BMJ Open

Risk factors for neuroendocrine neoplasms: a protocol for a case-control study based on a record linkage of registry and claims data

SCHOLARONE™
Manuscripts

1 **STUDY PROTOCOL**

2

3 **Risk factors for neuroendocrine neoplasms: a protocol for a case-control study based**

4 **on a record linkage of registry and claims data**

5

6 Sven Voigtländer,[1] Roman Gerlach,[2] Nina Grundmann,[1] Ewan Donnachie,[2] Ursula Berger,[3]

7 Amir Hakimhashemi,[1] Martin Meyer,[1] Martin Tauscher,[2] Jacqueline Müller-Nordhorn[1]

8

9 [1] Bavarian Cancer Registry, Bavarian Health and Food Safety Authority, Schweinauer

10 Hauptstraße 80, 90441 Nuremberg, Germany

11 [2] Bavarian Association of Statutory Health Insurance Accredited Physicians,

12 Elsenheimerstraße 39, 80687 Munich, Germany

13 [3] Institute for Medical Information Processing, Biometry, and Epidemiology, Faculty of

14 Medicine at Ludwig-Maximillians-University Munich, Marchioninistraße 15, 81377 Munich,

15 Germany

16

17 Word count (without title page, abstract, statements, and references):

18 2,547

19

20 **Correspondence:**

21 Dr. Sven Voigtländer

22 Bavarian Cancer Registry, Bavarian Health and Food Safety Authority

23 Schweinauer Hauptstraße 80, 90441 Nuremberg, Germany

24 Phone: +49-(0)9131-6808-2927

25 Email: sven.voigtlaender@lgl.bayern.de

26 ORCID ID: https://orcid.org/0000-0003-3808-9605

27

28

29

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

30 **ABSTRACT**

31 **Background:**

32 Recent studies showed an increase in neuroendocrine neoplasms, especially for the

33 digestive tract. Several risk factors have been suggested to explain this increase including a

34 family history of cancer, tobacco smoking, alcohol consumption, and metabolic disorders

35 including diabetes and obesity. Another risk factor may be depressive disorders, which could

36 increase the risk of neuroendocrine neoplasms either directly or mediated through

37 associated risk behaviours and/or antidepressant medication. Here, we outline the design of

38 our study to identify risk factors for neuroendocrine neoplasms in Germany.

39

40 **Methods and analysis:**

41 A case-control study of the resident population of Bavaria, the second most populous federal

42 state in Germany, based on a record linkage of data from the Bavarian Cancer Registry and

43 data from the Bavarian Association of Statutory Health Insurance Accredited Physicians.

44 Cases have a diagnosis of a malignant neuroendocrine neoplasm, either of the

45 bronchopulmonary system or the gastroenteropancreatic system, in 2020 or 2021. Controls

46 are sampled from the noncases and matched on sex, birth year (in 5-year intervals), and

47 time of diagnosis (by calendar quarter). Risk factor prevalence of cases and controls is

48 assessed on the basis of assured outpatient diagnoses, i.e. diagnoses documented in at

49 least two out of four consecutive quarters in the eight quarters preceding the diagnosis of a

50 neuroendocrine neoplasm. The analysis uses conditional logistic regression to estimates

51 odds ratios and 95% confidence intervals.

52

53 **Ethics and dissemination:**

54 This study protocol was approved by the Ethics Committee of the Bavarian State Chamber of

55 Physicians (reference number: 24008). Approval by supervisory authority has been obtained

56 from the Bavarian State Ministry of Health, Care, and Prevention (reference number: G35h-

57 A1080-2023/20-2) and also the Bavarian Data Protection Commissioner stated to have no

58 concerns after presentation of the study protocol (reference number: DSB/7-692/1-275). The

59 results of the case-control study will be presented at national as well as international

60 conferences and be published in the form of scientific articles in peer-reviewed journals.

61

62 **Registration details:**

63 ClinicalTrials.gov, NCT: 06282016.

64

65 **Keywords:**

2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

66   Neuroendocrine neoplasms, colorectal cancer, risk factors, case-control study, cancer

67   registry, claims data

68

69

70

3

ARTICLE SUMMARY

Strengths and limitations of this study

- Cases and controls will be drawn from an unselected source population, i.e. persons, who were insured with the Statutory Health Insurance (85% of all residents) and had at least one outpatient physician contact in 2020 or 2021 within Bavaria
- Cases will comprise all malignant neuroendocrine neoplasms of the gastroenteropancreatic and bronchopulmonary system in the source population in 2020 and 2021 based on high-quality cancer registry data
- The study will measure exposures based on cancer registry data and assured ICD-10 diagnoses and is, therefore, not prone to recall bias
- The study will, different to previous studies, also stratify for neuroendocrine tumours and neuroendocrine carcinomas based on high-quality cancer registry data.
- One limitation is that operationalisation of exposures is limited to exposures, which can be measured based on cancer registry as well claims data

4

88   **BACKGROUND**

89   Neuroendocrine neoplasms (NEN) are malignancies of neuroendocrine cells.[1]

90   Neuroendocrine cells can be found throughout the body, i.e. where there is epithelium,

91   excluding the central nervous system, bones, and soft tissue. NEN are rare and

92   heterogeneous tumours, which most commonly arise in the gastroenteropancreatic system

93   (GEPS) and the bronchopulmonary system (BPS).[1-3] NEN are categorised according to their

94   differentiation into differentiated neuroendocrine tumours (NET) and poorly differentiated

95   carcinomas (NEC).[4] NET can be further subdivided based on grade while NEC are further

96   categorised into small cell and large cell carcinomas. In addition, there are also mixed

97   neuroendocrine-non-neuroendocrine neoplasms (MiNEN).

98

99   Recent studies from the United States (US) and Germany showed an increase in NEN,

100 especially for the digestive tract.[5-9] According to an analysis of data from the US Surveillance,

101 Epidemiology and End Results (SEER) program, age-standardised incidence rates (ASR, US

102 standard population 2000) significantly increased for most gastro-intestinal sites between

103 1975 and 2008.[7] For instance, the age-adjusted incidence rate for rectal NEN increased from

104 0.1 in 1975 to 1.1 per 100,000 persons in 2008.[7] The risk for NEN is known to increase with

105 age.[2] Several other risk factors have been suggested by a limited number of studies

106 including a family history of cancer, tobacco smoking, alcohol consumption, and metabolic

107 disorders including diabetes and obesity.[2 10] Apart from this, there may be other risk factors,

108 which have not been confirmed yet, such as depressive disorders, which could increase the

109 risk of neuroendocrine neoplasms either directly or mediated through associated risk

110 behaviours and antidepressant medication. Kenner,[11] for instance, discusses the role of

111 depressive disorders for pancreatic cancer.

112

113 The aim of this study is to identify risk factors for neuroendocrine neoplasms based on data

114 of the Bavarian Cancer Registry, Germany, to help understand the increase in

115 neuroendocrine neoplasms.

116

117 **METHODS**

118 **Study design**

119 The study design is a case-control study based on data from the population-based Bavarian

120 Cancer Registry and data from the Bavarian Association of Statutory Health Insurance

121 Accredited Physicians (KVB, German: *Kassenärztliche Vereinigung Bayerns*). The Bavarian

122 Cancer Registry is based on mandatory notifications by physicians and health care providers

123 regarding the diagnosis and treatment of cancer.[12] The KVB collects, on a quarterly basis,

124 diagnosis and treatment data related to its main task, i.e. ensuring and reimbursing

5

125 outpatient medical care and psychotherapy for patients with Statutory Health Insurance

126 (GKV, German: *Gesetzliche Krankenversicherung*) in Bavaria. Bavarian Cancer Registry

127 data and KVB data are linked by pseudonymised record linkage following the probability

128 linkage procedure established by the German cancer registries.[13][14] The probability linkage is

129 based on the conversion of KVB identity data into unique tokens, i.e. pseudonymisation, that

130 are compared with the tokens already present in the Bavarian Cancer Registry by

131 probabilistic linkage.[14] After linkage, the pseudonyms are removed resulting in an

132 anonymous dataset.

133

134 **Setting**

135 The setting is the resident population of Bavaria in the years 2020 and 2021. Bavaria is the

136 second most populous federal state in Germany with about 13 million residents. Though the

137 Bavarian Cancer Registry covers the complete resident population of Bavaria, the source

138 population of the study is limited to persons, who are insured with the GKV and had at least

139 one outpatient physician contact in 2020 or 2021 within Bavaria. Persons who are not

140 insured in the GKV as well as persons without outpatient physician contact in 2020 and 2021

141 within Bavaria are not included in the KVB data. In Bavaria, about 11.5 million residents,

142 which is 85% of all residents, are insured with the GKV.[15]

143

144 **Definition and recruitment of cases and controls**

145 Eligible cases are defined as persons with a malignant NEN of the BPS or GEPS diagnosed

146 in 2020 or 2021. Definition of malignant NEN as well as of BPS and GEPS is based on the

147 fifth edition of the World Health Organization's (WHO) Classification of Tumours, also known

148 as the WHO Blue Books,[4][16] and the International Statistical Classification of Diseases and

149 Related Health Problems, 10th Revision (ICD-10),[17] respectively. Detailed information on the

150 eligible combinations of histology code and tumour site is provided in the Supplement

151 (Supplement 1). The selection of cases is based on the record linkage of Bavarian Cancer

152 Registry and KVB data, i.e. persons with a malignant NEN in 2020 or 2021 according to the

153 Bavarian Cancer Registry data that had a GKV insurance as well as an outpatient physician

154 contact in 2020 or 2021. Eligibility is not limited to age.

