Original research

Gastrointestinal syndromes preceding a diagnosis of Parkinson’s disease: testing Braak’s hypothesis using a nationwide database for comparison with Alzheimer’s disease and cerebrovascular diseases

Bo Konings,1 Luisa Villatoro,2 Jef Van den Eynde,3 Guillermo Barahona,2 Robert Burns,4 Megan McKnight,4 Ken Hui,4 Gayane Yenokyan,5 Jan Tack,1 Pankaj Jay Pasricha2

ABSTRACT

Objective Braak’s hypothesis states that Parkinson’s disease (PD) originates in the gastrointestinal (GI) tract, and similar associations have been established for Alzheimer’s disease (AD) and cerebrovascular diseases (CVD). We aimed to determine the incidence of GI syndromes and interventions preceding PD compared with negative controls (NCs), AD and CVD.

Design We performed a combined case-control and cohort study using TriNetX, a US based nationwide medical record network. Firstly, we compared subjects with new onset idiopathic PD with matched NCs and patients with contemporary diagnoses of AD and CVD, to investigate preceding GI syndromes, appendectomy and vagotomy. Secondly, we compared cohorts with these exposures to matched NCs for the development of PD, AD and CVD within 5 years.

Results We identified 24 624 PD patients in the case-control analysis and matched 18 cohorts with each exposure to their NCs. Gastroparesis, dysphagia, irritable bowel syndrome (IBS) without diarrhoea and constipation showed specific associations with PD (vs NCs, AD and CVD) in both the case-control (odds ratios (ORs) vs NCs 4.64, 3.58, 3.53 and 3.32, respectively, all p<0.0001) and cohort analyses (relative risks (RRs) vs NCs 2.43, 2.27, 1.17 and 2.38, respectively, all p<0.05). While functional dyspepsia, IBS with diarrhoea, diarrhoea and faecal incontinence were not PD specific, IBS with constipation and intestinal pseudo-obstruction showed PD specificity in the case-control (OR 4.11) and cohort analysis (RR 1.84), respectively. Appendectomy decreased the risk of PD in the cohort analysis (RR 0.48). Neither inflammatory bowel disease nor vagotomy were associated with PD.

Conclusion Dysphagia, gastroparesis, IBS without diarrhoea and constipation might specifically predict Parkinson’s disease.

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Braak’s hypothesis states that Parkinson’s disease (PD) originates in the gut in a subset of patients, but no studies to date have systematically investigated a broad range of gastrointestinal (GI) symptoms and diagnoses before a diagnosis of PD.

WHAT THIS STUDY ADDS
⇒ This is the first multicentre study to establish that dysphagia, gastroparesis, constipation and irritable bowel syndrome without diarrhoea specifically increase the risk of a subsequent new onset diagnosis of idiopathic Parkinson’s disease, even compared with other neurological diseases, such as Alzheimer’s disease and cerebrovascular diseases.

INTRODUCTION

Parkinsonism is a clinical syndrome characterised by bradykinesia, rest tremor, rigidity and postural instability.1 Its most common cause is Parkinson’s disease (PD), the pathological hallmark of which is thought to be cytoplasmatic eosinophilic Lewy body (LB) depositions. These depositions, mainly consisting of misfolded α-synuclein, have not only been found in the CNS but also in the vagus nerve and enteric nervous system (ENS) of patients with PD.4

These findings led Braak et al to state the neuroanatomical hypothesis that α-synuclein pathology progresses from peripheral sites such as the ENS to the CNS via vagal or olfactory pathways, thereby introducing the concept that the gastrointestinal (GI) tract might serve as a gateway for environmental factors that induce α-synuclein misfolding and lead to PD.5 A large body of evidence has since then accumulated to support this claim. Even in early untreated stages of the disease, neuropathological studies have found that α-synuclein concentrations in the ENS of patients with PD were higher than those of otherwise healthy individuals, in a
characteristic rostrocaudal gradient following visceromotor projections of the vagus nerve. Complementary studies have shown that various motility disorders and inflammatory bowel disease (IBD) can precede PD and therefore may be risk factors for its development. Moreover, since Gray and colleagues first identified the vermiform appendix as a potential source of misfolded α-synuclein, conflicting observational studies have been published about the impact of an appendicectomy on the risk of idiopathic PD. Finally, two recent registry based studies have strengthened the concept of retrograde vagal α-synuclein propagation by showing that a truncal vagotomy might be protective against the development of PD. Apart from the bottom up link formulated by Braak, a top down aetiology in which GI symptoms are present in early phases when neurological manifestations are still unnoticed is also supported by experimental evidence. Even if no causal link exists, GI syndromes might still represent a risk factor through other mechanisms, or both might be related to a yet unknown third factor.

