Diabetes/Metabolism Research and Reviews

WILEY

RESEARCH ARTICLE OPEN ACCESS

Prediction Model for Polyneuropathy in Recent-Onset Diabetes Based on Serum Neurofilament Light Chain, Fibroblast Growth Factor-19 and Standard Anthropometric and Clinical Variables

Haifa Maalmi[1,2] | Phong B. H. Nguyen[2,3,4] | Alexander Strom[1,2] | Gidon J. Bönhof[1,2,5] | Wolfgang Rathmann[2,6] | Dan Ziegler[1,2,5] | Michael P. Menden[2,3,4] | Michael Roden[1,2,5] | Christian Herder[1,2,5] | GDS Group

[1]Institute for Clinical Diabetology, German Diabetes Center (Deutsches Diabetes-Zentrum/DDZ), Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany | [2]German Center for Diabetes Research (DZD), München-Neuherberg, Germany | [3]Department of Computational Health, Helmholtz Center Munich, Neuherberg, Germany | [4]Department of Biology, Ludwig-Maximilians University Munich, Martinsried, Germany | [5]Department of Endocrinology and Diabetology, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany | [6]Institute for Biometrics and Epidemiology, German Diabetes Center (Deutsches Diabetes-Zentrum/DDZ), Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Correspondence: Haifa Maalmi (Haifa.maalmi@DDZ.de) | Christian Herder (christian.herder@ddz.de)

Received: 19 April 2023 | **Revised:** 20 August 2024 | **Accepted:** 28 October 2024

Funding: This work was initially exclusively supported by DDZ, which is funded by the German Federal Ministry of Health (Berlin, Germany) and the Ministry of Culture and Science of the state North Rhine-Westphalia (Düsseldorf, Germany) and as a multicenter study now receives additional funding by the German Federal Ministry of Education and Research (BMBF) through the German Center for Diabetes Research (DZD e.V.). Third-party funding supports sub-projects within GDS. The German Diabetes Study (GDS) was initiated and financed by the German Diabetes Center (DDZ), which is funded by (i) the German Federal Ministry of Health (Berlin, Germany), (ii) the Ministry of Culture and Science of the state of North Rhine-Westphalia (Düsseldorf, Germany) and (iii) grants from the German Federal Ministry of Education and Research (Berlin, Germany) to the German Center for Diabetes Research e.V. (DZD). This study was funded by a grant from the German Diabetes Association (DDG) to H.M. The funders had no role in the study design, data collection, analysis, interpretation, or report writing.

Keywords: diabetes | diabetic neuropathy | inflammation | machine learning | nerve conduction study | neurological biomarkers | peripheral nervous system | peripheral neuropathy | quantitative sensory tests

ABSTRACT

Background: Diabetic sensorimotor polyneuropathy (DSPN) is often asymptomatic and remains undiagnosed. The ability of clinical and anthropometric variables to identify individuals likely to have DSPN might be limited. Here, we aimed to integrate protein biomarkers for reliably predicting present DSPN.

Methods: Using the proximity extension assay, we measured 135 neurological and protein biomarkers of inflammation in blood samples of 423 individuals with recent-onset diabetes from the German Diabetes Study (GDS). DSPN was diagnosed based on the Toronto Consensus Criteria. We constructed (i) a protein-based prediction model using LASSO logistic regression, (ii) an optimised traditional risk model with age, sex, waist circumference, height and diabetes type and (iii) a model combining both. All models were bootstrapped to assess the robustness, and optimism-corrected AUCs (95% CI) were reported.

Abbreviations: ADA, American Diabetes Association; AUC, area under the ROC curve; DSPN, diabetic sensorimotor polyneuropathy; FGF-19, fibroblast growth factor-19; GDS, German Diabetes Study; NCS, nerve conduction study; NDS, Neuropathy Disability Score; NFL, neurofilament light; NPX, normalised protein expression; NSS, Neuropathy Symptom Score; QST, quantitative sensory testing.

For a complete list of the GDS Group, see the Acknowledgements section.

Results: DSPN was present in 16% of the study population. LASSO logistic regression selected the neurofilament light chain (NFL) and fibroblast growth factor-19 (FGF-19) as the most predictive protein biomarkers for detecting DSPN in individuals with recent-onset diabetes. The protein-based model achieved an AUC of 0.66 (0.59, 0.73), while the traditional risk model had an AUC of 0.66 (0.61, 0.74). However, combined features boosted the model performance to an AUC of 0.72 (0.67, 0.79).
Conclusion: We developed a prediction model for DSPN in recent-onset diabetes based on two protein biomarkers and five standard anthropometric, demographic and clinical variables. The model has a fair discrimination performance and might be used to inform the referral of patients for further testing.