155

156 Eligible controls are defined as noncases in the same source population as for cases, i.e. the

157 KVB data. Selection of controls is done by random sampling.

158

159 Cases will be matched 1:2 with controls on sex, birth year (5-year intervals), and time of

160 diagnosis (by calendar quarter). Two controls will be matched to each case.

161

6

Definition of outcomes, exposures, and confounders

<u>Outcomes</u>

Primary outcomes are NEN of the BPS and the GEPS as defined by eligible combinations of histology code and tumour site (see Supplement 1). Secondary outcomes are the two subcategories of NEN, i.e. NET and NEC.

<u>Exposures and potential confounders</u>

The following exposures and potential confounders will be assessed:

Exposures: disease-related exposures

1. Depression or persistent mood disorder
2. Obesity
3. Metabolic syndrome
4. Previous malignant neoplasms
5. Alcohol abuse (only for NEN of the GEPS)
6. Diabetes (only for NEN of the GEPS)
7. Ulcerative colitis (only for NEN of the GEPS)
8. Crohn disease (only for NEN of the GEPS)
9. Allergies (only for NEN of the BPS)
10. Asthma (only for NEN of the BPS)

Exposures: sociodemographic

1. Area deprivation of the residence municipality

Potential confounders: sociodemographic

1. Birth year (5-year intervals)
2. Sex
3. Rurality of the residence district

Potential confounders: time of NEN diagnosis

1. Time of diagnosis (calendar quarter)

Measurement of outcomes, exposures, and confounders

Outcomes are measured using the ICD-10 as well as histological information in the data of the Bavarian Cancer Registry (see Supplement 1). Previous malignant neoplasms, birth year (5-year intervals), sex, and time of diagnosis (by calendar quarter) are also taken from the BKR data. Previous malignant neoplasms includes all neoplasms with ICD-10 codes C00-

7

199 C97 (malignant neoplasms) except C44 (other malignant neoplasms of skin).[17] For cases

200 and controls, the prevalence of the disease-related exposures, apart from previous malignant

201 neoplasms, is assessed by specific ICD-10 diagnoses based on the KVB data (see

202 Supplement 2 for further details). Only assured ICD-10 diagnoses will be considered,[18] i.e.

203 diagnoses that were recorded in at least two out of four consecutive quarters in the eight

204 quarters preceding the quarter of the NEN diagnoses for cases as well as controls, which are

205 matched by quarter of NEN diagnosis. Area deprivation of the residence municipality of

206 cases and controls is measured by the Bavarian Index of Multiple Deprivation, which is

207 based on official data.[19] In particular, we use the deprivation quintile of the residence

208 municipality at the time of NEN diagnosis. Rurality of the residence district at the time of NEN

209 diagnosis is based on the dichotomous categorisation of districts into urban area and rural

210 area by the German Federal Institute for Research on Building, Urban Affairs and Spatial

211 Development.[20]

212

213 **Bias**

214 The case-control study is based on a record linkage of registry data, claims data, and

215 administrative data. These data sources are, unlike survey data, not prone to recall bias. The

216 potential of selection bias is considerably reduced by a small number of matching variables,

217 i.e. birth year, sex, and time of NEN diagnosis, as well as the use of an almost unselected

218 source population for cases and controls, i.e. the KVB data. The KVB data covers all persons

219 with GKV (about 85% of all residents in Bavaria) and at least one outpatient physician

220 contact in Bavaria within the study period of 2020-2021. According to the KVB, 94% of all

221 GKV insured persons had at least one contact with a general practitioner in 2021,[21] so that

222 about 80% of all residents are included in our source population. It is known that the

223 proportion of persons with GKV, who use outpatient care, is higher for women compared to

224 men and lower for younger persons compared to older persons.[18] Regarding potential

225 differential misclassification bias, assessment of disease-related exposures is done in the

226 eight quarters preceding the quarter of the NEN diagnosis to ensure that the prevalence and

227 number of disease-related exposures is not influenced by potentially increased clinical

228 investigation related to the NEN diagnosis.

229

230 Confounding is controlled by matching for birth year, sex, and time of NEN diagnosis.

231 Conditional logistic regression will be employed to account for matching of controls to cases.

232 Furthermore, differences in existing infrastructure of outpatient care between urban areas

233 and rural areas may be associated with the prevalence of diagnosis-related exposures. This

234 potential confounding is additionally controlled in sensitivity analyses by adjusting models for

235 the rurality of the residence district.

8

236

Study size

238 Based on the data of the population-based Bavarian Cancer Registry (until 21st June 2024),

239 the number of incident NEN cases in Bavaria, Germany, is 3,935 cases for the study period

240 2020 and 2021, of which 2,258 were NEN of the BPS and 1,677 were NEN of the GEPS.

241 Taking into account that the KVB data covers 85% of all residents in Bavaria, Germany,

242 3,345 cases may be expected at maximum after the record linkage of Bavarian Cancer

243 Registry and KVB data. As not all residents with incident NEN in 2020 and 2021 may receive

244 outpatient treatment in Bavaria and as the record linkage may not identify all possible NEN

245 cases in the KVB data, a record linkage for 80% of all cases is probably more realistic.

246 Based on this assumption, about 3,150 cases and 6,300 matched controls would be

247 expected. With more than 3,000 expected cases (about 1,800 BPS cases and 1,340 GEPS

248 cases) and more than 6,000 expected controls, the linked dataset has a considerable size

249 and is the best available approach to exploit outpatient data for a risk factor analysis for

250 NEN. Even for GEPS tumour sites, such as the small intestine and the pancreas, we may

251 expect 360 and 325 cases, respectively.

252

Analysis

254 The descriptive analysis, stratified by BPS and GEPS, includes the calculation of frequencies

255 and percentages for categorical variables, the mean and median (with standard deviation

256 (SD) and interquartile range (IQR), respectively) for birth year as well bivariate 2 x 2 tables

257 for the combinations of case-control status and disease-related exposures. In addition, the

258 bivariate analyses are stratified by the matching variables and by the subcategories of the

259 outcome, i.e. NET and NEC. For the stratified analyses, odds ratios (ORs) will be calculated

260 according to Mantel & Haenszel.[22]

261

262 Conditional logistic regression models will be estimated separately for BPS and GEPS to

263 obtain estimates of ORs (and 95% confidence intervals (CIs)) for multiple exposures. All

264 models will adjust for birth year (5-year intervals), sex, and time of diagnosis (calendar

265 quarter). After stepwise inclusion of exposure variables, interaction terms will be added to the

266 models to investigate effect modification. All models will also be stratified by the NEN

267 subcategories NET and NEC and all analyses for GEPS will additionally be stratified by

268 tumour site.

269

270 In a sensitivity analysis, rurality of the residence district and area deprivation of the residence

271 municipality will be added to the regression models. Rurality of the residence district may be

272 associated with the likelihood of receiving an assured diagnosis of disease-related exposures

9

273 as well as outcome measures. Area deprivation of the residence municipality may be linked

274 to patterns of disease-related exposures and, thus, influence the outcome measures.

275

276 **Patient and public involvement**

277 Patients and/or the public were not involved in the design, or conduct or reporting or

278 dissemination plans of this research.

279

280 **DISCUSSION**

281 This large, population-based case-control study fully exploits the potential of linking cancer

282 registry with outpatient data to investigate risk factors for NEN, which is a rare disease as of

283 now, and to help understand their increase in Germany. Thus, the study will add to previous

284 studies from other countries, of which many were suffering from small study size, had a

285 hospital-based study design, and analysed only selected tumour sites.[2] The study will also

286 investigate exposures, such as depression and metabolic syndrome, which have not been

287 extensively studied so far.[2 10 23] An additional advantage of the study design is that

288 assessment of exposures does not rely on self-reported exposures but is based on assured

289 outpatient diagnoses in the two years preceding the NEN diagnosis, and thus are not subject

290 to recall bias nor subjective perception. With regard to the outcomes, the study will,unlike

291 previous studies, additionally stratify for NET and NEC based on high-quality cancer registry

292 data, allowing for the analysis of potential differences in risk factors between these two

293 entities. The results of this study should provide risk ratios for potential risk factors of NEN

294 and, thus, help to understand the recent NEN increase. The findings of the study may

295 provide valuable insights for government policy on potential preventive measures, while also

296 initiating further research. The study design may also serve as a flagship example of how the

297 linkage of health data of different data sources can yield substantial epidemiological insights,

298 especially in the case of rare diseases.

299

300 **ETHICS AND DISSEMINATION**

301 This study protocol was approved by the Ethics Committee of the Bavarian State Chamber of

302 Physicians (reference number: 24008). Approval by supervisory authority has been obtained

303 from the Bavarian State Ministry of Health, Care, and Prevention (reference number: G35h-

304 A1080-2023/20-2) and also the Bavarian Data Protection Commissioner stated to have no

305 concerns after presentation of the study protocol (reference number: DSB/7-692/1-275). This

306 study is based on registry and claims data, which are collected on a legal basis without the

307 explicit consent of the patients and which can be used for research purposes by the registry

308 and, under certain conditions, by third parties. Patient consent for a specific study is only

309 required for the use or linkage of plain data, but not for the study protocol presented, which is

10

310 based on an anonymized data set that does not contain any personal data. In accordance

311 with point (b) of Article 14(5) of the European Union General Data Protection Regulation, it is

312 not necessary to inform the patients in this case. The study will be conducted in accordance

313 with the Helsinki declaration of the World Medical Association as well as the guidelines and

314 recommendations for ensuring good epidemiological practice.[24]

315

316 The data that support the findings of this study will not be publicly accessible because the

317 study partners, i.e. the Bavarian Cancer Registry and the Bavarian Association of Statutory

318 Health Insurance Accredited Physicians, are subject to strict legal regulations regarding the

319 disclosure of data. Upon reasonable request, however, the permissibility of the data provision

320 will be reviewed by the Bavarian Cancer Registry and the Bavarian Association of Statutory

321 Health Insurance Accredited Physicians in accordance with the applicable legal

322 requirements.

