

# Prediagnostic expressions in health records predict mortality in Parkinson's disease: A proof-of-concept study

Tomi Kuusimäki<sup>a,b,\*</sup>, Jani Sainio<sup>c</sup>, Samu Kurki<sup>c</sup>, Tero Vahlberg<sup>d</sup>, Valtteri Kaasinen<sup>a,b</sup>

<sup>a</sup> Clinical Neurosciences, University of Turku, Turku, Finland

<sup>b</sup> Neurocenter, Turku University Hospital, Turku, Finland

<sup>c</sup> Auria Biobank, University of Turku and Turku University Hospital, Turku, Finland

<sup>d</sup> Department of Biostatistics, University of Turku, Turku, Finland

## ARTICLE INFO

### Keywords:

Parkinson's disease  
Prodromal  
Data mining  
Electronical health records  
Mortality  
Prognosis

## ABSTRACT

**Introduction:** The relationship of prodromal markers of PD with PD mortality is unclear. Electronic health records (EHRs) provide a large source of raw data that could be useful in the identification of novel relevant prognostic factors in PD. We aimed to provide a proof of concept for automated data mining and pattern recognition of EHRs of PD patients and to study associations between prodromal markers and PD mortality.

**Methods:** Data from EHRs of PD patients (n = 2522) were collected from the Turku University Hospital database between 2006 and 2016. The data contained >27 million words/numbers and >750000 unique expressions. The 5000 most common words were identified in three-year time period before PD diagnosis. Cox regression was used to investigate the association of expressions with the 5-year survival of PD patients.

**Results:** During the five-year period after PD diagnosis, 839 patients died (33.3%). If expressions associated with psychosis/hallucinations were identified within 3 years before the diagnosis, worse survival was observed (hazard ratio = 1.71, 95%CI = 1.46–1.99, p < 0.001). Similar effects were observed for words associated with cognition (1.23, 1.05–1.43, p = 0.009), constipation (1.34, 1.15–1.56, p = 0.0002) and pain (1.34, 1.12–1.60, p = 0.001).

**Conclusions:** Automated mining of EHRs can predict relevant clinical outcomes in PD. The approach can identify factors that have previously been associated with survival and detect novel associations, as observed in the link between poor survival and prediagnostic pain. The significance of early pain in PD prognosis should be the focus of future studies with alternate methods.

## 1. Introduction

Although Parkinson's disease (PD) is associated with an early and progressive loss of dopaminergic neurons in the substantia nigra, PD can be characterized as a multisystem progressive disease that affects several neurotransmitter systems [1,2]. Importantly, PD-specific pathology seems to initiate years before classical motor symptoms are clinically observable [3], and a marked loss of dopaminergic neurons is already present when the diagnosis is established [4]. In addition to cardinal motor symptoms, patients with PD also suffer from numerous nonmotor symptoms such as psychiatric problems, cognitive impairment, dementia, dysautonomia, constipation and sleep problems [2]. Some of the nonmotor symptoms may occur several years before motor symptoms in the prodromal phase of PD [2].

One of the key challenges in PD pharmacotherapy is the lack of neuroprotective agents that could alter the disease course. Another related problem is the identification of prodromal patients who could be targeted for neuroprotective treatments. Previous studies have identified several prodromal markers in PD [5]. The first published prodromal markers were derived from studies investigating asymptomatic relatives of PD patients [6] and epidemiological studies designed for other purposes [7]. Later, prospective studies specifically investigating prodromal PD [8,9], and studies with retrospective reviews of electronic health records (EHRs) became available [10,11]. These studies have specifically identified REM sleep behavior disorder, depression, constipation, olfactory dysfunction, erectile dysfunction, somnolence, orthostatic hypotension and urinary dysfunction as prodromal markers of PD. Although prodromal symptoms increase the risk of PD diagnosis, the

\* Corresponding author. Neurocenter, Turku University Hospital, Hämeentie 11, POB 52, FIN-20521, Turku, Finland.

E-mail address: [tomi.kuusimaki@tyks.fi](mailto:tomi.kuusimaki@tyks.fi) (T. Kuusimäki).

