 different neuroendocrine cell types from which they can derive [83]. From the recent epidemiologic standpoint, the stomach is the most commonly affected organ [84,85]. Also, while WDNETs prevail in most of the GI tract, PDNECs tend to occur with a much greater frequency than their well-differentiated counterpart in the colon and the squamous-lined portions of the tubular GI tract (esophagus and anus) [86].

It has also been suggested that primary tumor location *per se* is a predictive factor of outcome. Two large registry-based studies, one from the US (Surveillance, Epidemiology and End Results [SEER] data) and one from Spain (National Cancer Registry), demonstrated the strong correlation between primary tumor site and disease stage. However, given the different populations studied, their findings differed slightly. In the Spanish study, the most common primary sites associated with distant metastases were jejunum/ileum (65%), colon (48%) and rectum (40%), whereas in the SEER registry, the pancreas outnumbered other sites (65%), and cecum/colon and jejunum/ileum followed (44%/32% and 30%, respectively) [1,87]. Other smaller series have had somewhat similar results, with jejunoileal tumors more commonly presenting with widespread disease than appendiceal and rectal tumors [88]. SEER-based studies have also shown differing 5-year overall survival rates for NETs depending on primary site, even for advanced stage disease, varying from 10% for gastric NETs to 40% for small bowel NETs, emphasizing the stage-independent biological differences among primaries [89]. The link between pancreatic origin and poorer outcome, as shown in the SEER-based study, is still controversial. One group reported pancreatic origin as an independent adverse prognostic variable [90], whereas another group demonstrated a 42.9% rate of metastatic disease among primary pancreatic NETs in comparison to 91.9% of those primary in the small bowel [66]. In the former study however, PDNECs were included, which may contribute to the difference in outcome rather than primary location itself.

Within a single anatomic site there are also variations depending on the presence of predisposing conditions or type of hormone produced. In the stomach for example, there are three main types of WDNETS, each with a particular clinical behavior. Type I NETs are associated with autoimmune atrophic gastritis while type II NETs, the rarest of the three, arise in association with multiple endocrine neoplasia type 1 (MEN1) and Zollinger-Ellison syndrome (ZES). Both tend to be small and multiple and have an excellent prognosis, whereas type III gastric NETs arise in a normal background mucosa, are solitary and have a much worse prognosis [84]. Pancreatic insulinomas are regarded as indolent tumors compared with other functioning tumors from the same site, likely related to hormone hypersecretion causing florid symptomatology even in small tumors, leading to detection at an earlier stage [84].

Conversely, tumors that secrete the same hormone may behave differently depending on the site of origin. Gastrinomas of the duodenum are known to give rise to nodal metastases early

in the course of disease (60–80% of cases at the time of diagnosis), even with microscopic primaries, but have a much lower rate of liver metastases (~10%). Despite detection at a late stage, 10-year survival rates of up to 84% have been reported for these tumors. Similarly to their duodenal equivalents, pancreatic gastrinomas also have a high rate of lymph node involvement. However, liver metastases are more common (in ~30% of cases) and their reported 10-year survival rate is only 51% [84].

The disparate molecular underpinnings of WDNETs depending on primary location may play an important role in these biologic differences. Chromosome 11q loss with or without *MEN1* gene mutations are involved in the initiation of a significant proportion of foregut NETs, while 18q defects are present almost exclusively in midgut and hindgut NETs. In addition, *p16^{INK4a}* methylation, *CTNNB1* (beta-catenin encoding gene) mutations and CpG island methylator phenotype (CIMP) are more frequent in GI than pancreatic NETs [91–93]. Furthermore, in addition to alterations in the MEN1 pathway, mutations in *DAXX* and *ATRX* are another common finding in WDNETs of the pancreas (up to 43%), as are mTOR pathway derangements (up to 14%) [94].

PDNECs, on the other hand, appear to be more closely related, independent of primary location. Although one group found primary colon G3 NECs to have a shorter survival than pancreatic primaries (8 vs 15 months) [72], this difference may reflect the higher incidence of G3 WDNETs in the pancreas (which have been classified as G3 NECs). In the study by Basturk *et al.* of 44 true PDNECs of the pancreas, the median overall survival was 11 months, similar to that reported for the same neoplasm in the colon [78].

These differences in biology according to anatomic topography lead to slightly different prognostic factors for each site, which is the rationale behind the site-specific TNM staging system currently in use. In view of these singularities, it is somewhat paradoxical that grading criteria are uniform throughout the entire GEP system. Although establishing different cutoffs for each location would make grading a troublesome task, particularly in the setting of metastases of unknown primaries, the site-specific clinical and molecular suggest that a universal grading system for these neoplasms may not be adequate. Further studies attempting to establish optimal cut-points for individual primary sites are needed to better reflect their inherent differences.

Proposals to change Ki-67 cut-points

The cut-points used in the ENETS/WHO 2010 grading system have long been a point of contention. Although most studies subsequent to the development of this proposal have shown good prognostic stratification for many anatomic sites, a number of groups have proposed alternate cut-points for the Ki-67 index, mostly for the G1/G2 threshold but also for G2/G3. Most of these studies have addressed pancreatic NENs. Proposed alternative values for the G1/G2 Ki-67 cut-point vary considerably in the literature, ranging from 2.5–15%, and from 9–15% for the G2/G3 cut-point (Table 6). These wide variations, especially those for the G1/G2 threshold, may be explained by limitations in some of the studies that have advanced those proposals. First and foremost, several of the studies that suggested a new cut-point for G1/G2 included tumors with a Ki-67 within the G3 range and/or with

poorly differentiated morphology in their analyses [21,23,32,95]. It is well established that PDNECs have a significantly higher Ki-67 index, usually above 50%, and a much worse outcome than WDNETs. Therefore, the inclusion of these tumors with higher Ki-67 indices would obviously skew the G1/G2 threshold towards a higher value. Secondly, some studies establishing cut-points had insufficient sample sizes for the cut-point estimates to be statistically reliable [23,95]. In addition, receiver operating characteristic (ROC) analysis has often been inappropriately utilized by using the binary outcome of dead versus alive, without taking into account survival time, time to recurrence or duration of follow-up [42,96,97], whereas other studies did not detail why the alternate cut-points were chosen [98,99]. As to location, as previously mentioned, most studies considered primary pancreatic tumors, but Faggiano *et al.* included NENs not only from the GI tract, but also from the lung and ovary, resulting in a markedly heterogeneous cohort [95].