Apart from PD, other neurological disorders have also been hypothesised to have GI precedents, either through similar neuroimmune pathways or translocation of microbiome derived neurotoxins into the CNS. A strong pathological link between microbiome derived neurotoxins, including Escherichia coli derived Lipopolysaccharide, has been established with disrupted intestinal cell adhesion, impaired synaptic signalling in the Alzheimer’s disease (AD) brain and exacerbation of inflammatory neuropsychology. Additionally, given the prominent role of reactive oxygen species induced inflammation in cerebrovascular diseases (CVD), proinflammatory intestinal and extraintestinal diseases have been linked to a higher risk of CVD than that predicted by conventional risk factors.

Previous studies on this topic have been limited by small sample sizes and inadequate controls. Therefore, we used a nationwide electronic health record (EHR) network to investigate the incidence of various GI syndromes and interventions, such as appendectomy and vagotomy, before the onset of PD. Because previous studies lacked specificity for exposures associated with PD, we used a case-control study design to compare patients with PD not only with negative controls (NCs), but also with patients diagnosed with AD and CVD. Additionally, we established a cohort study design for each exposure in the case-control design to validate these findings and establish relative risk (RR) estimates relevant in clinical practice.

**METHODS**

**Study design and data source**

To investigate the association between various GI syndromes and interventions with the subsequent development of new onset PD, we analysed electronic medical records from the TriNetX Analytics Research Network (Cambridge, Massachusetts, USA). At the moment of data collection, the network consisted of more than 80 million patients from 57 predominantly academic medical centres in the USA. Additional information can be found in the online supplemental methods.

**Study population and variables of interest**

In the case control analysis, we examined the incidence of exposures retrospectively (ie, before an initial diagnosis of PD compared with matched controls). Patients with PD were captured using a previously validated method. Patients were queried using the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10) diagnosis of PD (G20), if documented between 1 January 2005 and 1 July 2021; the first ever diagnosis was used as the index event. Only those with at least two prescriptions of an antiparkinsonian drug and a documented ambulatory visit at least 2 years before the first diagnosis of PD were included; secondary causes of PD were excluded. To determine PD specific exposures, control subjects comprised three groups: NCs, and patients with a diagnosis of AD and CVD. NCs consisted of patients without a recorded ICD-10 diagnosis of PD, with at least two documented ambulatory visits between the ages of 50 and 90 years, at least 2 years apart, recorded between 1 January 2005 and 1 July 2021. A minimum of 2 years of retrospective follow-up was ensured by using the second of these visits as the index event. Similarly, 2 years of follow-up was ensured for the AD and CVD groups, and the first ever documented respective ICD-10 diagnosis in the medical records after 1 January 2005 was chosen as the index event. In a pairwise fashion, these groups where then matched to the PD group for age, sex, race and ethnicity using a propensity score matching algorithm.

To cover the entirety of the GI tract, 18 exposures were investigated: achalasia, dysphagia, gastro-oesophageal reflux disease (GORD), gastroparesis (GP), functional dyspepsia (FD), paralytic ileus (PI), diarrhoea, irritable bowel syndrome (IBS) with constipation (IBS-C), IBS with diarrhoea (IBS-D), IBS without diarrhoea, intestinal pseudo-obstruction (approximate synonym of K59.8: other specified functional GI disorders), faecal incontinence (FI), Crohn’s disease (CD), ulcerative colitis (UC), microscopic colitis (MC), appendectomy and vagotomy. We conducted additional sensitivity analyses in the network, which included stratified analyses based on sex and age at the diagnosis of the index event. A detailed breakdown of the query, inclusion and exclusion criteria, stratified analyses and coding can be found in the online supplemental methods.