<https://doi.org/10.1016/j.parkreldis.2021.12.015>

Received 11 July 2021; Received in revised form 8 December 2021; Accepted 29 December 2021

Available online 1 January 2022

1353-8020/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

relationship between prodromal markers and mortality is unclear.

EHR systems are currently in routine use in many countries, including Finland [12], and the increasing amount of data can be used for several secondary purposes, such as statistics and decision support [12]. There is potentially valuable information inside EHRs that could be utilized using data mining of language processing [12]. Therefore, in this retrospective record-based cohort study, we sought to investigate whether data mining of prediagnostic language can be used for PD research. We investigated a large hospital-based cohort of Finnish PD patients and examined possible associations between these prodromal markers and the mortality of PD patients. The first aim was to provide proof-of-concept that text mining can identify relevant previously reported clinical prodromal symptoms in PD. The second aim was to investigate whether the method can identify novel associations that could be tested in other research settings.

## 2. Material and methods

The Turku Clinical Research Center (Turku CRC, <http://www.turku.crc.fi/en>) maintains the Turku University Hospital database, which contains detailed clinical data of patients who have visited Turku University Hospital or regional hospitals. In Finland, terms medical records/electronic health records are used interchangeably, and health records have been in electronic format from the January 1, 2004 onwards, enabling annual longitudinal analyses, and they include demographics, clinical diagnoses and procedures, inpatient periods and outpatient visits, pathology diagnoses and reports, imaging reports, chemotherapy and radiation treatments, inpatient medications and outpatient prescriptions, laboratory measurements, and clinical narratives. The data are pseudonymized, protecting the identity of the patients while making it possible to link data elements to individual patients.

The study population consisted of PD patients who were treated between 2006 and 2014 at Turku University Hospital district in southwestern Finland, and who had available EHR data from the prediagnostic period. PD patients were identified from the Turku University Hospital database on the basis of ICD-10 code G20\*. The date of diagnosis was defined as the first appearance of the diagnosis code in patient records. The diagnostic criteria of PD are homogeneous in different parts of Finland and are based on the Finnish Current Care Guidelines of PD [13]. To receive drug reimbursement for PD, the diagnosis must be made by a certified neurologist using either the UK Brain Bank criteria or the MDS clinical criteria [14] on the basis of clinical examination. Thus, all diagnoses of PD in our data were based on these criteria. In addition, dates and individual causes of death were obtained from the national authority Statistics Finland ([www.stat.fi](http://www.stat.fi)). The beginning of the follow-up was January 1, 2006, and the end of the follow-up was December 31, 2016.

After identifying patients and diagnosis dates, individual patient data were collected from the database. All electronic texts and lists were used including referrals, medical imaging reports and laboratory results. To extract the most common relevant words, Finnish stop words were removed from the tokens. We focused on the prediagnostic period of PD, and, consequently, we identified the 5000 most common words in time period of three years before PD diagnosis. The search allowed specific words to occur only once in the same time period. All 5000 words were reviewed by two investigators, and suitable words were categorized into ten categories that were determined in advance. Two word categories (hyposmia and sleep disorders) were excluded due to an insufficient number of words in categories. The final eight categories were words related to parkinsonism, depression, psychosis/hallucinations, cognition, constipation, pain, cancer or circulation.

## 3. Statistics

We used age- and sex-adjusted Cox regression analysis to investigate the association of word appearance within 3 years before PD diagnosis

with the 5-year survival of PD patients. The results are expressed by hazard ratios (HRs) with 95% confidence intervals (CIs). P-values less than 0.05 were considered statistically significant. Statistical analyses were performed using SAS System for Windows, version 9.4 (SAS Institute Inc., Cary, NC).

The authors confirm that the approval of an institutional review board was not required for this work. The study was approved by the Turku Clinical Research Center (TT263/2017, diary number TK-53-652-18).