The strongest evidence to support altering the G1/G2 cut-point has been set forth by Scarpa *et al.*, but not without a few shortcomings. In their study of 237 pancreatic NENs with known Ki-67 indices, Ki-67 values of 5% and 20% were among the four independent prognostic variables on multivariate analysis. In comparison, when the ENETS/WHO grading system was used, it did not distinguish between G1 and G2 outcomes on multivariate analysis. Furthermore, there was no difference in survival among patients whose tumors had a Ki-67 within the 0–5% range (10-year survival rate of 80% for patients with Ki-67 ≤ 2 vs 95% for Ki-67 2–5%; $p = 0.487$) [5]. However, their multivariate analysis models included too many variables, resulting in an overfit model in which true associations may be difficult to draw [103–106]. Another important consideration is that this study of primary tumors may not perfectly parallel the findings in the metastatic setting, in which the difference between G1 and G2 would more likely have clinical implications. The sampling of metastases for grading usually means assessing biopsies, in which the true grade is often underestimated [52]. Therefore, the cut-point to optimally stratify the behavior of metastatic NETs may be lower than that for primary tumors due to the artifactually lower Ki-67 values identified in small samples from these sites of disease.

Attempts to reproduce the 5% Ki-67 index for the G1/G2 cut-point proposed by Scarpa *et al.* have ensued. Khan and colleagues' study on stage IV pancreatic and midgut NETs found the 5% cut-point to better predict outcome of G1 and G2 tumors [102]. Boninsegna *et al.* reached similar conclusions [101]. It must be noted, however, that in both studies the 2% cut-point was able to stratify patients into clinically significant groups as well, albeit to a lesser degree, either by having higher p-values or by only reaching significance on univariate analyses, not on multivariate analyses.

Another important point is that Ki-67 is a continuous variable. It is well known to statisticians that dichotomization of continuous variables results in reduced power of a given variable in regression models and loss of information regarding individual differences [107–111]. As an example, in separating WDNETs into two groups according to the 3% threshold (or any cut-point used), one is assuming that tumors with a 3% Ki-67 and a 19% Ki-67 have similar prognoses. On the other hand, it leads one to infer that the tumor with the 3% Ki-67 behaves significantly worse than a tumor with a Ki-67 of 2%, despite the relatively minimal absolute difference of these values. Furthermore, it appears that Ki-67 carries a linear

relationship with prognosis and outcome (at least within the well-differentiated group), and therefore, the slope can be divided into two prognostically significant groups at essentially any given point. Studies that have included Ki-67 as a continuous variable have demonstrated it to be a significant predictor of recurrence and have shown its superiority to Ki-67 as a categorical variable, irrespective of the cut-points used [30,40,101].

Interestingly, while most studies have suggested increasing the Ki-67 cut-point to distinguish G1 and G2 tumors, Hochwald's group concluded the opposite with mitotic count, finding a mitotic rate of 2/50 high power fields (HPF) to optimally stratify WDNets of the pancreas [112]. This scheme has in fact been validated by subsequent investigations [29,30], and a similar proposal for rectal tumors has also been published [113].

Therefore, although there is some evidence to support changing the Ki-67 cut-point to 5% for pancreatic NETs, we believe it is still premature to institute any changes. Supportive evidence is still scarce, especially in the metastatic setting. Thus, further validations are warranted, as well as attempts to test the proposed cutoff in other anatomic sites of the GI tract and in metastatic cohorts. More importantly, however, is the fact that current treatment options for G1 and G2 GEP-NETs are largely overlapping, especially for resected primary tumors. Proposals to treat G2 metastases more aggressively than G1 are still evolving [73,114–115]. Additionally, data presented herein have shown that the current ENETS/WHO classification does correctly stratify patients with G1 and G2 WDNets into clinically significant subgroups, even though the statistical power of the proposed cut-point of 5% may be greater for pancreatic primaries. Therefore, we share the view of Rindi *et al.*: “The substantial efficacy overlap of both systems does not justify a change in the current WHO grading” [42]. Furthermore, given the continuous nature of Ki-67 as a variable, struggling to find an ultimate cut-point for WDNets may not be the best approach. Instead, future studies should attempt to address the impact of the specific Ki-67 index on response to various treatment regimens, which would be more clinically meaningful. For this reason, documentation of the actual Ki-67 index (and mitotic rate) in pathology reports – in addition to the corresponding grade – is recommended [51].

Conclusion

With emerging targeted therapies, accurate classification and prognostication assume seminal importance to guide clinicians towards providing the best available treatment for patients with NENs of the GEP system. While many of the prior classifications have used either tumor differentiation or proliferative rate as the mainstay for NEN subclassification, it has become clear that both parameters complement each other to adequately stratify and manage patients. Using only one parameter fails to recognize G3 WDNets. Although such cases are relatively rare, misclassifying them can lead to drastically different clinical approaches. Thus, a separate category to incorporate this new entity must be included in future classifications. Additionally, while the first NEN classifications have tried to acknowledge the biologic differences based on site of origin, that has been largely ignored in the current grading system (but not in the staging system) despite recent advances in the molecular realm supporting that genetic background varies with location. Therefore, further studies attempting to establish site-specific grading parameters should be undertaken. With

regards to changing the Ki-67 and/or mitotic count indices for G1 versus G2 separation, we believe implementing any changes at present would be premature and further studies are needed to support a more evidence-based decision.

The amount of effort put into unraveling these tumors so far has been undoubtedly immense. Clearly though, there are unexplored avenues and space for improvements, ultimately aiming at more precise classification and better patient care.