323

324 The results of the case-control study will be presented at national and international

325 conferences. After final analysis, the results will be published in the form of scientific articles

326 in peer-reviewed journals. In addition, the authors will seek opportunities to share the

327 findings with relevant stakeholders, such as clinicians in cancer centres, and the wider public

328 by using, for instance, newsletters, press releases, and social media platforms.

329

330 **TRIAL STATUS**

331 The study started 15th October 2024 with the recruitment, i.e. data extraction and record

332 linkage process of registry and claims data. The recruitment should be completed by the end

333 of 2024.

334

335

DECLARATIONS

Authors' contributions

Conceptualization: SV, MT, JMN; Methodology: SV, NG, UB, AH, JMN; Writing – Original

Draft: SV; Writing – Review & Editing: all authors.

All authors contributed to the study design, critically reviewed the manuscript for important

intellectual content, and revised the manuscript. All authors approved the final manuscript as

submitted and agreed to be accountable for all aspects of the work.

Competing interests

The authors declare that they have no competing interests.

Patient consent for publication

Not applicable, since no individual person's data was used for this study protocol.

Data availability statement

Not applicable, since no individual person's data was used for this study protocol.

336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365

REFERENCES

1. Luley K, Gebauer J, Gebauer N, et al. [Gastroenteropancreatic neuroendocrine neoplasms-Heterogeneity, management and perspectives of treatment and research]. *Internist (Berl)* 2020;61(8):875-90. doi: 10.1007/s00108-020-00832-x

2. Leoncini E, Carioli G, La Vecchia C, et al. Risk factors for neuroendocrine neoplasms: a systematic review and meta-analysis. *Ann Oncol* 2016;27(1):68-81. doi: 10.1093/annonc/mdv505

3. 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):3063-72. doi: 10.1200/JCO.2007.15.4377

4. WHO Classification of Tumours Editorial Board. Digestive system tumours [Internet]: International Agency for Research on Cancer; 2019 [5th ed.; vol. 1:[Available from: https://tumourclassification.iarc.who.int/chapters/31 accessed 14.06.2024.

5. Montminy EM, Zhou M, Maniscalco L, et al. Contributions of Adenocarcinoma and Carcinoid Tumors to Early-Onset Colorectal Cancer Incidence Rates in the United States. *Ann Intern Med* 2021;174(2):157-66. doi: 10.7326/M20-0068 [published Online First: 2020/12/15]

6. Emrich K, Kraywinkel K. Epidemiology of rectal cancer in Germany. Is the incidence of rectal cancer increasing in young adults? *Der Onkologe* 2020;26:1085-94. doi: 10.1007/s00761-020-00857-9

7. Tsikitis VL, Wertheim BC, Guerrero MA. Trends of incidence and survival of gastrointestinal neuroendocrine tumors in the United States: a seer analysis. *J Cancer* 2012;3:292-302. doi: 10.7150/jca.4502

8. Voigtländer S, Hakimhashemi A, Grundmann N, et al. Trends of colorectal cancer incidence according to age, anatomic site, and histological subgroup in Bavaria: A registry-based study. *Front Oncol* 2022;12:904546. doi: 10.3389/fonc.2022.904546 [published Online First: 20220920]

9. Grundmann N, Voigtländer S, Hakimhashemi A, et al. Site-specific trends in gastroenteropancreatic neuroendocrine neoplasms in Bavaria, Germany. *Cancer Med* 2023;12(19):19949-58. doi: 10.1002/cam4.6510 [published Online First: 20230922]

10. Haugvik SP, Basim Ibrahim I, Hedenström P, et al. Smoking, alcohol and family history of cancer as risk factors for small intestinal neuroendocrine tumors: a systematic review and meta-analysis. *Scand J Gastroenterol* 2017;52(8):797-802. doi: 10.1080/00365521.2017.1310290 [published Online First: 20170410]

11. Kenner BJ. Early Detection of Pancreatic Cancer: The Role of Depression and Anxiety as a Precursor for Disease. *Pancreas* 2018;47(4):363-67. doi: 10.1097/MPA.0000000000001024

12. Katalinic A, Halber M, Meyer M, et al. Population-Based Clinical Cancer Registration in Germany. *Cancers (Basel)* 2023;15(15) doi: 10.3390/cancers15153934 [published Online First: 20230802]

13. Meyer M, Altmann U, Bartholomäus S, et al. Datenzusammenführung, -speicherung und Best-of-Prozess [Data linkage, data storage, and best-of process]. In: Stegmaier C, Hentschel S, Hofstädter F, et al., eds. Das Manual der Krebsregistrierung: GEKID, ADT 2019:81-98.

14. Arndt V, Holleczek B, Kajuter H, et al. Data from Population-based Cancer Registration for Secondary Data Analysis: Methodological Challenges and Perspectives. *Gesundheitswesen* 2020;82(S 01):S62-S71. doi: 10.1055/a-1009-6466

15. Federal Ministry of Health. Mitglieder und Versicherte der gesetzlichen Krankenversicherung (GKV) [Members and insured persons in the Statutory Helth Insurance] 2024 [Available from: https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/mitglieder-und-versicherte accessed 27.05.2024.

16. World Health Organization (WHO), International Agency for Research on Cancer (IARC). WHO Classification of Tumours: Thoracic tumours (5th edition) (Blue Books) 2024 [Available from: https://tumourclassification.iarc.who.int/chapters/35 accessed 27.05.2024.

416 17. World Health Organization. International Statistical Classification of Diseases and Related Health
417   Problems, 10th Revision 2020 [Available from: https://icd.who.int/browse10/2019/en
418   accessed 14.09.2023.
419 18. Schubert I, Ihle P, Köster I. Interne Validierung von Diagnosen in GKV-Routinedaten: Konzeption
420   mit Beispielen und Falldefinition. [Internal confirmation of diagnoses in routine statutory
421   health insurance data: concept with examples and case definitions]. *Gesundheitswesen*
422   2010;72(6):316-22. doi: 10.1055/s-0030-1249688 [published Online First: 2010/05/19]
423 19. Maier W. [Indices of Multiple Deprivation for the analysis of regional health disparities in
424   Germany : Experiences from epidemiology and healthcare research]. *Bundesgesundheitsblatt*
425   *Gesundheitsforschung Gesundheitsschutz* 2017;60(12):1403-12. doi: 10.1007/s00103-017-
426   2646-2
427 20. German Federal Institute for Research on Building UAaSD. Städtischer und Ländlicher Raum
428   [Urban and rural areas] 2024 [Available from:
429   https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deu
430   tschland/kreise/staedtischer-laendlicher-raum/kreistypen.html accessed 27.05.2024.
431 21. Bavarian Health and Food Safety Authority. General practice: total patients (2021) [Hausärztliche
432   Praxis: Patienten/innen insgesamt (2021)] 2024 [Available from:
433   https://www.lgl.bayern.de/gesundheit/gesundheitsberichterstattung/gesundheitsatlas/ia_re
434   port/atlas.html accessed 27.05.2024.
435 22. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of
436   disease. *J Natl Cancer Inst* 1959;22(4):719-48.
437 23. Jung YS, Yun KE, Chang Y, et al. Risk factors associated with rectal neuroendocrine tumors: a
438   cross-sectional study. *Cancer Epidemiol Biomarkers Prev* 2014;23(7):1406-13. doi:
439   10.1158/1055-9965.EPI-14-0132 [published Online First: 20140509]
440 24. Hoffmann W, Latza U, Baumeister SE, et al. Guidelines and recommendations for ensuring Good
441   Epidemiological Practice (GEP): a guideline developed by the German Society for
442   Epidemiology. *Eur J Epidemiol* 2019;34(3):301-17. doi: 10.1007/s10654-019-00500-x
443   [published Online First: 20190304]

444

445

446

14

1
2
3   1   **SUPPLEMENT**
4
5   2   **Supplement 1**
6
7   3   For cases, eligible combinations of tumour site according to the ICD-10 [1] and histology code
8
9   4   according to the WHO Blue Books [2] [3] are:
10
11   5
12
13   6   -   Gastroenteropancreatic System (GEPS)
14
15   7   o   Tumour sites: ICD-10 codes C15-C25
16
17   8   o   Histological codes based on the WHO Blue Books, categorised according to
18
19   9   ICD-10 and excluding histological codes with behaviour codes 0 (benign) and
20
21   10   1 (borderline malignancy):
22
23   11   ▪   C15, Oesophagus: 8240/3 (NET), 8246/3, 8041/3, 8013/3 (all NEC),
24
25   12   8244/3 (MiNEN)
26
27   13   ▪   C16, Stomach: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3 (all
28
29   14   NEC), 8154/3 (MiNEN), 8241/3, 8153/3 (both NET)
30
31   15   •   Additional subtypes: 8242/3, 8156/3 (both NET), 8244/3
32
33   16   (MiNEN)
34
35   17   ▪   C17, Small intestine: 8240/3, 8249/3 (both NET), 8246/3, 8041/3,
36
37   18   8013/3 (all NEC), 8154/3 (MiNEN), 8241/3, 8153/3, 8156/3 (all NET),
38
39   19   8693/3
40
41   20   ▪   C18.1, Appendix: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3
42
43   21   (all NEC), 8154/3 (MiNEN), 8241/3 (NET)
44
45   22   ▪   C18-C20 excluding C18.1, Colorectum: 8240/3, 8249/3 (both NET),
46
47   23   8246/3, 8041/3, 8013/3 (all NEC), 8154/3 (MiNEN), 8241/3 (NET)
48
49   24   ▪   C21, Anus and anal canal: 8240/3, 8249/3 (both NET), 8246/3, 8041/3,
50
51   25   8013/3 (all NEC)
52
53   26   •   Additional subtypes: 8154/3 (MiNEN)
54
55   27   ▪   C22, Liver and intrahepatic bile ducts: 8240/3, 8249/3 (both NET),
56
57   28   8246/3, 8041/3, 8013/3 (all NEC), 8154/3 (MiNEN)
58
59   29   ▪   C23-C24, Gallbladder and other and unspecified parts of the biliary
60