## 4. Results

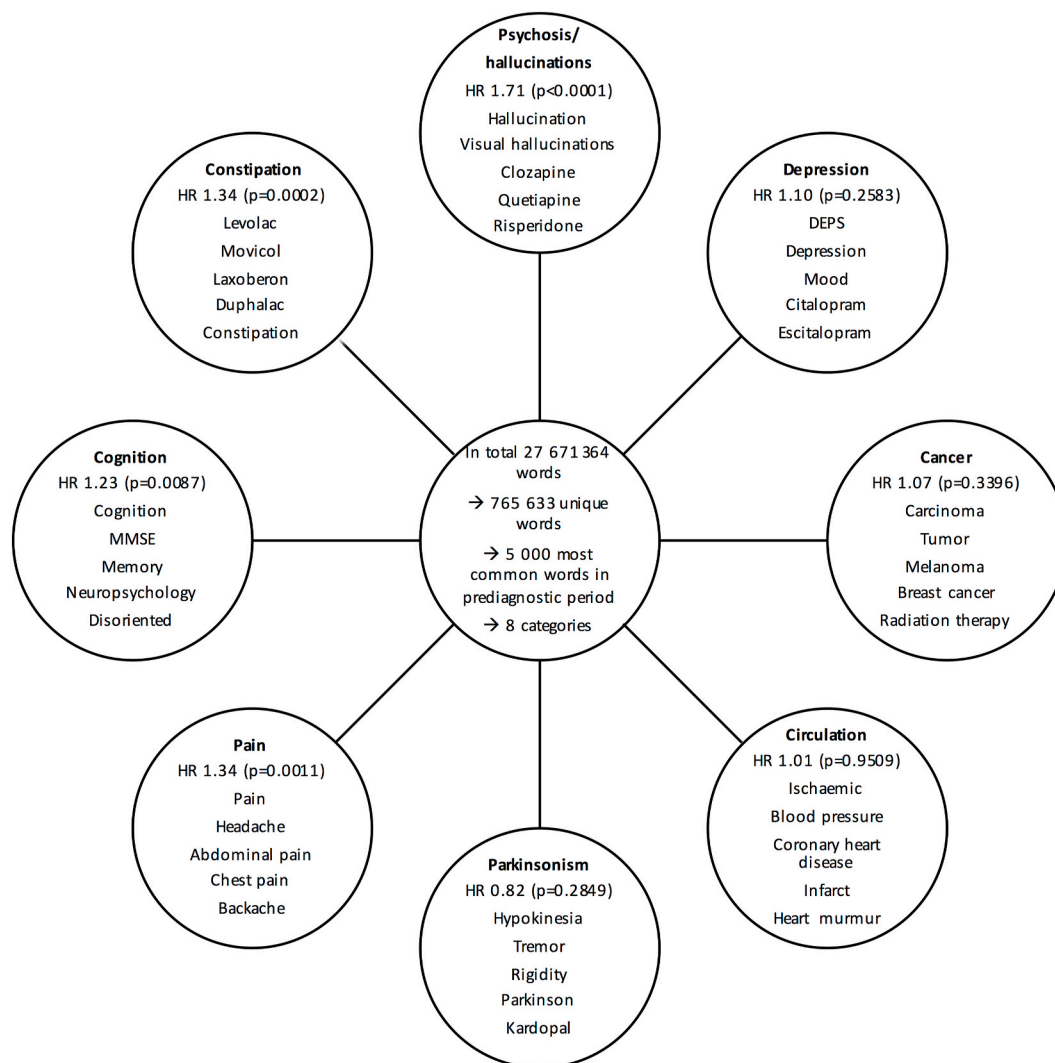
In the three-year time period before PD diagnosis, there were 2522 PD patients with available EHR data. In total, 27 671 364 words were identified from EHRs of PD patients, 765 633 of which were unique. Examples of common words in different word categories are presented in Fig. 1. During a five-year period after diagnosis, 33.3% of patients (839) died, of whom 56.7% (476) were men [Table 1]. Men were diagnosed at a younger age than women (76.9 years, SD 7.5 years vs. 79.3 years, SD 7.4 years). The mean age of PD patients at the time of death was 80.1 years (SD 7.5 years) (men: 79.0 years, SD 7.5 years, women: 81.6 years, SD 7.2 years). Pneumonia was the most common immediate cause of death in both men and women (39.3% of deaths in men and 27.3% of deaths in women). The second and third most common causes of death were myocardial infarction and heart failure, respectively.