References

1. Yao JC, Hassan M, Phan A, et al. One hundred years after “carcinoid”: epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. *J Clin Oncol*. 2008; 26(18):3063–3072. [PubMed: 18565894]
2. Fraenkel M, Kim MK, Faggiano A, Valk GD. Epidemiology of gastroenteropancreatic neuroendocrine tumours. *Best Pract Res Clin Gastroenterol*. 2012; 26(6):691–703. [PubMed: 23582913]
3. Hallet J, Law CH, Cukier M, Saskin R, Liu N, Singh S. Exploring the rising incidence of neuroendocrine tumors: a population-based analysis of epidemiology, metastatic presentation, and outcomes. *Cancer*. 2015; 121(4):589–597. [PubMed: 25312765]
4. Pape UF, Jann H, Muller-Nordhorn J, et al. Prognostic relevance of a novel TNM classification system for upper gastroenteropancreatic neuroendocrine tumors. *Cancer*. 2008; 113(2):256–265. [PubMed: 18506737]
5. Scarpa A, Mantovani W, Capelli P, et al. Pancreatic endocrine tumors: improved TNM staging and histopathological grading permit a clinically efficient prognostic stratification of patients. *Mod Pathol*. 2010; 23(6):824–833. [PubMed: 20305616]
6. Jann H, Roll S, Couvelard A, et al. Neuroendocrine tumors of midgut and hindgut origin: tumor-node-metastasis classification determines clinical outcome. *Cancer*. 2011; 117(15):3332–3341. [PubMed: 21246527]
7. Araujo PB, Cheng S, Mete O, et al. Evaluation of the WHO 2010 grading and AJCC/UICC staging systems in prognostic behavior of intestinal neuroendocrine tumors. *PLoS ONE*. 2013; 8(4):e61538. [PubMed: 23620762]
8. Merling F. Anatomie pathologique de l’appendice du caecum. *Experience (Paris)*. 1838; (1):337.
9. Langhans T. Ueber einen drüsenpolyp im ileum. *Archiv f Pathol Anat*. 1867; 38(4):559–560.
10. Lubarsch O. Ueber den primären krebs des ileum nebst bemerkungen über das gleichzeitige vorkommen von krebs und tuberculose. *Archiv f Pathol Anat*. 1888; 111(2):280–317.
11. Ransom WB. A case of primary carcinoma of the ileum. *Lancet*. 1890; ii:1020–1023.
12. Kloppel G. Oberndorfer and his successors: from carcinoid to neuroendocrine carcinoma. *Endocr Pathol*. 2007; 18(3):141–144. [PubMed: 17960501]
13. Modlin IM, Shapiro MD, Kidd M, Eick G. Siegfried oberndorfer and the evolution of carcinoid disease. *Arch Surg*. 2007; 142(2):187–197. [PubMed: 17309971]
14. Oberndorfer S. Karzinoide tumoren des dunndarms. *Frankf Z Pathol Int*. 1907; 1:425–432.
15. Williams ED, Sandler M. The classification of carcinoid tumours. *Lancet*. 1963; 1(7275):238–239. [PubMed: 14000847]
16. Soga J, Tazawa K. Pathologic analysis of carcinoids. Histologic reevaluation of 62 cases. *Cancer*. 1971; 28(4):990–998. [PubMed: 4106849]
17. Capella C, Heitz PU, Hofler H, Solcia E, Kloppel G. Revised classification of neuroendocrine tumours of the lung, pancreas and gut. *Virchows Arch*. 1995; 425(6):547–560. [PubMed: 7697211]
18. Williams ED. *Histological Typing of Endocrine Tumours*. Springer; Geneva, Switzerland: 1980.
19. Heymann MF, Joubert M, Nemeth J, et al. Prognostic and immunohistochemical validation of the capella classification of pancreatic neuroendocrine tumours: an analysis of 82 sporadic cases. *Histopathology*. 2000; 36(5):421–432. [PubMed: 10792483]