1 | Introduction

Identifying individuals with subclinical or asymptomatic diabetic sensorimotor polyneuropathy (DSPN) remains challenging despite the availability of various screening and diagnostic tools [1, 2]. Consequently, these individuals often remain undiagnosed until their first clinical signs or symptoms appear [3], driving a subsequent higher risk for associated burden (i.e., pain, foot ulceration, deformity and amputations) [1, 4].

The pathophysiology of DSPN implicates a complex interplay between metabolic and inflammatory processes triggered or amplified by multiple demographic, anthropometric, lifestyle and clinical risk factors (i.e., older age, obesity, higher height, smoking, alcohol, hypertension, dyslipidemia and diabetes duration) [2]. This complex pathogenic tableau leads to widespread detrimental modifications of distinct signalling pathways, ending with axonal loss and myelin damage [3]. Nerve conduction studies (NCS) are the gold standard methods to detect early subclinical morphologic abnormalities in large nerves [5], but unfortunately, they are not feasible in everyday clinical practice. Therefore, approaches using simple tools to identify patients with DSPN are needed.

Some studies have developed prediction models for DSPN using metabolic, demographic, anthropometric and clinical variables [6–9]. While these predictors represent crucial risk factors for DSPN development, they do not guarantee its presence, particularly in recent-onset diabetes. None of the available prediction models integrated predictors reflecting downstream pathogenic inflammatory pathways or morphologic alterations in peripheral nerves, which might be better indicators of present DSPN.

The present study aimed to integrate a large panel of protein biomarkers reflecting inflammatory and neurological processes and a small set of traditional risk factors to develop a prediction model that helps estimate the probability that DSPN is present in middle-aged individuals with recent-onset diabetes.

2 | Methods

2.1 | Study Design and Study Population

The German Diabetes Study (GDS) is an ongoing observational prospective study that evaluates the natural course of recently diagnosed diabetes and explores prognostic factors and mechanisms leading to the development of diabetes-related complications [10]. The inclusion criterion is a diagnosis of type 1 or type 2 diabetes according to ADA recommendations [11] within the last 12 months in individuals aged between 18 and 69 years. All participants underwent a comprehensive examination consisting of clinical tests, a face-to-face interview, standardised written questionnaires and detailed laboratory measurements at baseline. The GDS was conducted according to the Declaration of Helsinki, approved by the ethics committee of Heinrich Heine University, Düsseldorf, Germany (ref. 4508) and registered with ClinicalTrials.gov (registration no. NCT01055093). All participants provided written informed consent.

This analysis was based on the consecutive sampling of 504 participants with diabetes who entered the GDS cohort between September 2005 and December 2011, of whom 423 had available data for protein biomarkers measured with the OLINK inflammation [12] and the Neuro Exploratory panels [13] and complete data for age, sex, waist circumference, height and diabetes type. The study followed the Transparent Reporting of a multivariable prognostic model for Individual Prognosis or Diagnosis (TRIPOD) recommendations [14].

2.2 | Outcome Definition

We assessed DSPN with nerve conduction studies (NCSs), quantitative sensory testing (QST) and neurological examinations using the Neuropathy Disability Score (NDS) and the Neuropathy Symptom Score (NSS) as previously described [13]. We defined DSPN according to the Toronto Consensus criteria [5]. Following our previous report, we considered individuals with subclinical, confirmed asymptomatic and confirmed symptomatic DSPN as prevalent DSPN cases [13].

2.3 | Proteomic Profiling

Blood samples were collected from overnight fasted participants, centrifuged, aliquoted and stored at −80°C until further use. Aliquots were analysed with Olink Target 96 multiplex assays (OLINK, Uppsala, Sweden) on two commercially available panels named Inflammation and Neuro Exploratory, each consisting of 92 protein biomarkers.

The inflammation panel includes pro- and anti-inflammatory cytokines, chemokines, enzymes, receptors and growth factors covering biological processes of inflammation, angiogenesis, fibrosis and endothelial activation. The Neuro Exploratory panel includes a combination of exploratory and established markers (e.g., neurofilament light polypeptide [NFL]) focusing

on neurological processes such as axon development, neurogenesis and synapse assembly.

The Olink assays are based on the proximity extension assay (PEA) technology, where pairs of oligonucleotide-labelled antibodies bind to each targeted protein. If both antibodies bind in close proximity, a polymerase chain reaction (PCR) target is produced and quantified using standard real-time PCR. PEA technology allows the relative quantification of analyte concentrations, and results are reported as normalised protein expression (NPX) values (comparable to the log2 scaled values).

Of the 184 proteins included in the two panels, we used 135 biomarkers for the present analysis and excluded 49 biomarkers with values below the limit of detection (LOD) ≥ 25% of all samples or inter-/intra-assay coefficients of variation (CV) > 25%. A detailed description of the 135 protein biomarkers analysed is given in Table S1.