The HR for prediagnostic psychosis/hallucinations words was 1.71 (95% CI 1.46–1.99,  $p < 0.0001$ ), indicating worse survival for PD patients if psychosis/hallucinations-related words were identified within 3 years before PD diagnosis [Table 2, Fig. 2]. Additionally, the appearances of pain- (HR 1.34, 95% CI 1.12–1.60,  $p = 0.0011$ ), constipation- (HR 1.34, 95% CI 1.15–1.56,  $p = 0.0002$ ), and cognition- (HR 1.23, 95% CI 1.05–1.43,  $p = 0.0087$ ) related words over the prediagnostic period were associated with increased mortality. Parkinsonism, depression, circulation, and cancer-related words were not significantly associated with PD patient survival [Table 2].

## 5. Discussion

We demonstrate a methodological approach that enables analyses of massive data sets of large PD study populations. The results show that prediagnostic screening of medical terms in EHRs of PD patients can provide indicators of mortality. In particular, prediagnostic psychosis/hallucinations, cognition, constipation and pain-related word categories were associated with worse survival for PD patients. On the other hand, no associations with parkinsonism-, depression-, circulation-, or cancer-related words were observed.

Nonmotor symptoms are typically the main determinant of PD patients' quality of life [15], and certain nonmotor symptoms have been identified as risk factors for mortality [16]. Consistent with the present results, psychotic symptoms [17–19] and cognitive impairment/dementia [17,18,20,21] have previously been shown to be associated with an increased risk of death in patients with PD. Importantly, however, these associations have been demonstrated in diagnosed patients often years after the onset of first symptoms, whereas the present results focus on the prediagnostic period. Previous studies have also suggested that an early development of autonomic dysfunction or individual autonomic abnormalities (including constipation, orthostatic hypotension, urinary symptoms, upper gastrointestinal tract symptoms, sweating abnormalities, and erectile dysfunction in males) are independent determinants of more rapid disease progression and shorter survival in patients already diagnosed with PD [22]. Constipation was also associated with a higher risk of death in diagnosed PD patients in Kaplan-Meier analysis [16]. However, to our knowledge, there are no previous reports on the direct association between pain and survival in PD. PD-related pain has been reported to increase the risk of falls, which could secondarily affect the



**Fig. 1.** Examples of common words/expressions in different categories, and hazard ratios for 5-year survival. DEPS = Depression Scale, Duphalac = trade name for lactulose, HR = Hazard ratio, Kardopal = trade name for carbidopa/levodopa, Laxoberon = trade name for sodium picosulfate, Levolaac = trade name for lactulose, MMSE = Mini-Mental State Examination, Movicol = trade name for macrogol,  $p$  = P-value.

**Table 1**

Demographics and three most common causes of death (I, II, III) of the 839 patients who died during the five-year period after PD diagnosis.

Gender	Patients (n/%)	Mean age at diagnosis (SD)	Mean age at death (SD)	I (n/%)	II (n/ %)	III (n/ %)
Men	476/56.7	76.9 (7.5)	79.0 (7.5)	Pneumonia 187/39.3	MI 16/ 3.4	HF 12/ 2.5
Women	363/43.3	79.3 (7.4)	81.6 (7.2)	Pneumonia 99/27.3	MI 14/ 3.9	HF 12/ 3.3
All	839	77.9 (7.6)	80.1 (7.5)	Pneumonia 286/34.1	MI 30/ 3.6	HF 24/ 2.9

HF = heart failure, MI = myocardial infarction, PD = Parkinson's disease, SD = standard deviation.

The immediate cause of death was not reported for 41.2% of males, 47.7% of females, 44.2% of all patients.

occurrence of acute complications and deaths in PD patients [23].

