Appendix A:

Maximum likelihood estimation approach for estimating the parameters of the logistic regression model (Sogandi et al.[5])

The parameters of logistic regression model in Eq. (1) using MLE approach. The likelihood function for n patients at stage j^{th} is according to the following equation.

$$L\left(\boldsymbol{\gamma}_{jk_j}^T,\boldsymbol{\beta}_j\right)=\prod_{i=1}^{n}\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)^{y_{ij}}\left(1-\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)^{1-y_{ij}}\right). \tag{A1}$$

For the sake of simplicity, it is better to take the logarithm of the likelihood function. Hence, the log-likelihood function of the j^{th} stage can be shown as:

$$\begin{aligned}
\log\left(L\left(\boldsymbol{\gamma}_{jk_j}^T,\boldsymbol{\beta}_j\right)\right)&=\sum_{i=1}^{n}y_{ij}\log\left(\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)\right)+\left(1-y_{ij}\right)\log\left(1-\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)\right)\\
&=\sum_{i=1}^{n}\log\left(1-\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)\right)+\sum_{i=1}^{n}y_{ij}\log\left(\frac{\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)}{1-\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)}\right)\\
&=\sum_{i=1}^{n}\log\left(1-\pi_{ij}\left(\mathbf{d}_{ijk_j},\mathbf{x}_{ij},l_{ij}\right)\right)+\sum_{i=1}^{n}y_{ij}\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)\\
&=\sum_{i=1}^{n}-\log\left(1+\exp\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)\right)+\sum_{i=1}^{n}y_{ij}\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right).
\end{aligned} \tag{A2}$$

To find the maximum likelihood estimates, we usually differentiate the log likelihood with respect to the parameters. For this aim, we take the derivative with respect to the parameters vector and set it equal to zero as:

$$\frac{\partial L\left(\boldsymbol{\beta}_j,\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\right)}{\partial\boldsymbol{\varphi}_{jk_j}}=\left(\begin{aligned}
&-\sum_{i=1}^{n}\frac{\mathbf{x}_{ij}}{1+\exp\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)}\exp\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)+\sum_{i=1}^{n}y_{ij}\mathbf{x}_{ij},\\
&-\sum_{i=1}^{n}\frac{\sum_{k=1}^{K_j}\mathbf{d}_{ijk_j}}{1+\exp\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)}\exp\left(\sum_{k=1}^{K_j}\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\mathbf{d}_{ijk_j}+\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_{ij}+l_{ij}\right)+\sum_{i=1}^{n}\sum_{k=1}^{K_j}\mathbf{d}_{ijk_j}y_{ij}
\end{aligned}\right)=\mathbf{0} \tag{A3}$$

in which $\left(\boldsymbol{\beta}_j^{\mathrm{T}},\boldsymbol{\gamma}_{jk_j}^{\mathrm{T}}\right)=\boldsymbol{\varphi}_{jk_j}$. Since, there is no closed-form solution for equation (A3). We solve it numerically using Newton-Rophson approximation. In fact, the MLE is obtained by an iterative procedure known as the iterative weighted least squares.