2.4 | Statistical Analysis

We reported percentages (%) for categorical variables and means with standard deviations for continuous variables in descriptive statistics. We examined the differences between individuals with and without DSPN using chi-squared and *t*-tests. We evaluated the correlations between 135 protein biomarkers with Pearson correlation coefficients (*r*).

In the present study, we implemented a prediction approach previously reported elsewhere [15]. We constructed three prediction models. First, we used logistic regression with the least absolute shrinkage and selection operator (LASSO) to build the protein-based model with the SBC statistic's minimum value [16]. Second, we constructed an optimised traditional risk model including age, sex, waist circumference, height and diabetes type. Older age and higher waist circumference were reported as independent predictors of peripheral neuropathy in newly diagnosed patients with type 2 diabetes from the Anglo-Danish-Dutch study of Intensive Treatment of Diabetes in Primary Care (ADDITION) [17]. We included height because taller people were found to have a higher risk for DSPN [3]. We included sex and diabetes type because DSPN presence might differ in men versus women and type 1 versus type 2 diabetes [18]. Third, we built a final combined model integrating the variables of the first and the second models.

We performed internal validation of each of the three models by bootstrapping [19, 20]. Model performance was evaluated by assessing discrimination with the area under the receiving operating characteristic (ROC) curve and their 95% confidence intervals. We reported naïve AUCs with their 95% confidence intervals, optimism value and optimism-corrected AUCs with their location-shifted bootstrap confidence intervals [21]. To determine to what extent the combined model (main model) adds performance value compared to the protein-based or the traditional risk model, we compared its AUC to those of the protein-based and traditional risk models with the method of DeLong et al., which takes into account that the three models were derived from the same study sample [22]. The maximum of Youden's index was calculated to determine the optimal threshold value for maximising sensitivity + (specificity −1). Based on this cut-off point, we calculated the predicted probability of having DSPN, sensitivity, specificity and positive and negative predictive values of the three prediction scores.

We conducted all statistical analyses with SAS version 9.4 (SAS Institute, Cary, NC, USA). p values < 0.05 were considered to indicate statistically significant differences.

3 | Results

3.1 | Study Population

Prevalent DSPN was present in 66 (16%) out of 423 study participants (subclinical DSPN, $n = 41$; confirmed asymptomatic DSPN, $n = 11$; and confirmed symptomatic DSPN, $n = 14$). The baseline characteristics of individuals with and without DSPN are shown in Table 1. Briefly and as shown before [13], prevalent DSPN cases were more likely to be older, taller and male, and they were more likely to have higher waist circumferences and to be diagnosed with type 2 diabetes. However, individuals with and without DSPN did not differ in their BMI, HbA1c, or triglyceride levels.

Pairwise Pearson correlations showed weak correlations between most protein biomarkers (most correlation coefficients were < 0.25; Figure S1).

3.2 | Prediction of DSPN

LASSO logistic regression selected the neurofilament light chain (NFL) and the fibroblast growth factor-19 (FGF-19) as the most predictive protein biomarkers for detecting DSPN in individuals with recent-onset diabetes. Serum NPX levels of these two proteins were significantly higher in the 66 prevalent DSPN cases (NFL, mean ± SD = 4.1 ± 0.8; FGF-19, 8.3 ± 0.9) compared to 357 non-DSPN individuals (NFL, 3.7 ± 0.6; FGF-19, 7.9 ± 0.9) ($p = 0.000084$ and 0.0045 for serum NFL and FGF-19, respectively) (Figure 1). Pearson correlation shows a weak correlation between serum NFL and FGF-19 ($r = 0.15$, $p = 0.0023$).

The prediction performances of the three developed models are reported in Table 2. The protein-based model resulted in a naive AUC (95% CI) of 0.67 (0.59, 0.74) and an optimism-corrected AUC (95% CI) of 0.66 (0.59, 0.73). The traditional risk model yielded a naïve AUC of 0.69 (0.62, 0.76) and an optimism-corrected AUC of 0.66 (0.61, 0.74). The combined model yielded a naïve AUC of 0.75 (0.68, 0.82) and an optimism-corrected AUC of 0.72 (0.67, 0.79). The comparison of the AUCs of the combined model versus the protein-based or the traditional risk models was statistically significant ($p = 0.008$). Sensitivity at the maximum of Youden's index was 0.58, 0.77 and 0.77 for the protein-based, the traditional risk and the combined models, respectively; specificity was 0.69, 0.58, and 0.65 for the protein-based, the traditional risk and the combined models, respectively. Positive predictive values ranged from 0.26 to 0.29, while

TABLE 1 | Baseline characteristics of the study population (n = 423).