Over recent years, the role and clinical importance of pain in PD has become apparent. Studies have shown that PD patients suffer from pain

more than age-matched controls [24], and as many as 61.8% of PD patients have been reported to experience at least one form of chronic pain according to the results of the DoPaMiP study [25]. The clinical importance of the issue has been demonstrated by studies that have shown that pain is a significant nonmotor symptom of PD, but it remains an underdiagnosed and undertreated symptom negatively influencing the quality of life of PD patients [24]. In the present study, the results point to a possible link between pain in the prediagnostic period of PD and mortality within the first five years post diagnosis. It is possible that poor survival is related to early balance problems or comorbidities. In addition, pain may also be an indirect motor sign of rigidity. One of the most common term in pain category was backache that may associate with axial rigidity, which may explain reduced survival. However, future studies with other methods are needed to investigate this in detail.

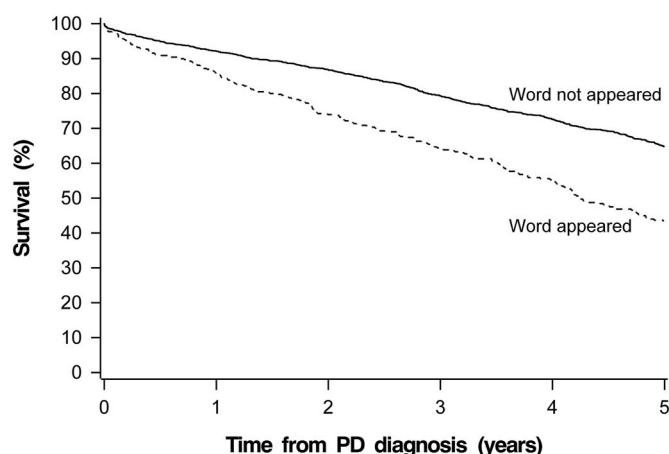
It is important to note that although our study focused on the pre-diagnostic time period, it was not the same as the premotor stage. The PD diagnosis date in this study was the first appearance of the diagnosis code in patient records, and the diagnostic latency from motor symptom occurrence to clinical diagnosis of PD typically varies from months to a few years [26,27]. Therefore, the present results should be interpreted to indicate that psychosis/hallucinations, cognition, constipation and pain-related expressions are associated with worse survival in the early

**Table 2**

Associations between prediagnostic word categories and 5-year survival of PD patients.

Word category	Deaths/ Patients (%)	Age and sex adjusted HR (95% CI)	P-value
<b>Circulation</b>			
Suitable word not appeared	67/249 (26.9)		
Suitable word appeared	772/2273 (34.0)	1.01 (0.78–1.30)	0.9509
<b>Cancer</b>			
Suitable word not appeared	537/1676 (32.0)		
Suitable word appeared	302/846 (35.7)	1.07 (0.93–1.24)	0.3396
<b>Parkinsonism</b>			
Suitable word not appeared	31/71 (43.7)		
Suitable word appeared	808/2451 (33.0)	0.82 (0.57–1.18)	0.2849
<b>Depression</b>			
Suitable word not appeared	647/1915 (33.8)		
Suitable word appeared	192/607 (31.6)	1.10 (0.93–1.29)	0.2583
<b>Cognition</b>			
Suitable word not appeared	244/899 (27.4)		
Suitable word appeared	595/1633 (36.4)	1.23 (1.05–1.43)	0.0087
<b>Constipation</b>			
Suitable word not appeared	613/1956 (31.3)		
Suitable word appeared	226/566 (39.9)	1.34 (1.15–1.56)	0.0002
<b>Pain</b>			
Suitable word not appeared	154/601 (25.6)		
Suitable word appeared	685/1921 (35.7)	1.34 (1.12–1.60)	0.0011
<b>Psychosis/hallucinations</b>			
Suitable word not appeared	624/2083 (30.0)		
Suitable word appeared	215/439 (49.0)	1.71 (1.46–1.99)	<0.0001

CI = Confidence interval, HR= Hazard ratio.