To validate the results from the case-control analyses, we set up a complementary cohort study design. Eighteen cohorts, each diagnosed with one of the investigated exposures in the case-control analysis, were queried and compared with a respective NC cohort (ie, without the exposure) for the prospective risk of developing PD, AD or CVD. Only those with at least 5 years of prospective follow-up were included, and were propensity score matched for age, sex, race and ethnicity, and additionally for a set of potential risk factors and risk modifiers for the development of PD, AD and CVD: arterial hypertension, diabetes mellitus, atrial fibrillation and flutter, and nicotine dependence.

**Statistical analysis**

In the case-control analyses, patients were counted as positive for an exposure if the respective ICD-10 code was documented any time before the first diagnosis of PD or the control health event. To approximate the diagnostic interval between each exposure and the first PD diagnosis, a yearly cross sectional prevalence for each exposure was calculated for the PD and NC groups, up to 6 years before the index event. To detect and quantify potential surveillance bias in our case-control analyses, we collected an agnostic set of negative exposures (Charlson comorbidities). This allowed us to determine the OR that should be considered as indicative of no association. Additionally, we collected positive exposures based on a previous case-control study that identified prodromal motor and non-motor symptoms of PD. This enabled us to assess the ability of our dataset to reproduce existing associations. The coding can be found in the online supplemental methods.

In the cohort analyses, patients diagnosed with the exposure of interest and their NCs were counted as positive for an outcome...
(PD, AD or CVD) if the respective new onset ICD-10 diagnosis occurred within a 5 year follow-up. Subjects who already had the outcome of interest before the index event were excluded after propensity score matching.

Exposures and outcomes were collected as absolute numbers; ORs and RRs were calculated with 95% CIs. Standardised mean differences (SMDs) were used to compare baseline characteristics; an SMD of <0.2 was considered well balanced. A Pearson χ² test was calculated to compare outcomes, and a two sided p value of <0.05 was used to indicate statistical significance. Correction for false discovery rate (FDR) was performed using the step up procedure by Benjamini and Yekutieli, with the Stats P value of <0.05 was used to indicate statistical significance.

RESULTS
For the case-control study, 24,624 patients with PD met all of the criteria and were matched with 8,267,744 NCs, and 36,187 AD and 528,207 CVD patients, giving 24,624 patients in the comparison with NCs, 19,046 with AD and 23,942 with CVD. Baseline characteristics after pairwise matching are presented in table 1; minimal differences in age at index persisted. SMDs and p values before and after matching can be found in online supplemental tables 1,2.

The results of the case-control analyses are presented in figure 1 and online supplemental tables 3. All GI syndromes were significantly increased in the PD group compared with NCs (OR >1; p<0.05). However, only dysphagia (OR 3.58), GP (OR 4.64), FD (OR 3.39), intestinal pseudo-obstruction (OR 3.01), diarrhoea (OR 2.85), constipation (OR 3.32), IBS-C (OR 4.11), IBS-D (OR 4.31), IBS without diarrhoea (OR 3.53) and FI (OR 3.76) gave ORs that were numerically greater than the upper limit of what is expected by surveillance bias (OR range 1.20–2.79). Only for GORD and appendectomy we observed significant differences compared with AD (OR 1.14, p<0.0001) and CVD (OR 0.57, p=0.03), respectively. The latter did not remain significant after correction for FDR (p=0.21). Prior vagotomy did not impact the risk of PD.