    30   tract: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3 (all NEC),

    31   8154/3 (MiNEN)

    32   ▪   C25, Pancreas:

    33   •   Non-functioning pancreatic NET: 8150/3

    34   •   Functioning pancreatic NET: 8151/3, 8153/3, 8155/3, 8152/3,

    35   8156/3, 8158/3, 8241/3

1

36   • Other pancreatic NET[1]: 8240/3, 8249/3

37   • Pancreatic NEC: 8246/3, 8041/3, 8013/3

38   • Pancreatic MiNEN: 8154/3

39   - Bronchopulmonary System (BPS)

40   o Tumour site: ICD-10 code C34

41   o Histological codes based on the WHO Blue Books, categorised according to

42   ICD-10:

43   ▪ C34, Bronchus and lung:

44   • NET: 8240/3, 8249/3

45   • NEC: 8041/3, 8045/3 (both small cell NEC), 8013/3 (large cell

46   NEC)

---

[1] The histological codes 8240/3 and 8249/3 are not included in the WHO Blue Books, likely because they cannot be subdivided into non-functioning and functioning tumours. Therefore, both histological codes will be included as NET.

2

48   **Supplement 2**

49   The prevalence of the disease-related exposures, apart from previous malignant neoplasms,

50   will be measured based on the following ICD-10 [1] diagnoses:

51

52   -   Depression or persistent mood disorder
53   o   ICD-10 codes F32 (depressive episode), F33 (recurrent depressive disorder),
54   or F34 (Persistent mood [affective] disorders), each including all four-
55   character categories
56   -   Obesity
57   o   ICD-10 codes E65 (localised adiposity) or E66 (obesity) including all four-
58   character categories
59   -   Metabolic syndrome (if any three out of the five following diagnoses is present)
60   o   ICD-10 code E66 (obesity), each including all four-character categories
61   o   ICD-10 code E78.1 (pure hyperglyceridaemia)
62   o   ICD-10 code E78.0 (pure hypercholesterolaemia)
63   o   ICD-10 code I10 (essential (primary) hypertension)
64   o   ICD-10 codes E10-E14 (diabetes mellitus) or R73 (elevated blood glucose
65   level), each including all four-character categories
66   -   Alcohol abuse (only for NEN of the GEPS)
67   o   ICD-10 code F10 (mental and behavioural disorders due to use of alcohol),
68   including all four-character categories
69   -   Diabetes (only for NEN of the GEPS)
70   o   ICD-10 codes E10-E14 (diabetes mellitus), each including all four-character
71   categories
72   -   Ulcerative colitis (only for NEN of the GEPS)
73   o   ICD-10 code K51 (ulcerative colitis) including all four-character categories
74   -   Crohn disease (only for NEN of the GEPS)
75   o   ICD-10 code K50 (Chron disease) including all four-character categories
76   -   Allergies (only for NEN of the BPS)
77   o   ICD-10 J30.1 (allergic rhinitis due to pollen) or J30.4 (allergic rhinitis,
78   unspecified)
79   -   Asthma (only for NEN of the BPS)
80   o   ICD-10 code J45 (asthma) including all four-character categories

81

82   The metabolic syndrome is defined as the presence of any three out of five constituting risk

83   factors,[4] i.e. elevated waist circumference, elevated triglycerides (including drug treatment for

84   elevated triglycerides), reduced high-density lipoprotein cholesterol (HDL-C, including

3

85 treatment for reduced HDL-C), elevated blood pressure (including antihypertensive drug
86 treatment), and elevated fasting glucose (including treatment for elevated fasting glucose).
87 As diagnosis data does not include information on the defined cut points, the measurement
88 of the metabolic syndrome is approximated by the presence of any three out of the five
89 following ICD-10 diagnoses: E66 approximates elevated waist circumference, E78.1
90 approximates elevated triglycerides, E78.0 approximates reduced HDL-C, I10 approximates
91 elevated blood pressure, and E10-E14 or R73 approximates elevated fasting glucose. In
92 case, the measurement of the metabolic syndrome will be limited due to missing four-
93 character categories, such as E78.1, the measurement will be based on the three-character
94 categories E66, E78, I10, and E10-E14.

95

96

4

97 **References**

98

99 1. World Health Organization. International Statistical Classification of Diseases and Related Health
100   Problems, 10th Revision 2020 [Available from: https://icd.who.int/browse10/2019/en
101   accessed 14.09.2023.
102 2. WHO Classification of Tumours Editorial Board. Digestive system tumours [Internet]: International
103   Agency for Research on Cancer; 2019 [5th ed.; vol. 1:[Available from:
104   https://tumourclassification.iarc.who.int/chapters/31 accessed 14.06.2024.
105 3. World Health Organization (WHO), International Agency for Research on Cancer (IARC). WHO
106   Classification of Tumours: Thoracic tumours (5th edition) (Blue Books) 2024 [Available from:
107   https://tumourclassification.iarc.who.int/chapters/35 accessed 27.05.2024.
108 4. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim
109   statement of the International Diabetes Federation Task Force on Epidemiology and
110   Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World
111   Heart Federation; International Atherosclerosis Society; and International Association for the
112   Study of Obesity. *Circulation* 2009;120(16):1640-5. doi:
113   10.1161/CIRCULATIONAHA.109.192644 [published Online First: 20091005]

114

5

BMJ Open

Risk factors for neuroendocrine neoplasms: a protocol for a case-control study based on a record linkage of registry and claims data

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-096273.R1
Article Type:	Protocol
Date Submitted by the Author:	17-Apr-2025
Complete List of Authors:	Voigtländer, Sven; Bavarian Health and Food Safety Authority, Bavarian Cancer Registry Gerlach, Roman; Bavarian Association of Statutory Health Insurance Accredited Physicians Grundmann, Nina; Bavarian Health and Food Safety Authority, Bavarian Cancer Registry Donnachie, Ewan; Bavarian Association of Statutory Health Insurance Accredited Physicians Berger, Ursula; Ludwig Maximilian University of Munich Faculty of Medicine, Institute for Medical Information Processing, Biometry, and Epidemiology Hakimhashemi, Amir; Bavarian Health and Food Safety Authority, Bavarian Cancer Registry Meyer, Martin; Bavarian Health and Food Safety Authority, Bavarian Cancer Registry Tauscher, Martin; Bavarian Association of Statutory Health Insurance Accredited Physicians Müller-Nordhorn, Jacqueline; Bavarian Health and Food Safety Authority, Bavarian Cancer Registry
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Oncology
Keywords:	Gastrointestinal tumours < ONCOLOGY, Respiratory tract tumours < ONCOLOGY, Risk Factors, Case-Control Studies, REGISTRIES

SCHOLARONE™
Manuscripts

1 **STUDY PROTOCOL**

2

3 **Risk factors for neuroendocrine neoplasms: a protocol for a case-control study based**

4 **on a record linkage of registry and claims data**

5

6 Sven Voigtländer[1], Roman Gerlach[2], Nina Grundmann[1], Ewan Donnachie[2], Ursula Berger[3],

7 Amir Hakimhashemi[1], Martin Meyer[1], Martin Tauscher[2], Jacqueline Müller-Nordhorn[1]

8

9 [1] Bavarian Cancer Registry, Bavarian Health and Food Safety Authority, Schweinauer

10 Hauptstraße 80, 90441 Nuremberg, Germany

11 [2] Bavarian Association of Statutory Health Insurance Accredited Physicians,

12 Elsenheimerstraße 39, 80687 Munich, Germany

13 [3] Institute for Medical Information Processing, Biometry, and Epidemiology, Faculty of

14 Medicine at Ludwig-Maximillians-University Munich, Marchioninistraße 15, 81377 Munich,

15 Germany

16

17 Word count (without title page, abstract, statements, and references):

18 3,145

19

20 **Correspondence:**

21 Dr. Sven Voigtländer

22 Bavarian Cancer Registry, Bavarian Health and Food Safety Authority

23 Schweinauer Hauptstraße 80, 90441 Nuremberg, Germany

24 Phone: +49-(0)9131-6808-2927

25 Email: sven.voigtlaender@lgl.bayern.de

26 ORCID ID: https://orcid.org/0000-0003-3808-9605

27

28

29

1

ABSTRACT

Background:

Recent studies showed an increase in neuroendocrine neoplasms, especially for the digestive tract. Several risk factors have been suggested to explain this increase including a family history of cancer, tobacco smoking, alcohol consumption, and metabolic disorders including diabetes and obesity. Another risk factor may be depressive disorders, which could increase the risk of neuroendocrine neoplasms either directly or mediated through associated risk behaviours and/or antidepressant medication. Here, we outline the design of our study to identify risk factors for neuroendocrine neoplasms in Germany.