20. Capella C, La Rosa S, Solcia E. Criteria for malignancy in pancreatic endocrine tumors. *Endocr Pathol.* 1997; (8):87–90.
21. Pelosi G, Bresaola E, Bogina G, et al. Endocrine tumors of the pancreas: Ki-67 immunoreactivity on paraffin sections is an independent predictor for malignancy: a comparative study with proliferating-cell nuclear antigen and progesterone receptor protein immunostaining, mitotic index, and other clinicopathologic variables. *Hum Pathol.* 1996; 27(11):1124–1134. [PubMed: 8912819]
22. Clarke MR, Baker EE, Weyant RJ, Hill L, Carty SE. Proliferative activity in pancreatic endocrine tumors: association with function, metastases, and survival. *Endocr Pathol.* 1997; 8(3):181–187. [PubMed: 12114721]
23. Gentil Perret A, Moshier JF, Buono JP, et al. The relationship between MIB-1 proliferation index and outcome in pancreatic neuroendocrine tumors. *Am J Clin Pathol.* 1998; 109(3):286–293. [PubMed: 9495200]
24. La Rosa S, Sessa F, Capella C, et al. Prognostic criteria in nonfunctioning pancreatic endocrine tumours. *Virchows Arch.* 1996; 429(6):323–333. [PubMed: 8982376]
25. Solcia E, Klöppel G, et al. WHO Histological Typing of Endocrine Tumours. 2. Springer; Berlin, Germany: 2000.
26. Kloppel G, Perren A, Heitz PU. The gastroenteropancreatic neuroendocrine cell system and its tumors: the WHO classification. *Ann NY Acad Sci.* 2004; 1014:13–27. [PubMed: 15153416]
27. Heitz P, Komminoth P, Perren A, et al. Tumours of the Endocrine Pancreas. International Agency for Research on Cancer (IARC); Lyon, France: 2004. 177–182.
28. Madeira I, Terris B, Voss M, et al. Prognostic factors in patients with endocrine tumours of the duodenopancreatic area. *Gut.* 1998; 43(3):422–427. [PubMed: 9863490]
29. Ferrone CR, Tang LH, Tomlinson J, et al. Determining prognosis in patients with pancreatic endocrine neoplasms: can the WHO classification system be simplified? *J Clin Oncol.* 2007; 25(35):5609–5615. [PubMed: 18065733]
30. Liu TC, Hamilton N, Hawkins W, Gao F, Cao D. Comparison of WHO Classifications (2010), the Hochwald grading system, and AJCC and ENETS staging systems in predicting prognosis in locoregional well-differentiated pancreatic neuroendocrine tumors. *Am J Surg Pathol.* 2013; 37(6): 853–859. (2004). [PubMed: 23598967]
31. Schindl M, Kaczirek K, Kaserer K, Niederle B. Is the new classification of neuroendocrine pancreatic tumors of clinical help? *World J Surg.* 2000; 24(11):1312–1318. [PubMed: 11038199]
32. Bettini R, Boninsegna L, Mantovani W, et al. Prognostic factors at diagnosis and value of WHO classification in a mono-institutional series of 180 non-functioning pancreatic endocrine tumours. *Ann Oncol.* 2008; 19(5):903–908. [PubMed: 18209014]
33. Rindi G, Kloppel G, Alhman H, et al. TNM staging of foregut (neuro)endocrine tumors: a consensus proposal including a grading system. *Virchows Arch.* 2006; 449(4):395–401. [PubMed: 16967267]
34. Rindi G, Kloppel G, Couvelard A, et al. TNM staging of midgut and hindgut (neuro) endocrine tumors: a consensus proposal including a grading system. *Virchows Arch.* 2007; 451(4):757–762. [PubMed: 17674042]
35. Dolcetta-Capuzzo A, Villa V, Albarello L, et al. Gastroenteric neuroendocrine neoplasms classification: comparison of prognostic models. *Cancer.* 2013; 119(1):36–44. [PubMed: 22744914]
36. Fischer L, Kleeff J, Esposito I, et al. Clinical outcome and long-term survival in 118 consecutive patients with neuroendocrine tumours of the pancreas. *Br J Surg.* 2008; 95(5):627–635. [PubMed: 18306152]
37. Ekeblad S, Skogseid B, Dunder K, Oberg K, Eriksson B. Prognostic factors and survival in 324 patients with pancreatic endocrine tumor treated at a single institution. *Clin Cancer Res.* 2008; 14(23):7798–7803. [PubMed: 19047107]
38. La Rosa S, Klersy C, Uccella S, et al. Improved histologic and clinicopathologic criteria for prognostic evaluation of pancreatic endocrine tumors. *Hum Pathol.* 2009; 40(1):30–40. [PubMed: 18715612]

39. Srirajaskanthan R, Ahmed A, Prachialias A, et al. ENETS TNM staging predicts prognosis in small bowel neuroendocrine tumours. *ISRN Oncol.* 2013; 2013:420795. [PubMed: 23533809]
40. Ellison TA, Wolfgang CL, Shi C, et al. A single institution's 26-year experience with nonfunctional pancreatic neuroendocrine tumors: a validation of current staging systems and a new prognostic nomogram. *Ann Surg.* 2014; 259(2):204–212. [PubMed: 23673766]
41. Edge SB. *AJCC Cancer Staging Manual.* 7. Springer; NY, USA: 2010. American Joint Committee on Cancer.
42. Rindi G, Falconi M, Klersy C, et al. TNM staging of neoplasms of the endocrine pancreas: results from a large international cohort study. *J Natl Cancer Inst.* 2012; 104(10):764–777. [PubMed: 22525418]
43. Strosberg JR, Cheema A, Weber J, Han G, Coppola D, Kvols LK. Prognostic validity of a novel American Joint Committee on Cancer Staging Classification for pancreatic neuroendocrine tumors. *J Clin Oncol.* 2011; 29(22):3044–3049. [PubMed: 21709192]
44. Liszka L, Pajak J, Mrowiec S, Zielinska-Pajak E, Golka D, Lampe P. Discrepancies between two alternative staging systems (European Neuroendocrine Tumor Society 2006 and American Joint Committee on Cancer/Union for International Cancer Control 2010) of neuroendocrine neoplasms of the pancreas. A study of 50 cases. *Pathol Res Pract.* 2011; 207(4):220–224. [PubMed: 21354717]
45. Kloppel G, Rindi G, Perren A, Komminoth P, Klimstra DS. The ENETS and AJCC/UICC TNM classifications of the neuroendocrine tumors of the gastrointestinal tract and the pancreas: a statement. *Virchows Arch.* 2010; 456(6):595–597. [PubMed: 20422210]
46. Klimstra DS, Modlin IR, Coppola D, Lloyd RV, Suster S. The pathologic classification of neuroendocrine tumors: a review of nomenclature, grading, and staging systems. *Pancreas.* 2010; 39(6):707–712. [PubMed: 20664470]
47. Couvelard A. Neuroendocrine tumours of the pancreas: recent developments in staging and grading. *Diagn Pathol.* 2011; 18(1):1–7.
48. Yang M, Tian BL, Zhang Y, et al. Evaluation of the World Health Organization 2010 grading system in surgical outcome and prognosis of pancreatic neuroendocrine tumors. *Pancreas.* 2014; 43(7):1003–1008. [PubMed: 24945681]
49. Strosberg J, Nasir A, Coppola D, Wick M, Kvols L. Correlation between grade and prognosis in metastatic gastroenteropancreatic neuroendocrine tumors. *Hum Pathol.* 2009; 40(9):1262–1268. [PubMed: 19368957]
50. Brown DC, Gatter KC. Monoclonal-antibody Ki-67 – its use in histopathology. *Histopathology.* 1990; 17(6):489–503. [PubMed: 2076881]
51. Klimstra DS, Modlin IR, Adsay NV, et al. Pathology reporting of neuroendocrine tumors: application of the Delphic consensus process to the development of a minimum pathology data set. *Am J Surg Pathol.* 2010; 34(3):300–313. [PubMed: 20118772]
52. Yang Z, Tang LH, Klimstra DS. Effect of tumor heterogeneity on the assessment of Ki67 labeling index in well-differentiated neuroendocrine tumors metastatic to the liver: implications for prognostic stratification. *Am J Surg Pathol.* 2011; 35(6):853–860. [PubMed: 21566513]
53. Lopez F, Belloc F, Lacombe F, et al. Modalities of synthesis of Ki67 antigen during the stimulation of lymphocytes. *Cytometry.* 1991; 12(1):42–49. [PubMed: 1999122]
54. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. *J Clin Oncol.* 2005; 23(28):7212–7220. [PubMed: 16192605]
55. Arima N, Nishimura R, Osako T, et al. The importance of tissue handling of surgically removed breast cancer for an accurate assessment of the Ki-67 index. *J Clin Pathol.* 2015; 69(3):255–259. [PubMed: 26420767]
56. Mengel M, Von Wasielewski R, Wiese B, Rudiger T, Muller-Hermelink HK, Kreipe H. Inter-laboratory and inter-observer reproducibility of immunohistochemical assessment of the Ki-67 labelling index in a large multi-centre trial. *J Pathol.* 2002; 198(3):292–299. [PubMed: 12375261]
57. Blank A, Wehweck L, Marinoni I, et al. Interlaboratory variability of MIB1 staining in well-differentiated pancreatic neuroendocrine tumors. *Virchows Arch.* 2015; 467(5):543–550. [PubMed: 26384025]