	Prevalent DSPN (n = 66, 16%)	No prevalent DSPN (n = 357, 84%)	p-value
Age, years	49.1 (12.5)	45.6 (14.7)	0.043
Sex (men/women), %	85/15	61/39	0.0002
Waist circumference, cm	103.2 (16.1)	97.1 (16.5)	0.005
Height, cm	176.1 (9.3)	173.3 (9.7)	0.029
BMI, kg/m^2	30.2 (6.0)	28.7 (6.2)	0.060
Diabetes type (type 1/type 2), %	20/80	40/60	0.002
HbA1c, mmol/mol	50.2 (14.4)	47.6 (11.8)	0.165
Triglycerides, mmol/L	1.2 (0.8, 1.7)	1.1 (0.7, 1.9)	0.686

Note: Data are presented as mean (SD), median (25th, 75th percentiles) or percentages (%).

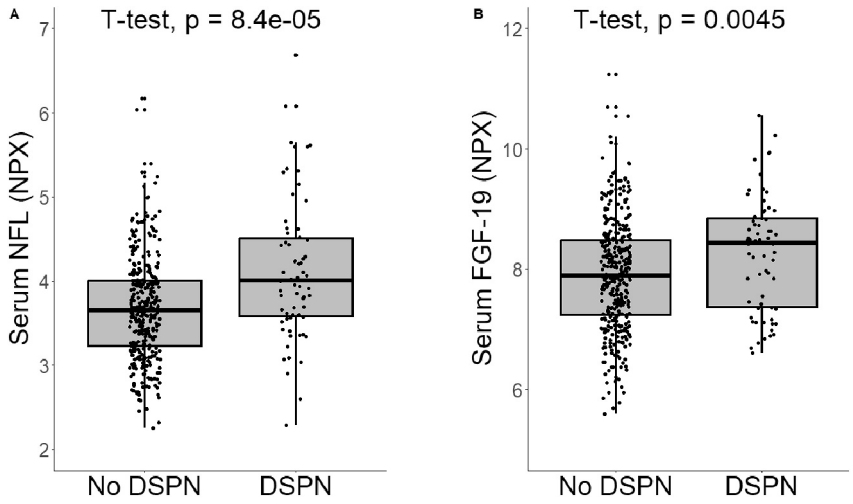

FIGURE 1 | Serum NFL and FGF-19 levels by DSPN status. The line that divides the box into two parts represents the median of the data. The top and bottom of the box show the upper (Q3) and lower (Q1) quartiles. The extreme line shows Q3 + 1.5 × IQR to Q1 − 1.5 × IQR. Serum NFL and FGF-19 are expressed as NPX values. DSPN, diabetic sensorimotor polyneuropathy; FGF-19, fibroblast growth factor-19; NFL, neurofilament light.

TABLE 2 | Discrimination performances of prediction models.

	Protein-based model[a]	Traditional risk model[b]	Combined model[c]
Discrimination performance parameters			
Naive AUC (95% CI)	0.67 (0.59, 0.74)	0.69 (0.62, 0.76)	0.75 (0.68, 0.82)
Maximum Youden Index	0.26	0.35	0.42
Sensitivity	0.58	0.77	0.77
Specificity	0.69	0.58	0.65
Positive predictive value	0.26	0.25	0.29
Negative predictive value	0.90	0.93	0.94
Internal validation			
Optimism	0.008	0.03	0.03
Optimism-corrected AUC (95% CI)	0.66 (0.59, 0.73)	0.66 (0.61, 0.74)	0.72 (0.67, 0.79)

Note: p-value for the comparison of the optimism-corrected AUCs of the combined model versus the protein-based or the traditional risk models is 0.008.
Abbreviations: AUC, area under the ROC curve; CI, confidence interval.
[a]The protein-based model comprises serum NFL and serum FGF-19.
[b]The traditional risk model comprises age, sex, waist circumference, height and diabetes type.
[c]The combined model comprises serum NFL, serum FGF-19, age, sex, waist circumference, height and diabetes type.

negative predictive values were much higher across the three models (0.90 for the protein-based model, 0.93 for the traditional risk model and 0.94 for the combined model).

The confusion matrix (Table 3) showed that the protein-based model scored 65% of our study population as not having DSPN, with 67% of participants with and without DSPN being correctly diagnosed (true positive + true negative). Moreover, the traditional risk model scored 52% of our study population as not having DSPN, with 61% of participants with and without DSPN being correctly diagnosed. Finally, our main predictive model (combined model) scored 58% of our study population as not having DSPN, with 67% of participants with and without DSPN being correctly diagnosed (true positive = 12%, true negative = 55%).

4 | Discussion

A prediction model integrating serum NFL and FGF-19 and age, sex, waist circumference, height and diabetes type performs well in predicting prevalent DSPN in recent-onset diabetes.