For GP, women were approximately twice as likely as men to develop PD (OR 7.3 for women, 3.05 for men, both p<0.0001 online supplemental tables 7–9 and online supplemental figure 4), and the OR for GP was especially high for early onset PD (online supplemental tables 10–11). Exclusion of previous antiparkinsonian drug use did not significantly alter any associations with PD compared with NCs (online supplemental table 12 and online supplemental figure 5). The ORs of all PD specific exposures were positioned well within the range of those of established motor and non-motor prodromes of PD (ie, positive outcomes; OR PD vs NCs 2.04–7.41). An approximation of the diagnostic interval for each exposure can be found in online supplemental figure 6.

Figure 2 and online supplemental tables 13–14 show the results of the cohort analyses. A significantly increased RR of new onset PD (RR >1; p<0.05) was found after a diagnosis of dysphagia (RR 2.27), GORD (RR 1.13), GP (RR 2.43), FD (RR 1.15), intestinal pseudo-obstruction (RR 1.84), diarrhoea (RR 1.32), constipation (RR 2.38), IBS without diarrhoea (RR 1.17) and FI (RR 1.74); all except for FD (p=0.09) remained significant after correction for FDR. However, only for dysphagia, GP, intestinal pseudo-obstruction, IBS without diarrhoea and constipation was this RR numerically higher than the RR of developing AD and

Table 1 Baseline characteristics for subjects with PD and controls in the case-control analyses after pairwise matching.

<table>
<thead>
<tr>
<th>Characteristic *</th>
<th>Patients (n=24 624)</th>
<th>Negative controls (n=24 624)</th>
<th>Parkinson’s disease (n=19 046)</th>
<th>Alzheimer’s disease (n=19 046)</th>
<th>Parkinson’s disease (n=23 942)</th>
<th>Cerebrovascular diseases (n=23 942)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) (mean (SD))</td>
<td>70.8 (8.49)</td>
<td>70.3 (8.68)</td>
<td>72.7 (7.93)</td>
<td>73.4 (7.99)</td>
<td>70.8 (8.51)</td>
<td>70.6 (8.68)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>14254 (57.89)</td>
<td>14254 (57.89)</td>
<td>9948 (52.23)</td>
<td>9979 (52.39)</td>
<td>13874 (57.95)</td>
<td>13874 (57.95)</td>
</tr>
<tr>
<td>Women</td>
<td>10103 (41.03)</td>
<td>10103 (41.03)</td>
<td>8869 (46.57)</td>
<td>8835 (46.39)</td>
<td>9801 (40.94)</td>
<td>9801 (40.94)</td>
</tr>
<tr>
<td>Unknown</td>
<td>267 (1.08)</td>
<td>267 (1.08)</td>
<td>229 (1.22)</td>
<td>232 (1.22)</td>
<td>267 (1.12)</td>
<td>267 (1.12)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>20476 (83.15)</td>
<td>20476 (83.15)</td>
<td>15458 (81.16)</td>
<td>15461 (81.18)</td>
<td>19906 (83.14)</td>
<td>19906 (83.14)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>1294 (5.26)</td>
<td>1294 (5.26)</td>
<td>1166 (6.12)</td>
<td>1176 (6.17)</td>
<td>1182 (4.94)</td>
<td>1182 (4.94)</td>
</tr>
<tr>
<td>Other and unknown†</td>
<td>2854 (11.59)</td>
<td>2854 (11.59)</td>
<td>2425 (12.73)</td>
<td>2413 (12.67)</td>
<td>2856 (11.93)</td>
<td>2854 (11.92)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>939 (3.81)</td>
<td>939 (3.81)</td>
<td>810 (4.25)</td>
<td>806 (4.23)</td>
<td>942 (3.93)</td>
<td>942 (3.93)</td>
</tr>
<tr>
<td>Not Hispanic of Latino</td>
<td>19746 (80.19)</td>
<td>19746 (80.19)</td>
<td>15419 (80.96)</td>
<td>15411 (80.91)</td>
<td>19771 (82.58)</td>
<td>19771 (82.58)</td>
</tr>
<tr>
<td>Unknown</td>
<td>3939 (16)</td>
<td>3939(16)</td>
<td>2817 (14.79)</td>
<td>2829 (14.85)</td>
<td>3229 (13.49)</td>
<td>3229 (13.49)</td>
</tr>
</tbody>
</table>

