**SupplementalDigital Content 1.** Model selection strategy

To estimate the annual number of dorsal spinal fusions via Poisson regression, we used age, sex, and calendar year as covariates. The expected mean incidence was adjusted for population size using the age-specific logged population numbers as an offset. To overcome overdispersion problems that could result in an underestimation of variance, we used a robust sandwich covariance matrix estimator to calculate variance. To minimize underestimation of the error of variance of the estimated parameter because of overdispersion, we applied a quasi-Poisson regression to the data in accordance with the theory of quasi-likelihood. Given that Poisson regression can provide exponential growth to infinity, we also tried, in accordance with some other authors, an estimation model of future projections using logarithmic regression [30]. However, both models seem to be more economically driven than knowledge-driven, because regressions based on a logarithm or exponent will only fit optimally if this is the exact nature of the true relationship. To overcome this issue, time series forecasting has recently become increasingly popular. Exponential smoothing (ETS) and autoregressive integrated moving average (ARIMA) models are the two most widely used methods here and provide a complementary approach to the problem [8, 16, 17, 23]. Although ETS models are based on a description of the trend and seasonality in the data and weigh averages of past observations, with the weights decaying exponentially as the observations get older, ARIMA models aim to describe autocorrelation in the data. In both cases, future values are constrained to be linear functions of past observations. In this study, a grouped time series analysis for modelling time effects and accounting for respective sex and age groups was used. These models also contain combinations of ARIMA and ETS as introduced by Hyndman et al. [14]. Based on the stepwise Hyndman–Khandakar algorithm, the best model is automatically selected using a combination of the Kwiatkowski–Phillips–Schmidt–Shin unit root test, minimization of the Akaike information criterion for small samples, and maximum likelihood method algorithm [2].

The prediction accuracy of each model we used was verified by splitting the dataset into training (years 2005 to 2014) and testing subsets (years 2015 to 2019) and comparing the mean absolute percentage error (MAPE), the root mean squared error (RMSE), and the mean absolute error (MAE) of the models. The MAPE is a percentage error that measures the relative difference between the actual and predicted values of a given model. The lower the MAPE, the more accurate the model’s forecast is. The RMSE is defined as the square root of the mean square error. Smaller values indicate better accuracy. The RMSE value is in the same unit as the predicted value. The MAE is defined as the average of the absolute difference between predicted values and true values. It shows how large an error can be expected from the forecast on average. The error values are in the original units of the forecast values. The lower the MAE value, the better the model; a value of zero means there is no error in the forecast. The greater the difference between RMSE and MAE, the more inconsistent the error size is [13].