NFL is the most abundant cytoskeleton filament protein in myelinated axons, providing structural support for the axon [23]. Neuronal degeneration releases NFL into the cerebrospinal fluid and blood, and substantial evidence from epidemiological studies has reported increased NFL levels in various neurodegenerative diseases, including traumatic brain injury, multiple sclerosis, frontotemporal dementia and Alzheimer's disease [24]. Axonal damage of peripheral nerves also releases NFL from peripheral nerves into the blood, with a growing body of evidence linking higher serum NFL levels to peripheral neuropathies from different aetiologies [25]. We recently showed that higher serum NFL levels were associated with higher prevalence of DSPN and nerve dysfunction in individuals with diabetes free of other neurodegenerative diseases [13]. It is worth noting that the present study is a secondary analysis of data used in our aforementioned primary study but aims to address a different research question through different statistical approaches. While our primary study aimed to identify individual neurological proteins involved in the disease process leading to DSPN using aetiological models (multivariable regression models) [13], the present study aimed to build a prediction model that estimates the probability of having DSPN based on a small set of variables selected based on their ability to improve prediction and model performance. Differences between aetiological and prediction models are explained elsewhere [26]. It is indeed important to note that compared to the primary study, the present study extended data on protein biomarkers by adding 92 biomarkers of inflammation.

FGF-19 is a hormone secreted by the small intestine and is a member of a subfamily of FGFs that includes FGF-21 and FGF-23. Animal models reported insulin-like actions for FGF-19 in the liver (i.e., inducing protein and glycogen synthesis and suppressing gluconeogenesis) [27]. Indeed, animal studies have reported an upregulation in the expression of FGF-19 and other growth factors in the dorsal root ganglia of rat sciatic nerve after sciatic nerve injury, and this upregulation promotes axon regrowth and nerve repair [28]. In humans, a previous epidemiological study showed that higher serum FGF-19 was associated with an increased risk of incident DSPN in older individuals [29].

In the present study, higher serum FGF-19 and NFL in individuals with DSPN versus those without DSPN suggest some evidence for an ongoing repair process in peripheral nerves after nerve injury.

We previously reported that six protein biomarkers of inflammation (i.e., chemokines and soluble receptors) had a fair performance in predicting incident DSPN in older individuals when added to clinical parameters (AUC of 0.783) [29]. However, this study selected protein biomarkers based on single associations of each protein with DSPN after adjustment for confounders and multiple testing. In contrast, our study used automatic variable selection where all protein biomarkers were assessed simultaneously, considering multicollinearity. Indeed, compared to

TABLE 3 | Confusion matrix summarising the predicted probabilities of the three predictive scores according to the presence or absence of DSPN.

	Prevalent DSPN		
	Yes	**No**	**Total**
Protein-based model[a]			
High likelihood of DSPN	TP (38, 9%)	FP (111, 26%)	149 (35%)
Low likelihood of DSPN	FN (28, 7%)	TN (246, 58%)	274 (65%)
Traditional risk model[b]			
High likelihood of DSPN	TP (51, 12%)	FP (150, 35%)	201 (48%)
Low likelihood of DSPN	FN (15, 4%)	TN (207, 49%)	222 (52%)
Combined model[c]			
High likelihood of DSPN	TP (51, 12%)	FP (126, 29%)	177 (42%)
Low likelihood of DSPN	FN (15, 4%)	TN (231, 55%)	246 (58%)

Abbreviations: DSPN, diabetic sensorimotor polyneuropathy; FN, false negative; FP, false positive, TN, true negative; TP, true positive.
[a]The protein-based model comprises serum NFL and serum FGF-19.
[b]The traditional risk model comprises age, sex, waist circumference, height and diabetes type.
[c]The combined model comprises serum NFL, serum FGF-19, age, sex, waist circumference, height and diabetes type.

other studies [6–9, 30], our study incorporated a relatively minor number of anthropometric, demographic and clinical variables, which are easy to collect in a routine clinical setting [1, 31]. Our study constructed a prediction model for DSPN in individuals with type 1 and type 2 diabetes, while other studies were restricted to individuals with type 2 diabetes [6–9]. Though DSPN affects 6%–20% of adults with type 1 diabetes at the onset of the disease [18, 32], current guidelines recommend starting DSPN screening for these individuals 5 years after disease onset [1], a lag that might hamper DSPN management and treatment [33] in those developing the disease earlier. Based on the above, it is of added value that we tested the performance of our predictive model in type 1 and type 2 diabetes. However, because of the small number of cases, subgroup analysis by diabetes type was not feasible.

We used logistic regression to construct our prediction model, while two other studies used non-linear machine-learning classifiers with random forest found as the best algorithm [7, 8]. Nevertheless, some evidence in the literature shows that conventional logistic regression performs similarly to optimised non-linear machine-learning algorithms in clinical settings [34].

