Background: There is no recognized gold standard for diagnosing nerve root involvement (NRI) in patients with low back–related leg pain (LLBP) and no consensus on which cluster of items might best identify NRI. Diagnostic models have mainly been developed in secondary care with conflicting reference standards and predictor selection. In the absence of a gold standard, clinical opinion and/or imaging findings are often used. Clinical opinion has issues of incorporation bias and imaging can incorrectly classify patients. This study explores the challenges of reference standard selection in NRI diagnostic modelling and aims to ascertain which combination of clinical assessment items best identify NRI in primary care LLBP consulters. Classification using clinical diagnosis is compared with a statistical approach that circumvents the need for a reference standard.

Methods: Two definitions of NRI formed the reference standards for diagnostic modelling: high confidence (≥80%) NRI clinical diagnosis; and high confidence (≥80%) NRI clinical diagnosis with confirmatory MRI findings. Cross-sectional data on 394 LLBP consulters were used to develop the model. Potential NRI indicators were seven clinical assessment items. Multivariable logistic regression models were constructed and compared for both reference standards. Model performances were summarized using the Hosmer–Lemeshow statistic and area under the curve (AUC). Bootstrapping assessed internal validity. The same variables were analysed using latent class analysis (LCA).

Results: The NRI clinical diagnosis model 1 retained five items. The clinical diagnosis plus MRI model 2 retained six items. Four items remained in both models: below knee pain, leg pain worse, positive neural tension and neurological deficit. The NRI clinical diagnosis model was well calibrated (P = 0.17) and discrimination was an AUC of 0.95 (95% CI 0.93, 0.98). The clinical diagnosis plus MRI model showed good discrimination [AUC 0.82 (95% CI 0.78, 0.86)] but poor calibration (P = 0.004). Overfitting was minimal in both. LCA identified three LLBP groups. Group 1 (n = 147) had high pain intensity (6.9/10) and very high probability (≥0.85) of three clinical indicators being positive [leg pain worse (P = 0.86), below knee pain (P = 0.92), positive neural tension (P = 0.92)]. Group 2 (n = 165) had moderate pain (4.8/10) and high probability (≥0.7) of below knee pain (P = 0.78) and neurological deficit (P = 0.73). Group 3 (n = 83) had low pain (3.3/10) and low probability (≤0.33) of any positive clinical indicators. There was very high agreement between the clinically diagnosed NRI group and latent class groups 1 and 2.

Conclusion: Following diagnostic modelling, four clinical assessment items were common in both reference standard definitions. Three of these items were highly probable in one LCA subgroup and two of the four items were probable in the second subgroup. This work suggests a combination of items that identify two diagnostic subgroups of LLBP consulters with NRI that could be used clinically and in research to improve homogeneous patient identification.

Disclosure statement: The authors have declared no conflicts of interest.