**Fig. 2.** Kaplan-Meier 5-year survival curves if psychosis/hallucinations-related words appeared or not appeared within 3 years before PD diagnosis.

motor phase of PD, albeit belonging to the prediagnostic period. This probably explains also the relatively high number of deceased patients, as the true time period between disease onset and death was likely considerably longer than 5 years. Moreover, another issue affecting

mortality is that the study population included only PD patients treated at the Turku University Hospital (main hospital and TUH regional hospitals). PD patients with mild symptoms are usually receiving treatment outside the university hospital (general practitioners and private practitioners), whereas PD patients with severe symptoms and advanced disease who have reduced survival are more often treated at university hospitals.

Parkinsonism-, depression-, circulation-, and cancer-related expressions within 3 years before PD diagnosis were not associated with increased mortality. As expected, the appearance of parkinsonism-related words was high in PD patients in the prediagnostic time period (97.2%) and the very high prevalence of parkinsonism-related expressions made longitudinal analyses less meaningful. Moreover, it is important to note that the search allowed specific words to occur only once in the same time period. Therefore, the appearance of an expression indicates merely that there was a certain symptom or health-related problem without information on its severity. This, together with the duration of follow-up (max 8 years), may explain why depression-, circulation- or cancer-related expression was not associated with PD mortality. Previous results about the association of depression with mortality in PD are somewhat mixed. Some studies have reported that depression is a predictor of decreased survival in PD [28], whereas some studies have not found a significant association between depression and PD mortality [29].

A limitation in any register-based PD study is the suboptimal diagnostic accuracy. As psychosis/hallucinations-related words included expressions related to antipsychotic drugs (clozapine, quetiapine, risperidone), there is a possibility of drug-induced parkinsonism in some patients. In addition, some patients with an early presentation of hallucinations may eventually receive a diagnosis of Lewy body dementia. However, all PD diagnoses in the study cohort were based on individual clinical examinations by neurologists using formal diagnostic criteria which reduces the possibility of misdiagnoses.

As EHRs are becoming more prevalent in medical care worldwide, a similar automated analysis of PD medical texts is possible in many regions. EHRs may be used to assess study feasibility, facilitate patient recruitment, and streamline data collection at baseline and follow-up [30]. Although this method provides advantages by enabling investigations of massive sample sizes, there are limitations that should be taken into account when the results are interpreted, such as less reliable documentation, possibility of missing data, and difficulty in verifying documented information. Moreover, the variability in the quality of documentation among health care professionals affects data quality. Given these limitations, our method is best at generating research hypotheses to be tested using other methods.

To conclude, we have demonstrated a methodological approach for PD research that can be used to mine large clinical samples of patients. We suggest that the method can be used to bring previously undetected clinically relevant factors to the surface. The validity of the method can be observed in text mining of EHRs in early stages of PD, which shows that psychosis/hallucinations, cognition and constipation-related words are related to worse survival, as previously reported using other methods of investigation.

## Funding

This work was supported by Turku University Hospital (VTR-funds), Instrumentarium Science Foundation, The Finnish Parkinson Foundation, and the Päivikki and Sakari Sohlberg Foundation.

## Author contributions

- 1) Research project: A. Conception, B. Organization, C. Execution;
- 2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique;
- 3) Manuscript: A. Writing of the first draft, B. Review and Critique.



T.K.: 1B, 1C, 2A, 2B, 2C, 3A, 3B  
 J.S.: 1C, 2A, 2C, 3B  
 S.K.: 1C, 2C, 3B  
 T.V.: 1C, 2A, 2B, 2C, 3B  
 V.K.: 1A, 1B, 1C, 2C, 3B

## Declaration of competing interest

None.