The very high negative predictive value (0.94) makes using our combined model gives insight to the healthcare provider with respect to the non-need to pursue other more advanced testing to diagnose DSPN in individuals classified as having a low likelihood. However, due to the relatively low positive predictive value (0.29), our combined model might not be very informative in evaluating the chances of having DSPN among those classified as having a high likelihood.

The present study's strengths include using gold-standard outcome measures and the PEA technology to measure serum protein biomarkers. Another strength is evaluating the predictive performance of several nerve-specific protein biomarkers covering neurological processes such as axon development, neurogenesis and synapse assembly, and biomarkers of systemic inflammation covering biological processes of inflammation, angiogenesis, fibrosis and endothelial activation. The main limitation of the present study is the lack of external validation due to the lack of studies with similar data. Our study aimed to build a model that estimates the probability of present DSPN. Thus, estimating the probability that DSPN will occur in the future in DSPN-free individuals is beyond the scope of our study. We constructed our predictive model in middle-aged individuals with short diabetes duration; hence, to what extent our model applies to younger individuals or people with longer diabetes duration is unknown.

5 | Conclusion

We developed a prediction model for prevalent DSPN based on serum NFL, a protein biomarker of nerve injury, serum FGF-19, a growth factor hormone, and age, sex, waist circumference, height and diabetes type. Our model's high negative predictive and true negative values suggest that it can help healthcare providers prevent performing further non-beneficial testing of DSPN with costly and time-consuming tools.

Author Contributions

H.M. acquired funding, performed the statistical analysis and wrote the manuscript. H.M., M.P.M. and C.H. designed the study. A.S., G.J.B., D.Z., C.H., W.R. and M.R. contributed data. C.H., M.P.M., P.B.H.N., M. R., D.Z. and W.R. contributed to data interpretation. All authors reviewed and edited the manuscript and approved its submission.

Acknowledgements

We appreciate the contribution of all study participants and thank the Clinical Research Center staff at the Institute for Clinical Diabetology, German Diabetes Center, for the data collection. Data from this study were presented at the Diabetes congress of the German Diabetes Association (DDG) on 17-20 Mai 2023 in Berlin. The GDS Group consists of M. Roden (speaker), H. Al-Hasani, B. Belgardt, G. Bönhof, V. Burkart, A. E. Buyken, G. Geerling, C. Herder, A. Icks, K. Jandeleit-Dahm, J. Kotzka, O. Kuß, E. Lammert, W. Rathmann, V. Schrauwen-Hinderling, J. Szendrödi, S. Trenkamp, R. Wagner and their co-workers who are responsible for the design and conduct of the GDS.

Conflicts of Interest

H.M. received a research grant from EFSD/Novo Nordisk outside the submitted work and is an editorial board member of DMRR. CH is an editorial board member of DMRR. W.R. reports receiving consulting fees for attending educational sessions or advisory boards from AstraZeneca, Boehringer-Ingelheim and Novo Nordisk and institutional research grants from Novo Nordisk outside of the topic of the current work. M.R. received fees as a member of advisory boards or speaker from Allergan, Boehringer-Ingelheim Pharma, Bristol-Myers Squibb, Eli Lilly, Fishawack Group, Gilead Sciences, Novartis Pharma, Intercept Pharma, Inventiva, Novo Nordisk, AstraZeneca, Sanofi US, Prosciento, Target RWE and Terra Firma and has been involved with clinical trial research for Boehringer-Ingelheim, Danone Nutricia Research and Sanofi-Aventis, all outside the submitted work. All other authors declare that no relationships or activities might bias their work.

Data Availability Statement

The data are subject to national data protection laws. Thus, data cannot be made freely available in a public repository. Nevertheless, data can be requested through an individual project agreement with the GDS Steering Committee (speaker: M. Roden, michael.roden@ddz.de).

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/dmrr.70009.

References

1. R. Pop-Busui, A. J. M. Boulton, E. L. Feldman, et al., "Diabetic Neuropathy: A Position Statement by the American Diabetes Association," *Diabetes Care* 40, no. 1 (2016): 136–154, https://doi.org/10.2337/dc16-2042.

2. G. J. Bönhof, C. Herder, A. Strom, N. Papanas, M. Roden, and D. Ziegler, "Emerging Biomarkers, Tools, and Treatments for Diabetic Polyneuropathy," *Endocrine Reviews* 40, no. 1 (2019): 153–192, https://doi.org/10.1210/er.2018-00107.

3. E. L. Feldman, B. C. Callaghan, R. Pop-Busui, et al., "Diabetic Neuropathy," *Nature Reviews Disease Primers* 5, no. 1 (2019): 41, https://doi.org/10.1038/s41572-019-0092-1.