58. Polley MY, Leung SC, Mcshane LM, et al. An international Ki67 reproducibility study. *J Natl Cancer Inst.* 2013; 105(24):1897–1906. [PubMed: 24203987]
59. Tang LH, Gonen M, Hedvat C, Modlin IM, Klimstra DS. Objective quantification of the Ki67 proliferative index in neuroendocrine tumors of the gastroenteropancreatic system: a comparison of digital image analysis with manual methods. *Am J Surg Pathol.* 2012; 36(12):1761–1770. [PubMed: 23026928]
60. Young HT, Carr NJ, Green B, Tilley C, Bhargava V, Pearce N. Accuracy of visual assessments of proliferation indices in gastroenteropancreatic neuroendocrine tumours. *J Clin Pathol.* 2013; 66(8): 700–704. [PubMed: 23703851]
61. Reid MD, Bagci P, Ohike N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies. *Mod Pathol.* 2015; 28(5):686–694. [PubMed: 25412850]
62. Ong CW, Kim LG, Kong HH, et al. Computer-assisted pathological immunohistochemistry scoring is more time-effective than conventional scoring, but provides no analytical advantage. *Histopathology.* 2010; 56(4):523–529. [PubMed: 20459559]
63. Remes SM, Tuominen VJ, Helin H, Isola J, Arola J. Grading of neuroendocrine tumors with Ki-67 requires high-quality assessment practices. *Am J Surg Pathol.* 2012; 36(9):1359–1363. [PubMed: 22895268]
64. Couvelard A, Deschamps L, Ravaud P, et al. Heterogeneity of tumor prognostic markers: a reproducibility study applied to liver metastases of pancreatic endocrine tumors. *Mod Pathol.* 2009; 22(2):273–281. [PubMed: 18997736]
65. Zen Y, Heaton N. Elevated Ki-67 labeling index in ‘synchronous liver metastases’ of well differentiated enteropancreatic neuroendocrine tumor. *Pathol Int.* 2013; 63(11):532–538. [PubMed: 24274715]
66. Miller HC, Drymoussis P, Flora R, Goldin R, Spalding D, Frilling A. Role of Ki-67 proliferation index in the assessment of patients with neuroendocrine neoplasias regarding the stage of disease. *World J Surg.* 2014; 38(6):1353–1361. [PubMed: 24493070]
67. Shi C, Gonzalez RS, Zhao Z, et al. Liver metastases of small intestine neuroendocrine tumors: Ki-67 heterogeneity and World Health Organization grade discordance with primary tumors. *Am J Clin Pathol.* 2015; 143(3):398–404. [PubMed: 25696798]
68. Fung AD, Cohen C, Kavuri S, Lawson D, Gao X, Reid MD. Phosphohistone H3 and Ki-67 labeling indices in cytologic specimens from well-differentiated neuroendocrine tumors of the gastrointestinal tract and pancreas: a comparative analysis using automated image cytometry. *Acta Cytol.* 2013; 57(5):501–508. [PubMed: 24021213]
69. Voss SM, Riley MP, Lokhandwala PM, Wang M, Yang Z. Mitotic count by phosphohistone H3 immunohistochemical staining predicts survival and improves interobserver reproducibility in well-differentiated neuroendocrine tumors of the pancreas. *Am J Surg Pathol.* 2015; 39(1):13–24. [PubMed: 25353284]
70. Bosman FT. World Health Organization, International Agency for Research on Cancer. WHO Classification of Tumours of the Digestive System. 4. International Agency for Research on Cancer; Lyon, France: 2010.
71. McCall CM, Shi C, Cornish TC, et al. Grading of well-differentiated pancreatic neuroendocrine tumors is improved by the inclusion of both Ki67 proliferative index and mitotic rate. *Am J Surg Pathol.* 2013; 37(11):1671–1677. [PubMed: 24121170]
72. Sorbye H, Strosberg J, Baudin E, Klimstra DS, Yao JC. Gastroenteropancreatic high-grade neuroendocrine carcinoma. *Cancer.* 2014; 120(18):2814–2823. [PubMed: 24771552]
73. Kunz PL, Reidy-Lagunes D, Anthony LB, et al. Consensus guidelines for the management and treatment of neuroendocrine tumors. *Pancreas.* 2013; 42(4):557–577. [PubMed: 23591432]
74. Pavel M, Baudin E, Couvelard A, et al. ENETS Consensus Guidelines for the management of patients with liver and other distant metastases from neuroendocrine neoplasms of foregut, midgut, hindgut, and unknown primary. *Neuroendocrinology.* 2012; 95(2):157–176. [PubMed: 22262022]
75. Strosberg JR, Coppola D, Klimstra DS, et al. The NANETS consensus guidelines for the diagnosis and management of poorly differentiated (high-grade) extrapulmonary neuroendocrine carcinomas. *Pancreas.* 2010; 39(6):799–800. [PubMed: 20664477]