*Characteristics were identified using electronic medical health record data from the TriNetX Research Network. Baseline characteristics before matching can be found in online supplemental table 1, and p values and standardised mean differences between groups in online supplemental table 2.
†Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, or unknown.
### Figure 1

Case-control analyses. Odds ratios (ORs) of previous exposures in patients with Parkinson’s disease (PD) compared with matched negative controls (NCs), or patients with Alzheimer’s disease (AD) and cerebrovascular diseases (CVD). In the case-control analysis, the incidence of exposures was examined retrospectively (ie, before an initial diagnosis of PD compared with matched NCs, and AD and CVD patients). Patients were diagnosed with PD or the respective control health event between 1 January 2005 and 1 July 2021, and NCs were queried using two ambulatory visits during the same time window. Patients and exposures were identified with the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10) codes, and Current Procedural Terminology (CPT) codes, using electronic medical health record data from the TriNetX research network (online supplemental methods). Exposures were included if they were documented any time before the diagnosis of PD or the control health event, in the available medical records.

ORs were calculated as follows: odds of documented exposure in the PD cohort/odds of documented exposure in the control cohort. Absolute rates can be found in online supplemental table 3. *P* values were calculated using a Pearson $\chi^2$ test, after matching for baseline characteristics (ie, age, sex, race and ethnicity). $d$Correction for false discovery rate was performed using the step up procedure by Benjamini and Yekutieli, with the Stats package in R.

**The term intestinal pseudo-obstruction was used as an approximate synonym for the ICD-10 code K59.8 other specified functional intestinal disorders.

**The term IBS not otherwise specified is commonly used as an approximate synonym for the ICD-10 code K58.9 IBS without diarrhoea. IBs, irritable bowel syndrome.
Cohort analyses. Relative risks (RRs) of developing Parkinson’s disease (PD), Alzheimer’s disease (AD) or cerebrovascular diseases (CVD) within 5 years of the diagnosis of a given exposure, compared with negative controls (NCs) without the respective exposure. For each analysis, a cohort of patients identified by the diagnosis of a given exposure was compared with their respective NCs for the prospective risk of PD, AD and CVD within 5 years of the index event (ie, diagnosis of the given exposure, or a visit for NCs). After propensity score matching, patients that already had the investigated outcome (ie, PD, AD or CVD) documented before the index event were excluded from the analysis. Exposures and outcomes were identified using the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10,) and Current Procedural Terminology (CPT) codes. Electronic medical health record data were collected from the TriNetX research network. Diagnostic coding can be found in the online supplemental methods.

RRs were calculated as follows: risk of outcome in the exposure cohort/risk of outcome in the control cohort. Absolute rates can be found in online supplemental table 13. P values were calculated with a Pearson χ² test, after matching for baseline characteristics and risk factors. Baseline characteristics included age, sex, race and ethnicity; risk factors included arterial hypertension, diabetes mellitus, atrial fibrillation and flutter, and nicotine dependence. Correction for false discovery rate was performed with the step up procedure by Benjamini and Yekutieli, using the Stats package in R. The term intestinal pseudo-obstruction was used as an approximate synonym for the ICD-10 code K59.8 other specified functional intestinal disorders. The term IBS not otherwise specified is commonly used as an approximate synonym for the ICD-10 code K59.9 IBS without diarrhoea. IBS, Irritable bowel syndrome.
CVD, and statistical significance (p<0.05) was achieved only for constipation and dysphagia (online supplemental table 1). In line with the case control analyses, FI equally increased the risk of PD and AD (PD: RR 1.74; AD: RR 1.76, both p<0.0001), and in the GORD cohort the RR of developing CVD was higher than that of PD (CVD: RR 1.38; PD: RR 1.13, both p<0.0001). FD and diarrhoea were associated with all three disorders (p<0.05), and PI, CD and MC only increased the risk of CVD (RR 1.22, p<0.0001; RR 1.20, p<0.0001; RR 1.26, p=0.02, respectively). Although no specific association was found in the case-control analysis, appendectomy significantly reduced the risk of developing PD (RR 0.48, p=0.05), which did not remain significant after correction for FDR. For all other exposures (ie, achalasia, UC, IBS-C, IBS-D and vagotomy), no significant associations were found.