Methods and analysis:

A case-control study of the resident population of Bavaria, the second most populous federal state in Germany, based on a record linkage of data from the Bavarian Cancer Registry and data from the Bavarian Association of Statutory Health Insurance Accredited Physicians. Cases have a diagnosis of a malignant neuroendocrine neoplasm, either of the bronchopulmonary system or the gastroenteropancreatic system, in the period from 2021 to 2023. Controls are sampled from the noncases and matched on sex, birth year (in 5-year intervals), and time of diagnosis (by calendar quarter). Risk factor prevalence of cases and controls is assessed on the basis of assured outpatient diagnoses, i.e. diagnoses documented in at least two out of four consecutive quarters in the 16 quarters preceding the diagnosis of a neuroendocrine neoplasm. The analysis uses conditional logistic regression to estimates odds ratios and 95% confidence intervals.

Ethics and dissemination:

This study protocol was approved by the Ethics Committee of the Bavarian State Chamber of Physicians (reference number: 24008). Approval by supervisory authority has been obtained from the Bavarian State Ministry of Health, Care, and Prevention (reference number: G35h-A1080-2023/20-2) and also the Bavarian Data Protection Commissioner stated to have no concerns after presentation of the study protocol (reference number: DSB/7-692/1-275). The results of the case-control study will be presented at national as well as international conferences and be published in the form of scientific articles in peer-reviewed journals.

Registration details:

ClinicalTrials.gov, NCT: 06282016.

Keywords:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

68

69

70

3

ARTICLE SUMMARY

Strengths and limitations of this study

- Cases and controls will be drawn from an unselected source population, i.e. persons, who were insured with the Statutory Health Insurance (85% of all residents) and had at least one outpatient physician contact between 2021 and 2023 within Bavaria

- Cases will comprise all malignant neuroendocrine neoplasms of the gastroenteropancreatic and bronchopulmonary system in the source population in the period from 2021 to 2023 based on high-quality cancer registry data

- The study will measure exposures based on cancer registry data and assured ICD-10 diagnoses and is, therefore, not prone to recall bias

- The study will, different to previous studies, also stratify for neuroendocrine tumours and neuroendocrine carcinomas based on high-quality cancer registry data.

- One limitation is that operationalisation of exposures is limited to exposures, which can be measured based on cancer registry as well claims data

4

1
2
3   88   **BACKGROUND**
4
5   89   Neuroendocrine neoplasms (NEN) are malignancies of neuroendocrine cells.[1]
6
7   90   Neuroendocrine cells can be found throughout the body, i.e. where there is epithelium,
8
9   91   excluding the central nervous system, bones, and soft tissue. NEN are rare and
10
11   92   heterogeneous tumours, which most commonly arise in the gastroenteropancreatic system
12
13   93   (GEPS) and the bronchopulmonary system (BPS).[1-3] NEN are categorised according to their
14
15   94   differentiation into differentiated neuroendocrine tumours (NET) and poorly differentiated
16
17   95   carcinomas (NEC).[4] NET can be further subdivided based on grade while NEC are further
18
19   96   categorised into small cell and large cell carcinomas. In addition, there are also mixed
20
21   97   neuroendocrine-non-neuroendocrine neoplasms (MiNEN).
22
23   98
24
25   99   Recent studies from the United States (US) and Germany showed an increase in NEN,
26
27   100   especially for the digestive tract.[5-9] According to an analysis of data from the US Surveillance,
28
29   101   Epidemiology and End Results (SEER) program, age-standardised incidence rates (ASR, US
30
31   102   standard population 2000) significantly increased for most gastro-intestinal sites between
32
33   103   1975 and 2008.[7] For instance, the age-adjusted incidence rate for rectal NEN increased from
34
35   104   0.1 in 1975 to 1.1 per 100,000 persons in 2008.[7] The risk for NEN is known to increase with
36
37   105   age.[2] Several other risk factors have been suggested by a limited number of studies
38
39   106   including a family history of cancer, tobacco smoking, alcohol consumption, and metabolic
40
41   107   disorders including diabetes and obesity.[2 10] Apart from this, there may be other risk factors,
42
43   108   which have not been confirmed yet, such as depressive disorders, which could increase the
44
45   109   risk of neuroendocrine neoplasms either directly or mediated through associated risk
46
47   110   behaviours and antidepressant medication. Kenner,[11] for instance, discusses the role of
48
49   111   depressive disorders for pancreatic cancer. In the past two decades, several studies have
50
51   112   explored the relationship between depression and the metabolic syndrome,[12-14] which is a
52
53   113   cluster of risk factors including raised blood pressure, dyslipidemia, raised fasting glucose,
54
55   114   and central obesity.[15] Prospective cohort studies observed a bidirectional association for
56
57   115   depression and the metabolic syndrome.[13] Both diseases, depression and the metabolic
58
59   116   syndrome, are linked with insulin resistance and chronic inflammation involving the endocrine
60
       117   and immune systems.[13 14]

       118

       119   The aim of this study is to identify risk factors for neuroendocrine neoplasms based on data

       120   of the Bavarian Cancer Registry, Germany, to help understand the increase in

       121   neuroendocrine neoplasms.

       122

       123   **METHODS**

       124   **Study design**

5

125   The study design is a case-control study based on data from the population-based Bavarian

126   Cancer Registry and data from the Bavarian Association of Statutory Health Insurance

127   Accredited Physicians (KVB, German: *Kassenärztliche Vereinigung Bayerns*). The Bavarian

128   Cancer Registry is based on mandatory notifications by physicians and health care providers

129   regarding the diagnosis and treatment of cancer.[16] The KVB collects, on a quarterly basis,

130   diagnosis and treatment data related to its main task, i.e. ensuring and reimbursing

131   outpatient medical care and psychotherapy for patients with Statutory Health Insurance

132   (GKV, German: *Gesetzliche Krankenversicherung*) in Bavaria. As common for claims data,

133   the treatment data of the KVB, however, is limited to fee schedule items including non-

134   specific flat fees, why we will focus on the diagnosis data. Bavarian Cancer Registry data

135   and KVB data are linked by pseudonymised record linkage following the probability linkage

136   procedure established by the German cancer registries.[17 18] The probability linkage is based

137   on the conversion of KVB identity data into unique tokens, i.e. pseudonymisation, that are

138   compared with the tokens already present in the Bavarian Cancer Registry by probabilistic

139   linkage.[18] After linkage, the pseudonyms are removed resulting in an anonymous dataset.

140

141   **Setting**

142   The setting is the resident population of Bavaria in the period from 2021 to 2023. Bavaria is

143   the second most populous federal state in Germany with about 13 million residents. Though

144   the Bavarian Cancer Registry covers the complete resident population of Bavaria, the source

145   population of the study is limited to persons, who are insured with the GKV and had at least

146   one outpatient physician contact between 2021 and 2023 within Bavaria. Persons who are

147   not insured in the GKV as well as persons without outpatient physician contact between 2021

148   and 2023 within Bavaria are not included in the KVB data. In Bavaria, about 11.5 million

149   residents, which is 85% of all residents, are insured with the GKV.[19]

150

151   **Definition and recruitment of cases and controls**

152   Eligible cases are defined as persons with a malignant NEN of the BPS or GEPS diagnosed

153   in the period from 2021 to 2023. Definition of malignant NEN as well as of BPS and GEPS is

154   based on the fifth edition of the World Health Organization's (WHO) Classification of

155   Tumours, also known as the WHO Blue Books,[4 20] and the International Statistical

156   Classification of Diseases and Related Health Problems, 10th Revision (ICD-10),[21]

157   respectively. Detailed information on the eligible combinations of histology code and tumour

158   site is provided in the Supplement (Supplement 1). The selection of cases is based on the

159   record linkage of Bavarian Cancer Registry and KVB data, i.e. persons with a malignant NEN

160   in the period from 2021 to 2023 according to the Bavarian Cancer Registry data that had a

6

161 GKV insurance as well as an outpatient physician contact between 2021 and 2023. Eligibility

162 is not limited to age.

163

164 Eligible controls are defined as noncases in the same source population as for cases, i.e. the

165 KVB data. Selection of controls is done by random sampling.

166

167 Cases will be matched 1:2 with controls on sex, birth year (5-year intervals), and time of

168 diagnosis (by calendar quarter). Two controls will be matched to each case.

169

170 The recruitment of cases and controls is planned to start 15th May 2025 and end 31st August

171 2025, followed by the data analysis. The recruitment comprises the data extraction at the

172 Bavarian Cancer Registry and the KVB, the record linkage of both datasets and the merging

173 of outcomes, exposures, and confounders.

174

175 **Definition of outcomes, exposures, and confounders**

176 <u>Outcomes</u>

177 Primary outcomes are NEN of the BPS and the GEPS as defined by eligible combinations of

178 histology code and tumour site (see Supplement 1). Secondary outcomes are the two

179 subcategories of NEN, i.e. NET and NEC, and the tumour characteristics stage and grade.