## References

- [1] D.W. Dickson, Neuropathology of Parkinson disease, *Park. Relat. Disord.* 46 (Suppl 1) (2018) S30–S33, <https://doi.org/10.1016/j.parkreldis.2017.07.033>.
- [2] L.V. Kalia, A.E. Lang, Parkinson's disease, *Lancet* 386 (9996) (2015) 896–912, [https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3).
- [3] P. Mahlknecht, K. Seppi, W. Poewe, The concept of prodromal Parkinson's disease, *J. Parkinsons Dis.* 5 (4) (2015) 681–697, <https://doi.org/10.3233/JPD-150685>.
- [4] S. Greffard, M. Verny, A.M. Bonnet, et al., Motor score of the unified Parkinson disease rating scale as a good predictor of lewy body-associated neuronal loss in the substantia nigra, *Arch. Neurol.* 63 (4) (2006) 584–588, <https://doi.org/10.1001/archneur.63.4.584>.
- [5] R.B. Postuma, D. Berg, Prodromal Parkinson's disease: the decade past, the decade to come, *Mov. Disord.* 34 (5) (2019) 665–675, <https://doi.org/10.1002/mds.27670>.
- [6] M.M. Ponsen, D. Stoffers, J. Booij, B.L.F. Van Eck-Smit, E.C. Wolters, H. W. Berendse, Idiopathic hyposmia as a preclinical sign of Parkinson's disease, *Ann. Neurol.* 56 (2) (2004) 173–181, <https://doi.org/10.1002/ana.20160>.
- [7] A. Ascherio, S.M. Zhang, M.A. Hernán, et al., Prospective study of caffeine consumption and risk of Parkinson's disease in men and women, *Ann. Neurol.* 50 (1) (2001) 56–63, <https://doi.org/10.1002/ana.1052>.
- [8] A. Gaenslen, I. Wurster, K. Brockmann, et al., Prodromal features for Parkinson's disease - baseline data from the TREND study, *Eur. J. Neurol.* 21 (5) (2014) 766–772, <https://doi.org/10.1111/ene.12382>.
- [9] P. Mahlknecht, A. Gasperi, A. Djamshidian, et al., Performance of the Movement Disorders Society criteria for prodromal Parkinson's disease: a population-based 10-year study, *Mov. Disord.* 33 (3) (2018) 405–413, <https://doi.org/10.1002/mds.27281>.
- [10] A. Schrag, L. Horsfall, K. Walters, A. Noyce, I. Petersen, Prediagnostic presentations of Parkinson's disease in primary care: a case-control study, *Lancet Neurol.* 14 (1) (2015) 57–64, [https://doi.org/10.1016/S1474-4422\(14\)70287-X](https://doi.org/10.1016/S1474-4422(14)70287-X).
- [11] H. Gustafsson, A. Nordström, P. Nordström, Depression and subsequent risk of Parkinson disease A nationwide cohort study, *Neurology* 84 (24) (2015) 2422–2429, <https://doi.org/10.1212/WNL.0000000000001684>.
- [12] R. Vuokko, P. Mäkelä-Bengs, H. Hyppönen, M. Lindqvist, P. Doupi, Impacts of structuring the electronic health record: results of a systematic literature review from the perspective of secondary use of patient data, *Int. J. Med. Inf.* 97 (2017) 293–303, <https://doi.org/10.1016/j.ijmedinf.2016.10.004>.
- [13] Parkinson's disease, Current Care Guidelines. Working Group Set up by the Finnish Medical Society Duodecim and the Finnish Neurological Society, The Finnish Medical Society Duodecim, Helsinki, 2019 (referred Mar 31 2021). Available online at: <https://www.kaypahoito.fi>. Parkinson's disease: Current Care Guidelines Abstract, 2019.
- [14] R.B. Postuma, D. Berg, M. Stern, et al., MDS clinical diagnostic criteria for Parkinson's disease, *Mov. Disord.* 30 (12) (2015) 1591–1601, <https://doi.org/10.1002/mds.26424>.
- [15] A.H.V. Schapira, K.R. Chaudhuri, P. Jenner, Non-motor features of Parkinson disease, *Nat. Rev. Neurosci.* 18 (7) (2017) 435–450, <https://doi.org/10.1038/nrn.2017.62>.