**DISCUSSION**

We used a nationwide EHR network to comprehensively investigate disorders across the entire GI tract before a diagnosis of PD. We used two complementary study designs to establish that dysphagia, gastroparesis, constipation and IBS without diarrhoea specifically increase the risk of a subsequent new onset diagnosis of idiopathic PD, even compared with other neurological diseases, such as AD and CVD.

Surveillance bias is an inherent problem in observational studies. When not addressed appropriately, it can compromise the validity of causal inference and lead to irreproducible results.26 A broader implementation of empirical approaches to evaluate and correct for the presence of systematic error in observational studies is needed.27 Therefore, we set up an approach to understand the true extent of surveillance bias in our study and its potential contribution to implicating premorbid factors for PD. Hence we collected data on all diagnoses included in the Charlson comorbidity index. These premorbid conditions were considered as a comprehensive set of agnostic negative exposures.

In our case-control study, we observed statistically significant increases in most of these exposures in PD cases compared with NCs, but not compared with AD and CVD (online supplemental figure 1). To establish whether these increases represented surveillance bias or true associations (although unlikely based on the current literature), we investigated the same exposures for other neurological disorders (AD and CVD) compared with their NCs (online supplemental figure 3). Since the same significant correlations emerged, surveillance bias was likely. Subsequently, we determined the range of ORs that should be expected if they are a result of surveillance bias alone. These ORs ranged between 1.20 and 2.79 in the analysis of PD with NCs. To determine whether existing associations with PD could be replicated with ORs greater than those of negative exposures, we also collected prodromal motor and non-motor symptoms of PD (ie, positive exposures). These resulted in ORs ranging between 2.04 and 7.41 (online supplemental figure 1).

Having established a measure of surveillance bias in the case-control study, we then determined ORs for the putative GI pre-comorbidities of PD. Relative to the upper limit of the negative exposures, GI exposures fell into two categories. First were those for which the ORs overlapped with the ORs expected for surveillance bias (OR 1.20–2.79). For these exposures, including achalasia, GORD, PI, IBD (ie, CD, UC, and MC), appendectomy and vagotomy, we cannot be confident that these were true associations, although within the constraints of our study we cannot categorically state that they were not. Second were those for which the ORs were clearly higher than the ORs expected for surveillance bias (OR >2.79). For these exposures, including dysphagia, GP, FI, intestinal pseudo-obstruction, diarrhoea, constipation, IBS-C, IBS-D, IBS without diarrhoea and FI, we can confidently state to have established significant associations with new onset PD. To determine the specificity of the identified significant exposures for PD in the case-control analyses, we subsequently compared subjects with PD with subjects with AD and CVD. Only dysphagia, GP, constipation, IBS without diarrhoea and IBS-C remained specific for PD compared with both neurological diseases (table 2). Importantly, we cannot exclude the possibility that these factors might still be associated with these diseases, although at a smaller scale. Similarly, while other exposures were not specific for PD (ie, FI, intestinal pseudo-obstruction, diarrhoea, IBS-D and FI), we cannot strictly exclude the possibility that these conditions might still be risk factors for PD.

Finally, to validate these findings both in terms of their significance and specificity and establish one RR estimate of developing PD after the diagnosis of each exposure, we set up a complementary cohort study. Here, five exposures (ie, dysphagia, GP, IBS without diarrhoea, intestinal pseudo-obstruction and constipation) significantly increased the risk of PD and resulted in RRs that were numerically greater than those of AD and CVD.
(table 2). Four of these provide internal validation for PD specific exposures identified in the case-control analysis. These exposures are thus very unlikely to be a result of selection or surveillance bias and can therefore be considered the most significant findings from our study. Minor discrepancies can be explained by intrinsic differences in the study designs.