180

181 <u>Exposures and potential confounders</u>

182 The following exposures and potential confounders will be assessed:

183

184 Exposures: disease-related exposures

185 1. Depression or persistent mood disorder

186 2. Obesity

187 3. Metabolic syndrome

188 4. Previous malignant neoplasms

189 5. Alcohol abuse (only for NEN of the GEPS)

190 6. Diabetes (only for NEN of the GEPS)

191 7. Ulcerative colitis (only for NEN of the GEPS)

192 8. Crohn disease (only for NEN of the GEPS)

193 9. Allergies (only for NEN of the BPS)

194 10. Asthma (only for NEN of the BPS)

195

196 Exposures: sociodemographic

197 1. Area deprivation of the residence municipality

7

198

199 Potential confounders: sociodemographic and healthcare-related

200 1. Birth year (5-year intervals)

201 2. Sex

202 3. Rurality of the residence district

203 4. Healthcare utilisation frequency

204

205 Potential confounders: time of NEN diagnosis

206 1. Time of diagnosis (calendar quarter)

207

208 **Measurement of outcomes, exposures, and confounders**

209 Primary outcomes as well as their subcategories NET and NEC are measured using the ICD-

210 10 codes and histological information in the data of the Bavarian Cancer Registry (see

211 Supplement 1). Stage, grade, previous malignant neoplasms, birth year (5-year intervals),

212 sex, and time of diagnosis (by calendar quarter) are also taken from the Bavarian Cancer

213 Registry data. Stage and grade are measured according to the TNM classification of the

214 Union of International Cancer Control.[22] Previous malignant neoplasms include all neoplasms

215 with ICD-10 codes C00-C97 (malignant neoplasms) except C44 (other malignant neoplasms

216 of skin).[21] For cases and controls, the prevalence of the disease-related exposures, apart

217 from previous malignant neoplasms, is assessed by specific ICD-10 diagnoses based on the

218 KVB data (see Supplement 2 for further details). Only assured ICD-10 diagnoses will be

219 considered,[23] i.e. diagnoses that were recorded in at least two out of four consecutive

220 quarters in the 16 quarters preceding the quarter of the NEN diagnoses for cases as well as

221 controls, which are matched by quarter of NEN diagnosis. Healthcare utilisation frequency is

222 operationalised based on the use of outpatient medical care by calendar quarter and

223 physician group, as defined by the KVB. Area deprivation of the residence municipality of

224 cases and controls is measured by the Bavarian Index of Multiple Deprivation, which is

225 based on official data.[24] In particular, we use the deprivation quintile of the residence

226 municipality at the time of NEN diagnosis. Rurality of the residence district at the time of NEN

227 diagnosis is based on the dichotomous categorisation of districts into urban area and rural

228 area by the German Federal Institute for Research on Building, Urban Affairs and Spatial

229 Development.[25]

230

231 **Bias**

232 The case-control study is based on a record linkage of registry data, claims data, and

233 administrative data. These data sources are, unlike survey data, not prone to recall bias. The

234 potential of selection bias is considerably reduced by a small number of matching variables,

8

235 i.e. birth year, sex, and time of NEN diagnosis, as well as the use of an almost unselected

236 source population for cases and controls, i.e. the KVB data. The KVB data covers all persons

237 with GKV (about 85% of all residents in Bavaria) and at least one outpatient physician

238 contact in Bavaria within the study period from 2021 to 2023. According to the KVB, 94% of

239 all GKV insured persons had at least one contact with a general practitioner in 2021,[26] so that

240 about 80% of all residents are included in our source population. It is known that the

241 proportion of persons with GKV, who use outpatient care, is higher for women compared to

242 men and lower for younger persons compared to older persons.[23] Regarding potential

243 differential misclassification bias, assessment of disease-related exposures is done in the 16

244 quarters preceding the quarter of the NEN diagnosis to ensure that the prevalence and

245 number of disease-related exposures is not influenced by potentially increased clinical

246 investigation related to the NEN diagnosis. The potential of reverse causality, i.e. subclinical

247 NEN causing disease-related exposures such as depression and not vice versa, is

248 addressed by sensitivity analyses excluding the eight quarters preceding the quarter of the

249 NEN diagnosis from the assessment of disease-related exposures.

250

251 Confounding is controlled by matching for birth year, sex, and time of NEN diagnosis. We will

252 assess the matching process by comparing the distributions of the matching variables

253 between cases and controls graphically and by summary statistics of the distributions.

254 Conditional logistic regression will be employed to account for matching of controls to cases.

255 Furthermore, differences in existing infrastructure of outpatient care between urban areas

256 and rural areas may be associated with the prevalence of diagnosis-related exposures. This

257 potential confounding is additionally controlled in sensitivity analyses by adjusting models for

258 the rurality of the residence district.

259

Study size

261 Based on the data of the population-based Bavarian Cancer Registry (until 27th March 2025),

262 the number of incident NEN cases in Bavaria, Germany, is 5,943 cases for the study period

263 from 2021 to 2023, of which 3,274 were NEN of the BPS and 2,669 were NEN of the GEPS.

264 Taking into account that the KVB data covers 85% of all residents in Bavaria, Germany,

265 5,051 cases may be expected at maximum after the record linkage of Bavarian Cancer

266 Registry and KVB data. As not all residents with incident NEN between 2021 and 2023 may

267 receive outpatient treatment in Bavaria and as the record linkage may not identify all possible

268 NEN cases in the KVB data, a record linkage for 80% of all cases is probably more realistic.

269 Based on this assumption, about 4,750 cases and 9,500 matched controls would be

270 expected. With more than 4,500 expected cases (about 2,620 BPS cases and 2,130 GEPS

271 cases) and more than 9,000 expected controls, the linked dataset has a considerable size

9

272 and is the best available approach to exploit outpatient data for a risk factor analysis for

273 NEN. Even for GEPS tumour sites, such as the small intestine and the pancreas, we may

274 expect 590 and 465 cases, respectively.

275

276 **Analysis**

277 The descriptive analysis, stratified by BPS and GEPS, includes the calculation of frequencies

278 and percentages for categorical variables, the mean and median (with standard deviation

279 (SD) and interquartile range (IQR), respectively) for birth year as well bivariate 2 x 2 tables

280 for the combinations of case-control status and disease-related exposures. This allows to

281 investigate shared exposures of cases and controls. In addition, the bivariate analyses are

282 stratified by the matching variables and by the subcategories of the outcome, i.e. NET and

283 NEC. For the stratified analyses, odds ratios (ORs) will be calculated according to Mantel &

284 Haenszel.[27]

285

286 Conditional logistic regression models will be estimated separately for BPS and GEPS to

287 obtain estimates of ORs (and 95% confidence intervals (CIs)) for multiple exposures. All

288 models will adjust for birth year (5-year intervals), sex, and time of diagnosis (calendar

289 quarter). After stepwise inclusion of exposure variables, interaction terms will be added to the

290 models to investigate effect modification, for instance, between depression and the metabolic

291 syndrome. Models will be compared based on Akaike information criterion (AIC) and

292 validated examining their residual plots. The assumption of linearity in the predictors is

293 assessed using additive models with P-splines.[28] All models will also be stratified by the NEN

294 subcategories NET and NEC and all analyses for GEPS will additionally be stratified by

295 tumour site, stage, and grade.

296

297 Several sensitivity analyses will be performed. To address the potential of reverse causality,

298 the first sensitivity analysis will measure disease-related exposures based on the fourth and

299 third year preceding the quarter of the NEN diagnosis so that the eight quarters before the

300 NEN diagnosis are excluded. The second sensitivity analysis refers to the study period,

301 which partially coincides with the COVID-19 pandemic that disrupted healthcare utilisation

302 and diagnosis patterns[29 30] and may have led to an underdiagnosis of, both, outcomes as well

303 as disease-related exposures. To control for this potential confounding, we will add

304 healthcare utilisation frequency to the regression models. In a third sensitivity analysis,

305 rurality of the residence district and area deprivation of the residence municipality will be

306 added to the regression models. Rurality of the residence district may be associated with the

307 likelihood of receiving an assured diagnosis of disease-related exposures as well as outcome

308 measures. Area deprivation of the residence municipality may be linked to patterns of

309 disease-related exposures and, thus, influence the outcome measures.

310

311 Missing values may occur in the variables stage and grade of cases. If the number of missing

312 values exceeds an acceptable threshold, multiple imputation (using multiple imputation by

313 chained equation (MICE)) will be applied.[31] All variables, including the matching variables,

314 will be incorporated into the imputation model.

315

316 **Patient and public involvement**

317 Patients and/or the public were not involved in the design, or conduct or reporting or

318 dissemination plans of this research.

319

320 **DISCUSSION**

321 This large, population-based case-control study fully exploits the potential of linking cancer

322 registry with outpatient data to investigate risk factors for NEN, which is a rare disease as of

323 now, and to help understand their increase in Germany. Thus, the study will add to previous

324 studies from other countries, of which many were suffering from small study size, had a

325 hospital-based study design, analysed only selected tumour sites, and did not include

326 information on stage and grade.[2] The study will also investigate exposures, such as

327 depression and metabolic syndrome, which have not been extensively studied so far.[2 10 32] An

328 additional advantage of the study design is that assessment of exposures does not rely on

329 self-reported exposures but is based on assured outpatient diagnoses in the four years

330 preceding the NEN diagnosis, and thus are not subject to recall bias nor subjective

331 perception. With regard to the outcomes, the study will, unlike previous studies, additionally

332 stratify for NET and NEC based on high-quality cancer registry data, allowing for the analysis

333 of potential differences in risk factors between these two entities.

334

335 Limitations refer to the observation of specific causal pathways as well as the measurement

336 of disease-related exposures, as the outpatient data does not include information regarding

337 the disease onset and as ICD-10 codes lack diagnostic thresholds. To our knowledge,

338 however, this is the best currently available data in Germany to study such a large and

339 unselected population. The drawback is the lack of detailed information to observe specific

340 causal pathways. Regarding the measurement of disease-related exposures, we will use

341 assured ICD-10 diagnoses, i.e. diagnoses recorded in at least two out of four consecutive

342 quarters, to limit misclassification.

343

11

344 The results of this study should provide risk ratios for potential risk factors of NEN and, thus,

345 help to understand the recent NEN increase. The findings of the study may provide valuable

346 insights for government policy on potential preventive measures, while also initiating further

347 research. The study design may also serve as a flagship example of how the linkage of

348 health data of different data sources can yield substantial epidemiological insights, especially

349 in the case of rare diseases.