76. Velayoudom-Cephise FL, Duvillard P, Foucan L, et al. Are G3 ENETS neuroendocrine neoplasms heterogeneous? *Endocr Relat Cancer*. 2013; 20(5):649–657. [PubMed: 23845449]
77. Sorbye H, Welin S, Langer SW, et al. Predictive and prognostic factors for treatment and survival in 305 patients with advanced gastrointestinal neuroendocrine carcinoma (WHO G3): the NORDIC NEC study. *Ann Oncol*. 2013; 24(1):152–160. [PubMed: 22967994]
78. Basturk O, Tang L, Hruban RH, et al. Poorly differentiated neuroendocrine carcinomas of the pancreas: a clinicopathologic analysis of 44 cases. *Am J Surg Pathol*. 2014; 38(4):437–447. [PubMed: 24503751]
79. Basturk O, Yang Z, Tang LH, et al. The high-grade (WHO G3) pancreatic neuroendocrine tumor category is morphologically and biologically heterogeneous and includes both well differentiated and poorly differentiated neoplasms. *Am J Surg Pathol*. 2015; 39(5):683–690. [PubMed: 25723112]
80. Tang LH, Untch BR, Reidy DL, et al. Well differentiated neuroendocrine tumors with a morphologically apparent high grade component: a pathway distinct from poorly differentiated neuroendocrine carcinomas. *Clin Cancer Res*. 2016; 22(4):1011–1017. [PubMed: 26482044]
81. Tang LH, Basturk O, Sue JJ, Klimstra DS. A practical approach to the classification of WHO grade 3 (G3) well differentiated neuroendocrine tumor (WD-NET) and poorly differentiated neuroendocrine carcinoma (PD-NEC) of the pancreas. *Am J Surg Pathol*. 2016 (Epub ahead of print).
82. Yachida S, Vakiani E, White CM, et al. Small cell and large cell neuroendocrine carcinomas of the pancreas are genetically similar and distinct from well-differentiated pancreatic neuroendocrine tumors. *Am J Surg Pathol*. 2012; 36(2):173–184. [PubMed: 22251937]
83. Rindi G, Wiedenmann B. Neuroendocrine neoplasms of the gut and pancreas: new insights. *Nat Rev Endocrinol*. 2012; 8(1):54–64.
84. Kloppel G, Rindi G, Anlauf M, Perren A, Komminoth P. Site-specific biology and pathology of gastroenteropancreatic neuroendocrine tumors. *Virchows Arch*. 2007; 451(Suppl 1):S9–S27. [PubMed: 17684761]
85. Niederle MB, Hackl M, Kaserer K, Niederle B. Gastroenteropancreatic neuroendocrine tumours: the current incidence and staging based on the WHO and European Neuroendocrine Tumour Society classification: an analysis based on prospectively collected parameters. *Endocr Relat Cancer*. 2010; 17(4):909–918. [PubMed: 20702725]
86. Kloppel G. Classification and pathology of gastroenteropancreatic neuroendocrine neoplasms. *Endocr Relat Cancer*. 2011; 18(Suppl 1):S1–S16. [PubMed: 22005112]
87. Garcia-Carbonero R, Capdevila J, Crespo-Herrero G, 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):1794–1803. [PubMed: 20139156]
88. Moyana TN, Xiang J, Senthilselvan A, Kulaga A. The spectrum of neuroendocrine differentiation among gastrointestinal carcinoids: importance of histologic grading, MIB-1, p53, and bcl-2 immunoreactivity. *Arch Pathol Lab Med*. 2000; 124(4):570–576. [PubMed: 10747315]
89. Modlin IM, Kidd M, Latich I, Zikusoka MN, Shapiro MD. Current status of gastrointestinal carcinoids. *Gastroenterology*. 2005; 128(6):1717–1751. [PubMed: 15887161]
90. Panzuto F, Nasoni S, Falconi M, et al. Prognostic factors and survival in endocrine tumor patients: comparison between gastrointestinal and pancreatic localization. *Endocr Relat Cancer*. 2005; 12(4):1083–1092. [PubMed: 16322345]
91. Perren A, Anlauf M, Komminoth P. Molecular profiles of gastroenteropancreatic endocrine tumors. *Virchows Arch*. 2007; 451(Suppl 1):S39–S46. [PubMed: 17684763]
92. Oberg K. Genetics and molecular pathology of neuroendocrine gastrointestinal and pancreatic tumors (gastroenteropancreatic neuroendocrine tumors). *Curr Opin Endocrinol Diabetes Obes*. 2009; 16(1):72–78. [PubMed: 19115524]
93. Lubensky IA, Zhuang Z. Molecular genetic events in gastrointestinal and pancreatic neuroendocrine tumors. *Endocr Pathol*. 2007; 18(3):156–162. [PubMed: 18041590]

