Prediction of Long-Term Recovery From Disability Using Hemoglobin-Based Models: Results From a Cohort of 1,392 Patients Undergoing Spine Surgery

Matteo Briguglio 1,*, Paolo Perazzo 2, Francesco Langella 3, Tiziano Crespi 2, Elena De Vecchi 4, Patrizia Riso 5, Marisa Porrini 5, Laura Scaramuzzo 6, Roberto Bassani 7, Marco Brayda-Bruno 8, Giuseppe Banfi 1,9 and Pedro Berjano 3

Hemoglobin and its associated blood values are important laboratory biomarkers that mirror the strength of constitution of patients undergoing spine surgery. Along with the clinical determinants available during the preadmission visit, it is important to explore their potential for predicting clinical success from the patient’s perspective in order to make the pre-admission visit more patient-centered. We analyzed data from 1,392 patients with spine deformity, disc disease, or spondylolisthesis enrolled between 2016 and 2019 in our institutional Spine Registry. Patient-reported outcome measure at 17 months after surgery was referred to the Oswestry disability index. High preoperative hemoglobin was found to be the strongest biochemical determinant of clinical success along with high red blood cells count, while low baseline disability, prolonged hospitalization, and long surgical times were associated with poor recovery. The neural network model of these predictors showed a fair diagnostic performance, having an area under the curve of 0.726 and a sensitivity of 86.79%. However, the specificity of the model was 15.15%, thus providing to be unreliable in forecasting poor patient-reported outcomes. In conclusion, preoperative hemoglobin may be one of the key biomarkers on which to build appropriate predictive models of long-term recovery after spine surgery, but it is necessary to include multidimensional variables in the models to increase the reliability at the patient’s level.

Keywords: preoperative care, orthopedic procedures, vertebral column, hemoglobin, anemia, patient-reported outcome measures, enhanced recovery after surgery, spine surgery
INTRODUCTION
Spinal disorders are common and the prevalence increases in the aging population (1). In view of the recent advancements in spine treatments, surgery is considered clinically effective from a medical perspective. However, healthcare systems are progressively moving toward value-based business models (2), with the definition of clinical success increasingly based on patient-reported outcome measures (PROMs) (3). The Oswestry disability index (ODI) is a widely-used, validated, and self-administered questionnaire to assess a patient's functional impairment related to the spinal condition, encompassing questions about personal care, movements, sleeping, and social life (4), and can be used to evaluate the outcomes from the patient's perspective (5). The advent of PROMs calls for more patient-centric healthcare, which outlines the need for early preoperative pathways. Identifying the determinants that affect the individual's daily activities in the long term can help maintain high-quality standards (6). One of the parameters known to mirror the functional reservoirs required for proper recovery is hemoglobin (7–9). An abnormal circulating level before surgery is considered a risk factor for poor medical and surgical outcomes in spinal patients (10), also being a predictor of mortality in the most severe conditions (11). Hemoglobin is an assembly of four globular polypeptide chains that fill the warp of red blood cells, carrying up to four oxygen molecules attached to iron atoms. Together with the erythrocytes and their volume over total blood (i.e., hematocrit), hemoglobin reflects oxygen carrying capacity, functional iron levels, and correct erythropoiesis (12). Hemoglobin concentration is one of the benchmarks for planning transfusion therapy strategy (13) along with other laboratory parameters such as hematocrit (14). Therefore, preoperative iron optimization is considered a key aspect of patient blood management (15) and enhanced recovery after surgery (16). To the authors' knowledge, there are no studies in spine surgery that have investigated the potential of preoperative hemoglobin in predicting clinical success from a patient's point of view. We studied a large cohort of patients undergoing spine surgery for deformities, disc disease, and other back conditions to identify predictors of long-term functional status using hemoglobin-based models.

MATERIALS AND METHODS
Study Population
The research included patients enrolled in the institutional Spine Registry (SpineReg; ClinicalTrials.gov number: NCT03644407), which is a prospective observational registry recruiting patients undergoing spine surgery incepted in 2015 by our hospital IRCCS Orthopedic Institute Galeazzi of Milan (Italy). To answer the research question, we extracted from the registry the patients who met the following eligibility criteria: ≥18 years of age, enrolment between 2016 and 2019, the presence of at least one ODI assessment at 6-, 12-, or 24-month follow-up. The extraction excluded data of patients with a diagnosis of tumors, admission for complications, and surgical procedures involving the cervical spine. A 30% reduction from baseline was considered as the minimum clinically important difference (MCID) in the ODI score in order to categorize the outcomes (17). A secondary analysis was planned using the raw reduction of 12.7 points as classification threshold, which is more restrictive in categorizing successful surgeries (18). Patients who had a preoperative ODI < 12.7 were excluded from the extraction query. After data extraction and integration with routine parameters, the study sample comprised 1,392 patients with seventeen variables: gender, age, red blood cells (RBCs), hematocrit (Ht), serum concentration of hemoglobin (Hb), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), C-reactive protein (CRP), type of diagnosis, American society of anesthesiologists physical status classification system (ASA), duration of surgery (DS), length of hospital stay (LOS), preoperative ODI (PREOP ODI), 6-month ODI, 12-month ODI, and 24-month ODI. The number of missing values counted 85 CRP, 14 DS, 555 6-month ODI, 422 12-month ODI, and 731 24-month ODI.