350

351 **ETHICS AND DISSEMINATION**

352 This study protocol was approved by the Ethics Committee of the Bavarian State Chamber of

353 Physicians (reference number: 24008). Approval by supervisory authority has been obtained

354 from the Bavarian State Ministry of Health, Care, and Prevention (reference number: G35h-

355 A1080-2023/20-2) and also the Bavarian Data Protection Commissioner stated to have no

356 concerns after presentation of the study protocol (reference number: DSB/7-692/1-275). This

357 study is based on registry and claims data, which are collected on a legal basis without the

358 explicit consent of the patients and which can be used for research purposes by the registry

359 and, under certain conditions, by third parties. Patient consent for a specific study is only

360 required for the use or linkage of plain data, but not for the study protocol presented, which is

361 based on an anonymized data set that does not contain any personal data. In accordance

362 with point (b) of Article 14(5) of the European Union General Data Protection Regulation, it is

363 not necessary to inform the patients in this case. The study will be conducted in accordance

364 with the Helsinki declaration of the World Medical Association as well as the guidelines and

365 recommendations for ensuring good epidemiological practice.[33]

366

367 The data that support the findings of this study will not be publicly accessible because the

368 study partners, i.e. the Bavarian Cancer Registry and the Bavarian Association of Statutory

369 Health Insurance Accredited Physicians, are subject to strict legal regulations regarding the

370 disclosure of data. Upon reasonable request, however, the permissibility of the data provision

371 will be reviewed by the Bavarian Cancer Registry and the Bavarian Association of Statutory

372 Health Insurance Accredited Physicians in accordance with the applicable legal

373 requirements.

374

375 The results of the case-control study will be presented at national and international

376 conferences. After final analysis, the results will be published in the form of scientific articles

377 in peer-reviewed journals. In addition, the authors will seek opportunities to share the

378 findings with relevant stakeholders, such as clinicians in cancer centres, and the wider public

379 by using, for instance, newsletters, press releases, and social media platforms.

380

12

381 **TRIAL STATUS**

382 The study starts 15th May 2025 with the recruitment, i.e. data extraction and record linkage

383 process of registry and claims data. The recruitment should be completed by 31st August

384 2025.

385

386

387 **DECLARATIONS**

388 **Authors' contributions**

389 Conceptualization: SV, MT, JMN; Methodology: SV, NG, UB, AH, JMN; Writing – Original

390 Draft: SV; Writing – Review & Editing: all authors.

391

392 All authors contributed to the study design, critically reviewed the manuscript for important

393 intellectual content, and revised the manuscript. All authors approved the final manuscript as

394 submitted and agreed to be accountable for all aspects of the work.SV is the guarantor.

395

396 **Funding Statement**

397 This research received no specific grant from any funding agency in the public, commercial

398 or not-for-profit sectors.

399

400 **Competing interests**

401 The authors declare that they have no competing interests.

402

403 **Patient consent for publication**

404 Not applicable, since no individual person's data was used for this study protocol.

405

406 **Data availability statement**

407 Not applicable, since no individual person's data was used for this study protocol.

408

409 **Acknowledgements**

410 We are grateful to the Trust Centre of the Bavarian Cancer Registry (Dr Jana Johne, Mr

411 Stefan Möllenkamp) for substantial support regarding ethical approval and regulatory

412 approval. Moreover, we want to thank the Cancer Registry of North Rhine-Westphalia (Prof

413 Stang) for commenting on an earlier version of the study protocol.

414

415

416

14

REFERENCES

1. Luley K, Gebauer J, Gebauer N, et al. [Gastroenteropancreatic neuroendocrine neoplasms-Heterogeneity, management and perspectives of treatment and research]. *Internist (Berl)* 2020;61(8):875-90. doi: 10.1007/s00108-020-00832-x

2. Leoncini E, Carioli G, La Vecchia C, et al. Risk factors for neuroendocrine neoplasms: a systematic review and meta-analysis. *Ann Oncol* 2016;27(1):68-81. doi: 10.1093/annonc/mdv505

3. 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):3063-72. doi: 10.1200/JCO.2007.15.4377

4. WHO Classification of Tumours Editorial Board. Digestive system tumours [Internet]: International Agency for Research on Cancer; 2019 [5th ed.; vol. 1:[Available from: https://tumourclassification.iarc.who.int/chapters/31 accessed 14.06.2024.

5. Montminy EM, Zhou M, Maniscalco L, et al. Contributions of Adenocarcinoma and Carcinoid Tumors to Early-Onset Colorectal Cancer Incidence Rates in the United States. *Ann Intern Med* 2021;174(2):157-66. doi: 10.7326/M20-0068 [published Online First: 2020/12/15]

6. Emrich K, Kraywinkel K. Epidemiology of rectal cancer in Germany. Is the incidence of rectal cancer increasing in young adults? *Der Onkologe* 2020;26:1085-94. doi: 10.1007/s00761-020-00857-9

7. Tsikitis VL, Wertheim BC, Guerrero MA. Trends of incidence and survival of gastrointestinal neuroendocrine tumors in the United States: a seer analysis. *J Cancer* 2012;3:292-302. doi: 10.7150/jca.4502

8. Voigtländer S, Hakimhashemi A, Grundmann N, et al. Trends of colorectal cancer incidence according to age, anatomic site, and histological subgroup in Bavaria: A registry-based study. *Front Oncol* 2022;12:904546. doi: 10.3389/fonc.2022.904546 [published Online First: 20220920]

9. Grundmann N, Voigtländer S, Hakimhashemi A, et al. Site-specific trends in gastroenteropancreatic neuroendocrine neoplasms in Bavaria, Germany. *Cancer Med* 2023;12(19):19949-58. doi: 10.1002/cam4.6510 [published Online First: 20230922]

10. Haugvik SP, Basim Ibrahim I, Hedenström P, et al. Smoking, alcohol and family history of cancer as risk factors for small intestinal neuroendocrine tumors: a systematic review and meta-analysis. *Scand J Gastroenterol* 2017;52(8):797-802. doi: 10.1080/00365521.2017.1310290 [published Online First: 20170410]

11. Kenner BJ. Early Detection of Pancreatic Cancer: The Role of Depression and Anxiety as a Precursor for Disease. *Pancreas* 2018;47(4):363-67. doi: 10.1097/MPA.0000000000001024

12. Löwe B, Hochlehnert A, Nikendei C. [Metabolic syndrome and depression]. *Ther Umsch* 2006;63(8):521-7. doi: 10.1024/0040-5930.63.8.521

13. Pan A, Keum N, Okereke OI, et al. Bidirectional association between depression and metabolic syndrome: a systematic review and meta-analysis of epidemiological studies. *Diabetes Care* 2012;35(5):1171-80. doi: 10.2337/dc11-2055

14. Qiu W, Cai X, Zheng C, et al. Update on the Relationship Between Depression and Neuroendocrine Metabolism. *Front Neurosci* 2021;15:728810. doi: 10.3389/fnins.2021.728810 [published Online First: 20210831]

15. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009;120(16):1640-5. doi: 10.1161/CIRCULATIONAHA.109.192644 [published Online First: 20091005]

16. Katalinic A, Halber M, Meyer M, et al. Population-Based Clinical Cancer Registration in Germany. *Cancers (Basel)* 2023;15(15) doi: 10.3390/cancers15153934 [published Online First: 20230802]

15

468 17. Meyer M, Altmann U, Bartholomäus S, et al. Datenzusammenführung, -speicherung und Best-of-
469   Prozess [Data linkage, data storage, and best-of process]. In: Stegmaier C, Hentschel S,
470   Hofstädter F, et al., eds. Das Manual der Krebsregistrierung: GEKID, ADT 2019:81-98.
471 18. Arndt V, Holleczek B, Kajuter H, et al. Data from Population-based Cancer Registration for
472   Secondary Data Analysis: Methodological Challenges and Perspectives. *Gesundheitswesen*
473   2020;82(S 01):S62-S71. doi: 10.1055/a-1009-6466
474 19. Federal Ministry of Health. Mitglieder und Versicherte der gesetzlichen Krankenversicherung
475   (GKV) [Members and insured persons in the Statutory Helth Insurance] 2024 [Available from:
476   https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-
477   fakten-zur-krankenversicherung/mitglieder-und-versicherte accessed 27.05.2024.
478 20. World Health Organization (WHO), International Agency for Research on Cancer (IARC). WHO
479   Classification of Tumours: Thoracic tumours (5th edition) (Blue Books) 2024 [Available from:
480   https://tumourclassification.iarc.who.int/chapters/35 accessed 27.05.2024.
481 21. World Health Organization. International Statistical Classification of Diseases and Related Health
482   Problems, 10th Revision 2020 [Available from: https://icd.who.int/browse10/2019/en
483   accessed 14.09.2023.
484 22. Brierley JD, Gospodarowicz MK, Wittekind C, editors. *TNM Classification of Malignant Tumours,
485   8th edition*: Union for International Cancer Control, 2017.
486 23. Schubert I, Ihle P, Köster I. Interne Validierung von Diagnosen in GKV-Routinedaten: Konzeption
487   mit Beispielen und Falldefinition. [Internal confirmation of diagnoses in routine statutory
488   health insurance data: concept with examples and case definitions]. *Gesundheitswesen*
489   2010;72(6):316-22. doi: 10.1055/s-0030-1249688 [published Online First: 2010/05/19]
490 24. Maier W. [Indices of Multiple Deprivation for the analysis of regional health disparities in
491   Germany : Experiences from epidemiology and healthcare research]. *Bundesgesundheitsblatt
492   Gesundheitsforschung Gesundheitsschutz* 2017;60(12):1403-12. doi: 10.1007/s00103-017-
493   2646-2
494 25. German Federal Institute for Research on Building UAaSD. Städtischer und Ländlicher Raum
495   [Urban and rural areas] 2024 [Available from:
496   https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deu
497   tschland/kreise/staedtischer-laendlicher-raum/kreistypen.html accessed 27.05.2024.
498 26. Bavarian Health and Food Safety Authority. General practice: total patients (2021) [Hausärztliche
499   Praxis: Patienten/innen insgesamt (2021)] 2024 [Available from:
500   https://www.lgl.bayern.de/gesundheit/gesundheitsberichterstattung/gesundheitsatlas/ia_re
501   port/atlas.html accessed 27.05.2024.
502 27. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of
503   disease. *J Natl Cancer Inst* 1959;22(4):719-48.
504 28. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. *Statistical Science*
505   1996;11(2):89-121, 33.
506 29. Alom S, Chiu CM, Jha A, et al. The Effects of COVID-19 on Cancer Care Provision: A Systematic
507   Review. *Cancer Control* 2021;28:1073274821997425. doi: 10.1177/1073274821997425
508 30. Voigtländer S, Hakimhashemi A, Grundmann N, et al. Impact of the COVID-19 pandemic on
509   reported cancer diagnoses in Bavaria, Germany. *J Cancer Res Clin Oncol* 2023 doi:
510   10.1007/s00432-023-04707-0 [published Online First: 20230324]
511 31. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and
512   guidance for practice. *Stat Med* 2011;30(4):377-99. doi: 10.1002/sim.4067 [published Online
513   First: 20101130]
514 32. Jung YS, Yun KE, Chang Y, et al. Risk factors associated with rectal neuroendocrine tumors: a
515   cross-sectional study. *Cancer Epidemiol Biomarkers Prev* 2014;23(7):1406-13. doi:
516   10.1158/1055-9965.EPI-14-0132 [published Online First: 20140509]
517 33. Hoffmann W, Latza U, Baumeister SE, et al. Guidelines and recommendations for ensuring Good
518   Epidemiological Practice (GEP): a guideline developed by the German Society for
519   Epidemiology. *Eur J Epidemiol* 2019;34(3):301-17. doi: 10.1007/s10654-019-00500-x
520   [published Online First: 20190304]