- [16] D. Santos-García, E. Suárez-Castro, J. Hernandez, et al., Predictors of mortality in nondemented patients with Parkinson disease: motor symptoms versus nonmotor symptoms, *J. Geriatr. Psychiatr. Neurol.* 31 (1) (2018) 19–26, <https://doi.org/10.1177/0891988717743589>.
- [17] P. Bugalho, F. Ladeira, R. Barbosa, et al., Motor and non-motor function predictors of mortality in Parkinson's disease, *J. Neural. Transm.* 126 (11) (2019) 1409–1415, <https://doi.org/10.1007/s00702-019-02055-3>.
- [18] L.M.L. De Lau, D. Verbaan, J. Marinus, J.J. van Hilten, Survival in Parkinson's disease. Relation with motor and non-motor features, *Park. Relat. Disord.* 20 (6) (2014) 613–616, <https://doi.org/10.1016/j.parkreldis.2014.02.030>.
- [19] E.B. Forsaa, J.P. Larsen, T. Wentzel-Larsen, G. Alves, What predicts mortality in Parkinson disease?: a prospective population-based long-term study, *Neurology* 75 (14) (2010) 1270–1276, <https://doi.org/10.1212/WNL.0b013e3181f61311>.
- [20] G. Levy, M.X. Tang, E.D. Louis, et al., The association of incident dementia with mortality in PD, *Neurology* 59 (11) (2002) 1708–1713, <https://doi.org/10.1212/01.WNL.0000036610.36834.E0>.
- [21] D. Bäckström, G. Granäsén, M.E. Domellöf, et al., Early predictors of mortality in parkinsonism and Parkinson disease A population-based study, *Neurology* 91 (22) (2018) E2045–E2056, <https://doi.org/10.1212/WNL.0000000000006576>.
- [22] E. De Pablo-Fernandez, C. Tur, T. Revesz, A.J. Lees, J.L. Holton, T.T. Warner, Association of autonomic dysfunction with disease progression and survival in Parkinson disease, *JAMA Neurol.* 74 (8) (2017) 970–976, <https://doi.org/10.1001/jamaneurol.2017.1125>.
- [23] J. Xu, D.D. Gong, C.F. Man, Y. Fan, Parkinson's disease and risk of mortality: meta-analysis and systematic review, *Acta Neurol. Scand.* 129 (2) (2014) 71–79, <https://doi.org/10.1111/ane.12201>.
- [24] A. Antonini, M. Tinazzi, G. Abbuzzese, et al., Pain in Parkinson's disease: facts and uncertainties, *Eur. J. Neurol.* 25 (7) (2018) 917–e69, <https://doi.org/10.1111/ene.13624>.
- [25] L. Nègre-Pagès, W. Regragui, D. Bouhassira, H. Grandjean, O. Rascol, Chronic pain in Parkinson's disease: the cross-sectional French DoPaMiP survey, *Mov. Disord.* 23 (10) (2008) 1361–1369, <https://doi.org/10.1002/mds.22142>.
- [26] D.P. Breen, J.R. Evans, K. Farrell, C. Brayne, R.A. Barker, Determinants of delayed diagnosis in Parkinson's disease, *J. Neurol.* 260 (8) (2013) 1978–1981, <https://doi.org/10.1007/s00415-013-6905-3>.
- [27] Y. Wan, Y. Zhu, Y. Luo, et al., Determinants of diagnostic latency in Chinese people with Parkinson's disease, *BMC Neurol.* 19 (1) (2019), <https://doi.org/10.1186/s12883-019-1323-5>.
- [28] L.M.L. De Lau, D. Verbaan, J. Marinus, J.J. van Hilten, Survival in Parkinson's disease. Relation with motor and non-motor features, *Park. Relat. Disord.* 20 (6) (2014) 613–616, <https://doi.org/10.1016/j.parkreldis.2014.02.030>.
- [29] S.E. Starkstein, P.L. Bolduc, H.S. Mayberg, T.J. Preziosi, R.G. Robinson, Cognitive impairments and depression in Parkinson's disease: a follow up study, *J. Neurol. Neurosurg. Psychiatry* 53 (7) (1990) 597–602, <https://doi.org/10.1136/jnnp.53.7.597>.
- [30] M.R. Cowie, J.I. Blomster, L.H. Curtis, et al., Electronic health records to facilitate clinical research, *Clin. Res. Cardiol.* 106 (1) (2017) 1, <https://doi.org/10.1007/s00392-016-1025-6>.