94. Jiao Y, Shi C, Edil BH, et al. DAXX/ATRAX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science*. 2011; 331(6021):1199–1203. [PubMed: 21252315]
95. Faggiano A, Mansueto G, Ferolla P, et al. Diagnostic and prognostic implications of the World Health Organization classification of neuroendocrine tumors. *J Endocrinol Invest*. 2008; 31(3): 216–223. [PubMed: 18401203]
96. Ricci C, Casadei R, Taffurelli G, et al. WHO 2010 classification of pancreatic endocrine tumors. Is the new always better than the old? *Pancreatol*. 2014; 14(6):539–541. [PubMed: 25266640]
97. Ricci C, Casadei R, Taffurelli G, et al. Validation of the 2010 WHO classification and a new prognostic proposal: A single centre retrospective study of well-differentiated pancreatic neuroendocrine tumours. *Pancreatol*. 2016; 16(3):403–410. [PubMed: 26924664]
98. Hamilton NA, Liu TC, Cavataio A, et al. Ki-67 predicts disease recurrence and poor prognosis in pancreatic neuroendocrine neoplasms. *Surgery*. 2012; 152(1):107–113. [PubMed: 22503317]
99. Lowe K, Khithani A, Liu E, et al. Ki-67 labeling: a more sensitive indicator of malignant phenotype than mitotic count or tumor size? *J Surg Oncol*. 2012; 106(6):724–727. [PubMed: 22549809]
100. Goodell PP, Krasinskas AM, Davison JM, Hartman DJ. Comparison of methods for proliferative index analysis for grading pancreatic well-differentiated neuroendocrine tumors. *Am J Clin Pathol*. 2012; 137(4):576–582. [PubMed: 22431534]
101. Boninsegna L, Panzuto F, Partelli S, et al. Malignant pancreatic neuroendocrine tumour: lymph node ratio and Ki67 are predictors of recurrence after curative resections. *Eur J Cancer*. 2012; 48(11):1608–1615. [PubMed: 22129889]
102. Khan MS, Luong TV, Watkins J, Toumpanakis C, Caplin ME, Meyer T. A comparison of Ki-67 and mitotic count as prognostic markers for metastatic pancreatic and midgut neuroendocrine neoplasms. *Br J Cancer*. 2013; 108(9):1838–1845. [PubMed: 23579216]
103. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med*. 2004; 66(3):411–421. [PubMed: 15184705]
104. Concato J, Peduzzi P, Holford TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I Background, goals, and general strategy. *J Clin Epidemiol*. 1995; 48(12):1495–1501. [PubMed: 8543963]
105. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II Accuracy and precision of regression estimates. *J Clin Epidemiol*. 1995; 48(12):1503–1510. [PubMed: 8543964]
106. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15(4): 361–387. [PubMed: 8668867]
107. Irwin JR, McClelland GH. Negative consequences of dichotomizing continuous predictor variables. *J Mark Res*. 2003; 40(3):366–371.
108. Maccallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. *Psychol Methods*. 2002; 7(1):19–40. [PubMed: 11928888]
109. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006; 25(1):127–141. [PubMed: 16217841]
110. Weinberg CR. How bad is categorization? *Epidemiology*. 1995; 6(4):345–347. [PubMed: 7548338]
111. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using “optimal” cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst*. 1994; 86(11):829–835. [PubMed: 8182763]
112. Hochwald SN, Zee S, Conlon KC, et al. Prognostic factors in pancreatic endocrine neoplasms: an analysis of 136 cases with a proposal for low-grade and intermediate-grade groups. *J Clin Oncol*. 2002; 20(11):2633–2642. [PubMed: 12039924]
113. Fahy BN, Tang LH, Klimstra D, et al. Carcinoid of the rectum risk stratification (CaRRs): a strategy for preoperative outcome assessment. *Ann Surg Oncol*. 2007; 14(5):1735–1743. [PubMed: 17294074]

114. Janson ET, Sorbye H, Welin S, et al. Nordic guidelines 2014 for diagnosis and treatment of gastroenteropancreatic neuroendocrine neoplasms. *Acta Oncol.* 2014; 53(10):1284–1297. [PubMed: 25140861]
115. Oberg K, Knigge U, Kwekkeboom D, Perren A. Neuroendocrine gastroentero-pancreatic tumors: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2012; 23(Suppl 7):vii124–vii130. [PubMed: 22997445]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Practice points

- NENs are subdivided into well and poorly differentiated categories; well-differentiated NETs are a heterogeneous subgroup.
- Proliferative activity (mitotic rate or Ki-67 index) is a significant prognostic factor and is an integral part of more recent classification systems.
- The ENETS proposed separate grading and staging systems in 2006 and 2007; the same grading criteria were later endorsed by the WHO in 2010, giving rise to the ENETS/WHO grading system.
- Grading is based on Ki-67 and mitotic activity, whichever is highest.
- Determining Ki-67 proliferative index in 'hotspots' is more predictive than a random approach.
- Manual cell counting, preferably utilizing printed images, is currently the most accurate, practical and reproducible method for determining Ki-67 and is the recommended approach.
- Evidence has shown that the current WHO G3 category is comprised of two biologically distinct subgroups: (1) well-differentiated NETs with a high proliferative rate (grade-discordant WDNets or G3 WDNets) and (2) true poorly differentiated NECs (small-cell or large-cell types); G3 WDNets behave better than *bona fide* PDNECs and somewhat worse than grade concordant G2 WDNets.
- G3 WDNets may present as a morphologically well-differentiated tumor with G2 mitotic rate but G3 Ki-67 index or as a well-differentiated tumor with distinct high grade areas (morphologically high grade and with an elevated Ki-67 index).
- Data suggesting a modification to the G1/G2 cut-point are still scarce and require further supportive evidence before implementation.

Figure 1. Well-differentiated neuroendocrine tumors

They are characterized by a wide array of architectural patterns, such as gyriform (A), nested/organoid, cribriform (B), trabecular and tubular/glandular. Cytologically, they typically display 'salt and pepper' chromatin and finely granular cytoplasm. Mitotic index is low in these examples of low-grade NETs. (H&E, 200× magnification).

Figure 2. Poorly differentiated neuroendocrine carcinoma

(A) Small-cell type is characterized by solid/trabecular pattern of growth and cells with scant cytoplasm, inconspicuous nucleoli and nuclear molding. Note the mitoses and apoptotic bodies. An area of necrosis is partially represented in the upper left corner. (B) Large-cell type showing solid/trabecular architecture and cells with abundant basophilic cytoplasm and open chromatin with large nucleoli. Numerous mitotic figures and apoptotic bodies are again easily identified. (H&E, 200× magnification).

Figure 3. Well-differentiated neuroendocrine tumor with a high-grade component, one of the presentations of G3 WNET

(A) The tumor has areas with typical morphology of a WNET exhibiting gyriform architecture (center). However, distinctly separate areas display a solid pattern of growth (remainder of tumor lobules) with extensive areas of necrosis (not shown). (B) Corresponding Ki-67 highlights the difference of Ki-67 labeling in the well-differentiated component (within the G2 range) and the solid areas (Ki-67 >90%). (H&E and Ki-67 immunohistochemistry, 40× magnification). (C) Higher magnification of the well-differentiated area in (A), with typical architecture and a paucity of mitoses, and of the (D) high-grade component displaying a solid pattern, greater degree of cytologic atypia, numerous mitoses, and areas of necrosis. (H&E, 200× magnification).