16

1
2
3   521
4
5   522
6   523
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

17

1 **SUPPLEMENT**

2 **Supplement 1**

3 For cases, eligible combinations of tumour site according to the ICD-10 [1] and histology code

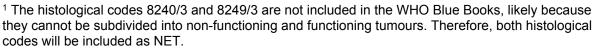4 according to the WHO Blue Books [2][3] are:

5

6 - Gastroenteropancreatic System (GEPS)

    o Tumour sites: ICD-10 codes C15-C25

    o Histological codes based on the WHO Blue Books, categorised according to

9 ICD-10 and excluding histological codes with behaviour codes 0 (benign) and

10 1 (borderline malignancy):

11 ▪ C15, Oesophagus: 8240/3 (NET), 8246/3, 8041/3, 8013/3 (all NEC),

12 8244/3 (MiNEN)

13 ▪ C16, Stomach: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3 (all

14 NEC), 8154/3 (MiNEN), 8241/3, 8153/3 (both NET)

15 • Additional subtypes: 8242/3, 8156/3 (both NET), 8244/3

16 (MiNEN)

17 ▪ C17, Small intestine: 8240/3, 8249/3 (both NET), 8246/3, 8041/3,

18 8013/3 (all NEC), 8154/3 (MiNEN), 8241/3, 8153/3, 8156/3 (all NET),

19 8693/3

20 ▪ C18.1, Appendix: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3

21 (all NEC), 8154/3 (MiNEN), 8241/3 (NET)

22 ▪ C18-C20 excluding C18.1, Colorectum: 8240/3, 8249/3 (both NET),

23 8246/3, 8041/3, 8013/3 (all NEC), 8154/3 (MiNEN), 8241/3 (NET)

24 ▪ C21, Anus and anal canal: 8240/3, 8249/3 (both NET), 8246/3, 8041/3,

25 8013/3 (all NEC)

26 • Additional subtypes: 8154/3 (MiNEN)

27 ▪ C22, Liver and intrahepatic bile ducts: 8240/3, 8249/3 (both NET),

28 8246/3, 8041/3, 8013/3 (all NEC), 8154/3 (MiNEN)

29 ▪ C23-C24, Gallbladder and other and unspecified parts of the biliary

30 tract: 8240/3, 8249/3 (both NET), 8246/3, 8041/3, 8013/3 (all NEC),

31 8154/3 (MiNEN)

32 ▪ C25, Pancreas:

33 • Non-functioning pancreatic NET: 8150/3

34 • Functioning pancreatic NET: 8151/3, 8153/3, 8155/3, 8152/3,

35 8156/3, 8158/3, 8241/3

1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

36   - Other pancreatic NET[1]: 8240/3, 8249/3

37   - Pancreatic NEC: 8246/3, 8041/3, 8013/3

38   - Pancreatic MiNEN: 8154/3

39   - Bronchopulmonary System (BPS)

40   o Tumour site: ICD-10 code C34

41   o Histological codes based on the WHO Blue Books, categorised according to

42   ICD-10:

43   ▪ C34, Bronchus and lung:

44   - NET: 8240/3, 8249/3

45   - NEC: 8041/3, 8045/3 (both small cell NEC), 8013/3 (large cell

46   NEC)

---

[1] The histological codes 8240/3 and 8249/3 are not included in the WHO Blue Books, likely because they cannot be subdivided into non-functioning and functioning tumours. Therefore, both histological codes will be included as NET.

2

48 **Supplement 2**

49 The prevalence of the disease-related exposures, apart from previous malignant neoplasms,

50 will be measured based on the following ICD-10 [1] diagnoses:

51

52 - Depression or persistent mood disorder

53   o ICD-10 codes F32 (depressive episode), F33 (recurrent depressive disorder),

54   or F34 (Persistent mood [affective] disorders), each including all four-

55   character categories

56 - Obesity

57   o ICD-10 codes E65 (localised adiposity) or E66 (obesity) including all four-

58   character categories

59 - Metabolic syndrome (if any three out of the five following diagnoses is present)

60   o ICD-10 code E66 (obesity), each including all four-character categories

61   o ICD-10 code E78.1 (pure hyperglyceridaemia)

62   o ICD-10 code E78.0 (pure hypercholesterolaemia)

63   o ICD-10 code I10 (essential (primary) hypertension)

64   o ICD-10 codes E10-E14 (diabetes mellitus) or R73 (elevated blood glucose

65   level), each including all four-character categories

66 - Alcohol abuse (only for NEN of the GEPS)

67   o ICD-10 code F10 (mental and behavioural disorders due to use of alcohol),

68   including all four-character categories

69 - Diabetes (only for NEN of the GEPS)

70   o ICD-10 codes E10-E14 (diabetes mellitus), each including all four-character

71   categories

72 - Ulcerative colitis (only for NEN of the GEPS)

73   o ICD-10 code K51 (ulcerative colitis) including all four-character categories

74 - Crohn disease (only for NEN of the GEPS)

75   o ICD-10 code K50 (Chron disease) including all four-character categories

76 - Allergies (only for NEN of the BPS)

77   o ICD-10 J30.1 (allergic rhinitis due to pollen) or J30.4 (allergic rhinitis,

78   unspecified)

79 - Asthma (only for NEN of the BPS)

80   o ICD-10 code J45 (asthma) including all four-character categories

81

82 The metabolic syndrome is defined as the presence of any three out of five constituting risk

83 factors,[4] i.e. elevated waist circumference, elevated triglycerides (including drug treatment for

84 elevated triglycerides), reduced high-density lipoprotein cholesterol (HDL-C, including

3

85   treatment for reduced HDL-C), elevated blood pressure (including antihypertensive drug

86   treatment), and elevated fasting glucose (including treatment for elevated fasting glucose).

87   As diagnosis data does not include information on the defined cut points, the measurement

88   of the metabolic syndrome is approximated by the presence of any three out of the five

89   following ICD-10 diagnoses: E66 approximates elevated waist circumference, E78.1

90   approximates elevated triglycerides, E78.0 approximates reduced HDL-C, I10 approximates

91   elevated blood pressure, and E10-E14 or R73 approximates elevated fasting glucose. In

92   case, the measurement of the metabolic syndrome will be limited due to missing four-

93   character categories, such as E78.1, the measurement will be based on the three-character

94   categories E66, E78, I10, and E10-E14.

95

96

4

97 **References**

98

99 1. World Health Organization. International Statistical Classification of Diseases and Related Health
100   Problems, 10th Revision 2020 [Available from: https://icd.who.int/browse10/2019/en
101   accessed 14.09.2023.
102 2. WHO Classification of Tumours Editorial Board. Digestive system tumours [Internet]: International
103   Agency for Research on Cancer; 2019 [5th ed.; vol. 1:[Available from:
104   https://tumourclassification.iarc.who.int/chapters/31 accessed 14.06.2024.
105 3. World Health Organization (WHO), International Agency for Research on Cancer (IARC). WHO
106   Classification of Tumours: Thoracic tumours (5th edition) (Blue Books) 2024 [Available from:
107   https://tumourclassification.iarc.who.int/chapters/35 accessed 27.05.2024.
108 4. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim
109   statement of the International Diabetes Federation Task Force on Epidemiology and
110   Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World
111   Heart Federation; International Atherosclerosis Society; and International Association for the
112   Study of Obesity. *Circulation* 2009;120(16):1640-5. doi:
113   10.1161/CIRCULATIONAHA.109.192644 [published Online First: 20091005]

114

5