4. S. D. Candrilli, K. L. Davis, H. J. Kan, M. A. Lucero, and M. D. Rousculp, "Prevalence and the Associated Burden of Illness of Symptoms of Diabetic Peripheral Neuropathy and Diabetic Retinopathy," *Journal of Diabetes and its Complications* 21, no. 5 (2007): 306–314, https://doi.org/10.1016/j.jdiacomp.2006.08.002.

5. P. J. Dyck, J. W. Albers, H. Andersen, et al., "Diabetic Polyneuropathies: Update on Research Definition, Diagnostic Criteria and Estimation of Severity," *Diabetes Metabolism Research and Reviews* 27, no. 7 (2011): 620–628, https://doi.org/10.1002/dmrr.1226.

6. B. W. Bongaerts, D. Ziegler, J. E. Shaw, et al., "A Clinical Screening Score for Diabetic Polyneuropathy: KORA F4 and AusDiab Studies," *Journal of Diabetes and its Complications* 29, no. 1 (2015): 44–49, https://doi.org/10.1016/j.jdiacomp.2014.09.014.

7. F. Haque, I. R. M. Bin, M. E. H. Chowdhury, et al., "Performance Analysis of Conventional Machine Learning Algorithms for Diabetic Sensorimotor Polyneuropathy Severity Classification," *Diagnostics* 11, no. 5 (2021): 801, https://doi.org/10.3390/diagnostics11050801.

8. D. Y. Shin, B. Lee, W. S. Yoo, J. W. Park, and J. K. Hyun, "Prediction of Diabetic Sensorimotor Polyneuropathy Using Machine Learning Techniques," *Journal of Clinical Medicine* 10, no. 19 (2021): 4576, https://doi.org/10.3390/jcm10194576.

9. W. Zhang and L. Chen, "A Nomogram for Predicting the Possibility of Peripheral Neuropathy in Patients With Type 2 Diabetes Mellitus," *Brain Sciences* 12, no. 10 (2022): 1328, https://doi.org/10.3390/brainsci12101328.

10. J. Szendroedi, A. Saxena, K. S. Weber, et al., "Cohort Profile: The German Diabetes Study (GDS)," *Cardiovascular Diabetology* 15, no. 1 (2016): 59, https://doi.org/10.1186/s12933-016-0374-9.

11. American Diabetes Association, "2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018," supplement, *Diabetes Care* 41, no. S1 (2017): S13–S27, https://doi.org/10.2337/dc18-s002.

12. H. Maalmi, C. Herder, K. Strassburger, et al., "Biomarkers of Inflammation and Glomerular Filtration Rate in Individuals With Recent-Onset Type 1 and Type 2 Diabetes," *Journal of Clinical Endocrinology and Metabolism* 105, no. 12 (2020): e4370–e4381, https://doi.org/10.1210/clinem/dgaa622.

13. H. Maalmi, A. Strom, A. Petrera, et al., "Serum Neurofilament Light Chain: A Novel Biomarker for Early Diabetic Sensorimotor Polyneuropathy," *Diabetologia* 66, no. 3 (2023): 579–589, https://doi.org/10.1007/s00125-022-05846-8.

14. K. G. Moons, D. G. Altman, J. B. Reitsma, et al., "Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): Explanation and Elaboration," *Annals of Internal Medicine* 162, no. 1 (2015): W1–W73, https://doi.org/10.7326/m14-0698.

15. S. C. Chung, B. O'Brien, G. Y. H. Lip, et al., "Prognostic Model for Atrial Fibrillation After Cardiac Surgery: A UK Cohort Study," *Clinical Research in Cardiology* 112, no. 2 (2023): 227–235, https://doi.org/10.1007/s00392-022-02068-1.

16. B. Lund, "Logistic Model Selection With SAS® PROC's Logistic, Hplogistic, Hpgenselect," in *Paper AA02, Paper Presented at: MidWest SAS Users Group* (St Louis, MO, 2017).

17. S. T. Andersen, D. R. Witte, E.-M. Dalsgaard, et al., "Risk Factors for Incident Diabetic Polyneuropathy in a Cohort With Screen-Detected Type 2 Diabetes Followed for 13 Years: ADDITION-Denmark," *Diabetes Care* 41, no. 5 (2018): 1068–1075, https://doi.org/10.2337/dc17-2062.

18. M. Sempere-Bigorra, I. Julián-Rochina, and O. Cauli, "Differences and Similarities in Neuropathy in Type 1 and 2 Diabetes: A Systematic Review," *Journal of Personalized Medicine* 11, no. 3 (2021): 230, https://doi.org/10.3390/jpm11030230.

19. B. Efron, "Bootstrap Methods: Another Look at the Jackknife," in *Breakthroughs in Statistics: Methodology and Distribution*, eds. S. Kotz and N. L. Johnson (New York, NY: Springer New York, 1992), 569–593.