The consistent correlation between constipation and PD (RR 2.38, 95% CI 2.24 to 2.54) confirms an abundance of existing literature. Previous reports have stated that constipation can even precede PD by up to 20 years. More surprising is the strong association for dysphagia (RR 2.27, 95% CI 2.10 to 2.45), which has so far mainly been reported after its diagnosis.13 The prevalence of oesophageal dysmotility in PD has been shown to be as high as 80% when using objective measures,29 but a delay in oesophageal transit has also been found in drug-naive and subjectively asymptomatic phases of the disease.30 31 While evidence supports that oesophageal function might be affected through brainstem and cortical areas,32 post mortem studies by Mu et al showed that pharyngeal muscles, sensory neurons and motor neurons are also often affected by LB pathology in PD.32 33 Interestingly, the highest RR was observed for GP (RR 1.05 to 1.3). Finally, we attempted to assess the proximity of each diagnosis to the diagnosis of PD in the case-control study (online supplemental figure 5). We found that the OR for dysphagia and constipation decreased considerably as the distance from the diagnosis of PD increased, while the OR for GP and IBS without diarrhoea remained relatively constant. This suggests that differences in lead time exist, but future longitudinal population based studies will be crucial to determine whether these PD specific GI syndromes are part of the early manifestation of PD or truly precede the disease. Importantly, the combination of two complementary study designs reduced the potential for selection bias. The case-control analysis consisted of patients with PD, AD or CVD without the requirement of any previous exposure, while the cohort analyses consisted of patients with newly diagnosed GI exposures without the requirement of a subsequent PD, AD or CVD diagnosis. This study is subject to intrinsic limitations of EHR data, including unknown completeness of records and absent validation of diagnoses. The multicentre character and inclusion of racially and ethnically diverse subjects ensured that these results are generalisable to patients at academic medical centres across the USA.
CONCLUSION
This study is the first to establish substantial observational evidence that the clinical diagnosis of not only constipation, but also dysphagia, GP and IBS without diarrhoea might specifically predict the development of PD, whereas other exposures were less specific. An appendixdectomy appeared protective, leading to further speculation about its role in PD pathophysiology. These findings warrant alertness for GI syndromes in patients at higher risk for PD and highlight the need for further investigation of GI precedents in AD and CVD. To establish a stronger body of clinicopathological evidence, we advocate for future studies to assess the sensitivity and specificity of these disorders and their clinicopathological correlates for the early detection of neuropathology.

Acknowledgements
BK and JV were supported by a fellowship of the Belgian American Educational Foundation (BAEF). Parts of this work have been presented as a conference abstract at Digestive Disease Week 2022, San Diego, California, USA (doi: 10.1016/S0016-5085(22)60469-4).

Contributors
LK: conceptualisation, methodology, and validation. GY: methodology and methodology, and visualisation. GB: conceptualisation, resources, and investigation. DJV: conceptualisation, resources, and supervision. MM: conceptualisation. KH: conceptualisation, methodology, and validation. GM: methodology and formal analysis. JT: writing-review and editing. PJP: guarantor, supervision, conceptualisation, methodology, validation, and writing-review and editing.

Funding
The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests
None declared.

Patient and public involvement
Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication
Not applicable.

Provenance and peer review
Not commissioned; externally peer reviewed.

Data availability statement
Data are available upon reasonable request. Data exported from TriNetX were saved in Excel files and archived. Every co-author affiliated to Johns Hopkins University was granted access to the TriNetX Research network by the Institute of Clinical and Translational Research (ICTR). The online supplemental materials contain an extensive list of tables representing the original data; researchers will be granted access to the original aggregated data upon reasonable request, with agreement of the corresponding author (PJP).

Supplemental material
This content has been supplied by the author(s). It has not been vetted by BMI Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iD
Bo Konings http://orcid.org/0000-0002-3889-6853

REFERENCES

18 Pogue AI, Jaber VR, Sharman NM, et al. Downregulation of neurofilament light chain expression in human neuronal-glia cell co-cultures by a microbiome-derived (lipo polysaccharide) induced mRNA-320-SP. Front Neurol 2022;13:90048.