Table 1

The evolution of classifications of neuroendocrine neoplasms of the gastroenteropancreatic system.

WHO 1980	Capella 1995	WHO 2000/2004	WHO 2010
Carcinoid: • EC cell • G cell • Others	Neuroendocrine tumor: • Benign behavior • Uncertain behavior • Low-grade malignant • High-grade malignant	Neuroendocrine neoplasm: • Well-differentiated neuroendocrine tumor – benign or uncertain malignant potential • Well-differentiated neuroendocrine carcinoma • Poorly differentiated neuroendocrine carcinoma/small-cell carcinoma	Neuroendocrine tumor: • Grade 1 – G1 (carcinoid) • Grade 2 – G2 Neuroendocrine carcinoma (small-cell or large-cell type): • Grade 3 – G3
Mucocarcinoid		Mixed exocrine-endocrine carcinoma	Mixed adenoneuroendocrine carcinoma
Mixed carcinoid-adenocarcinoma			
Pseudotumor lesions		Tumor-like lesions	Hyperplastic and preneoplastic lesions

EC cell: Enterochromaffin cell; G cell: Gastrin cell.

Table 2

Definitions for neuroendocrine neoplasms.

Term	Definition
Tumor differentiation	Degree a tumor resembles its normal tissue counterpart; based exclusively on morphologic criteria
Grade	Inherent tumor aggressiveness; largely based on proliferative rate of a tumor (mitotic count and/or Ki-67 labeling index)
Stage	Extent of disease at the time of diagnosis

Table 3

ENETS 2006/2007 grading proposal later endorsed by the WHO 2010 classification.

Grade	Mitotic count (per 10 HPF)	Ki-67 (%)
G1 (low)	<2	2
G2 (intermediate)	2–20	3–20
G3 (high)	>20	20

G1: Grade 1; G2: Grade 2; G3: Grade 3; HPF: High power fields.

Disease-specific survival rates of neuroendocrine neoplasms from diverse locations in relation to grade according to different studies.

Table 4

Study (year)	Nature of tumors	Location	Number of cases [†]	5-year survival (%)	10-year survival (%)	Ref.
Pape <i>et al.</i> (2008)	Primary	Foregut	158	G1: 95.7 G2: 73.4 G3: 27.7	G1: 83.7 G2: 69.4 G3: ND	[4]
Scarpa <i>et al.</i> (2010)	Primary	Pancreatic	237	G1: 90 G2: 63 G3: 12	G1: 80 G2: 46 G3: –	[5]
Jann <i>et al.</i> (2011)	Primary	Midgut and hindgut	189	G1: 95.2 G2: 82 G3: 51.4	G1: 80.3 G2: 66.2 G3: 34.3	[6]
Araujo <i>et al.</i> (2013)	Primary	Intestinal	77	G1: 92 G2: 92 G3: – [‡]	G1: 84 G2: 42 G3: – [‡]	[7]

All studies used grade as defined in Table 2.

[†]Number of cases included for each study includes only cases in which grading was possible and not necessarily entire cohort.

[‡]Not analyzable for survival rates due to small case numbers.

G1: Grade 1; G2: Grade 2; G3: Grade 3; ND: Not determined.

Table 5

Limitations of current WHO/ENETS grading system and recommendations.

Limitations	Recommendations
Issues with Ki-67	
Preanalytical factors: <ul style="list-style-type: none"> • Type of fixative • Fixation times • Different antibodies/clones used • Diverse staining protocols 	Standardization of Ki-67 immunohistochemical staining
<hr/>	
Analytical factors:	
<ul style="list-style-type: none"> • Diverse counting methods 	Manual counting (from printed image) of positive cells in 'hotspots'
<ul style="list-style-type: none"> • No definition for positivity (any intensity staining or only strong staining) 	Consider cells with any intensity of staining as positive
<ul style="list-style-type: none"> • Tumor heterogeneity 	Adequate tumor sampling on biopsies
<hr/>	
Heterogeneity of the G3 category	New category of G3 well-differentiated NETs must be incorporated and distinguished from PDNECs
<hr/>	
Site-specific biologic differences	Site-specific Ki-67 cut-points must be determined in future studies

PDNEC: Poorly differentiated neuroendocrine carcinoma.

Table 6

Studies that have tested and/or proposed different Ki-67 cut-point values.

Study (year)	Anatomic site	Sample size [‡]	G1/G2 Ki-67 cut-point (%)	G2/G3 Ki-67 cut-point (%)	Ref.
Pelosi <i>et al.</i> (1996)	Pancreas	54	5	–	[21]
Gentil Perret <i>et al.</i> (1998)	Pancreas	35	4	–	[23]
Faggiano <i>et al.</i> (2008)	GI, lung, ovary, unknown primary	31	15	–	[95]
Bettini <i>et al.</i> (2008)	Pancreas	180	5	–	[32]
La Rosa <i>et al.</i> (2009)	Pancreas	155	2.5	15	[38]
Scarpa <i>et al.</i> (2010)	Pancreas	237	5	20	[5]
Goodell <i>et al.</i> (2012) [‡]	Pancreas (WDNET only)	45	7.5	–	[100]
Boninsegna <i>et al.</i> (2012)	Pancreas (WDNET only)	57	5	–	[101]
Rindi <i>et al.</i> (2012)	Pancreas	892	4.85	–	[42]
Hamilton <i>et al.</i> (2012)	Pancreas	140	5	9	[98]
Lowe <i>et al.</i> (2012)	Pancreas	24	2	10	[99]
Khan <i>et al.</i> (2013)	Pancreas and midgut	267	5	20	[102]

[‡] Sample size for each study includes only cases in which Ki-67 assessment was possible and not necessarily entire cohort.

[‡] Used a computer-assisted system to assess Ki-67.

WDNET: Well-differentiated neuroendocrine tumor.