20. I. Lankham and M. Slaughter, "A Framework for Simple and Efficient Bootstrap Validation in SAS®, With Examples," *SAS Global Forum* (2021), Paper 1034-2021.

21. H. Noma, T. Shinozaki, K. Iba, S. Teramukai, and T. A. Furukawa, "Confidence Intervals of Prediction Accuracy Measures for Multivariable Prediction Models Based on the Bootstrap-Based Optimism Correction Methods," *Statistics in Medicine* 40, no. 26 (2021): 5691–5701, https://doi.org/10.1002/sim.9148.

22. E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach," *Biometrics* 44, no. 3 (1988): 837–845, https://doi.org/10.2307/2531595.

23. A. R. Gafson, N. R. Barthélemy, P. Bomont, et al., "Neurofilaments: Neurobiological Foundations for Biomarker Applications," *Brain* 143, no. 7 (2020): 1975–1998, https://doi.org/10.1093/brain/awaa098.

24. L. Gaetani, K. Blennow, P. Calabresi, M. Di Filippo, L. Parnetti, and H. Zetterberg, "Neurofilament Light Chain as a Biomarker in Neurological Disorders," *Journal of Neurology, Neurosurgery & Psychiatry* 90, no. 8 (2019): 870–881, https://doi.org/10.1136/jnnp-2018-320106.

25. J. Fundaun, M. Kolski, M. Molina-Álvarez, G. Baskozos, and A. B. Schmid, "Types and Concentrations of Blood-Based Biomarkers in Adults With Peripheral Neuropathies: A Systematic Review and Meta-Analysis," *JAMA Network Open* 5, no. 12 (2022): e2248593-e, https://doi.org/10.1001/jamanetworkopen.2022.48593.

26. E. H. Au, A. Francis, A. Bernier-Jean, and A. Teixeira-Pinto, "Prediction Modeling-Part 1: Regression Modeling," *Kidney International* 97, no. 5 (2020): 877–884, https://doi.org/10.1016/j.kint.2020.02.007.

27. S. Kir, S. A. Kliewer, and D. J. Mangelsdorf, "Roles of FGF19 in Liver Metabolism," *Cold Spring Harbor Symposia on Quantitative Biology* 76, (2011): 139–144, https://doi.org/10.1101/sqb.2011.76.010710.

28. Y. Y. Shen, X. K. Gu, R. R. Zhang, T. M. Qian, S. Y. Li, and S. Yi, "Biological Characteristics of Dynamic Expression of Nerve Regeneration Related Growth Factors in Dorsal Root Ganglia After Peripheral Nerve Injury," *Neural Regeneration Research* 15, no. 8 (2020): 1502–1509, https://doi.org/10.4103/1673-5374.274343.

29. C. Herder, J. M. Kannenberg, M. Carstensen-Kirberg, et al., "A Systemic Inflammatory Signature Reflecting Cross Talk Between Innate and Adaptive Immunity Is Associated With Incident Polyneuropathy: KORA F4/FF4 Study," *Diabetes* 67, no. 11 (2018): 2434–2442, https://doi.org/10.2337/db18-0060.

30. Y. Fan, E. Long, L. Cai, Q. Cao, X. Wu, and R. Tong, "Machine Learning Approaches to Predict Risks of Diabetic Complications and Poor Glycemic Control in Nonadherent Type 2 Diabetes," *Frontiers in Pharmacology* 12 (2021): 665951, https://doi.org/10.3389/fphar.2021.665951.

31. S. Schlesinger, C. Herder, J. M. Kannenberg, et al., "General and Abdominal Obesity and Incident Distal Sensorimotor Polyneuropathy: Insights into Inflammatory Biomarkers as Potential Mediators in the KORA F4/FF4 Cohort," *Diabetes Care* 42, no. 2 (2019): 240–247, https://doi.org/10.2337/dc18-1842.

32. A. Strom, K. Kaul, J. Brüggemann, et al., "Lower Serum Extracellular Superoxide Dismutase Levels Are Associated With Polyneuropathy in Recent-Onset Diabetes," *Experimental & Molecular Medicine* 49, no. 11 (2017): e394–e, https://doi.org/10.1038/emm.2017.173.

33. A. J. M. Boulton, "Management of Diabetic Peripheral Neuropathy," *Clinical Diabetes* 23, no. 1 (2005): 9–15, https://doi.org/10.2337/diaclin.23.1.9.

34. A. L. Lynam, J. M. Dennis, K. R. Owen, et al., "Logistic Regression Has Similar Performance to Optimised Machine Learning Algorithms in a Clinical Setting: Application to the Discrimination Between Type 1 and Type 2 Diabetes in Young Adults," *Diagnostic and Prognostic Research* 4, no. 1 (2020): 6, https://doi.org/10.1186/s41512-020-00075-2.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.