Data Handling
The simple imputation technique of the last observation carried forward (LOCF) was used to include patients with incomplete 24-month ODI, with the last reported scores being used in place of the missing values. The months of follow-up were then weighed on the imputations, thus obtaining the ODI scores at 17 months (661 24-month ODI + 576 12-month ODI + 155 6-month ODI). In Table 1 is reported the structure of the dataset. Subsequently, three new variables were imputed. The first variable was the difference between the baseline ODI and the last ODI at 17 months (ΔODI PRE vs. last 17). The second variable was the presence or absence (1,0) of clinical improvement at 17 months (17-month progress 30%), calculated as a 30% decrease from baseline ODI to the last ODI at 17 months. The third variable was the presence or absence (1,0) of clinical improvement at 17 months (17-month progress 12.7), calculated as a raw reduction in the ODI score≥12.7 from baseline ODI to the last ODI at 17 months. Two new sets classified the sample by gender code (males = 0, females = 1) and diagnosis (spine deformities: 4, disc diseases: 5; back surgeries: 6). Specifically, the group of spine deformities included kyphosis and scoliosis, and the group of back surgeries included spondylosis, spondylolisthesis, stenosis, and elective treatment of fractures. The dataset was explored for what concerned the presence of outliers among the primary variables, and a new dataset for regression analysis was created after the elimination of outliers. The following number of outliers were excluded based on the interquartile range rule of three: 4 RBCs, 27 MCV, 36 MCH, 2 MCHC, 96 CRP, 2 DS, and 21 LOS.

Statistics
The IBM SPSS 22 statistics package was used for all statistical analysis. The descriptive variables were reported as means, standard deviation, minimum and maximum values to be reported in Tables 2, 3 (baseline examination), regardless of the type of distribution. In Table 4 (outcome exploration) it was planned to report the most significant biochemical descriptors
Concerning outcome exploration, the Chi-Square test was used to investigate the differences in gender between outcome groups at 17 months using both 30% reduction and the raw 12.7 points reduction. Similarly, it was planned the Mann-Whitney U-test for the years of age and the biochemical parameters, taking into account the result of Levene’s test for equality of variances based on median (adjusted degree of freedom). The Chi-Square test was also used to investigate the different outcomes according to ASA, diagnosis, DS, and LOS. Subsequently, the probe of two prediction regression models was planned; the first model (PRE\textsubscript{Bio}) would be based on the seven preoperative biochemical markers, while the second model (PRE\textsubscript{BioDemClin}) would have included the biochemical, the demographic (gender and age), and the clinical (diagnosis, ASA, baseline ODI, DS, and LOS) variables. The prediction potential of each of the two models on the last ODI was explored through multiple regression analysis on the dataset without outliers. The models were planned to be based on the stepwise method to serially add the next strongest predictor feasibly removing the previously entered predictor not significant. The most significant predictors would be chosen to report the predictive equations. The Wilcoxon signed-rank test was run to match the regression predicted values from the estimated regression equation with the real values of the dataset. The main predictors of each model were tested for the clinically significant outcomes (1,0) by using binary logistic regression (enter method). Neural Networks (NN) analysis was planned to observe the forecasting outcome as a function of each variable-specific model: PRE\textsubscript{Bio}(NN) and PRE\textsubscript{BIODEMCLIN}(NN). Given the non-linear nature of this tool, the authentications between inputs and outputs have been run on the raw dataset through the supervised learning technique of Multilayer Perceptron (MLP) procedure to produce a predictive model for clinical success based on the values of the demographic, biochemical, and clinical predictors. After data whitening (the distributions were rescaled so that the mean was zero and the standard deviation was one), the training sample was set at 70%, the testing sample to track prediction at 20%, and the holdout sample to assess the final

### TABLE 1 | Structure and missing data of the study group dataset.

| Captions              | Cases n | Missing n (%) |
|-----------------------|---------|---------------|
| Age (years, continuous) | 1,392 | 0             |
| Gender (dichotomous)   | 1,392 | 0             |
| RBCs ($10^{12}$/µl, continuous) | 1,392 | 0             |
| Hematocrit (%) (continuous) | 1,392 | 0             |
| Hemoglobin (g/dl, continuous) | 1,392 | 0             |
| MCV (fl/cell, continuous) | 1,392 | 0             |
| MCH (pg/cell, continuous) | 1,392 | 0             |
| MCHC (g/dl, continuous) | 1,392 | 0             |
| CRP (mg/dl, continuous) | 1,307 | 1,807 (7.77%) |
| Diagnosis (categorical) | 1,392 | 0             |
| ASA (ordinal)          | 1,392 | 0             |
| DS (minutes, continuous) | 1,378 | 14 (1.01%)    |
| LOS (days, continuous)  | 1,392 | 0             |
| Baseline ODI (continuous) | 1,392 | 0             |
| 6-month ODI (continuous) | 837 | 555 (69.87%) |
| 12-month ODI (continuous) | 970 | 422 (30.32%) |
| 24-month ODI (continuous) | 661 | 731 (82.51%) |
| LOCF 17-month ODI (continuous) | 1,392 | 0             |
| TOTAL                  | 1,807 | (7.77%)       |

RBCs, red blood cells; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; CRP, C-reactive protein; ASA, American Society of Anesthesiologists physical status classification system (1, healthy; 2, mild, 3, severe; 4, life threatening; 5, moribund; 6, brain-dead); DS, duration of surgery; LOS, length of hospital stay; ODI, Oswestry Disability Index, ranging from 0 (no disability) to 100 (maximum disability); LOCF, last observation carried forward.

### TABLE 2 | Baseline demographic and biochemical characteristics of the study group.

| Descriptors      | Values                     | Cases n |
|------------------|----------------------------|---------|
| Age (years)      | 56.41 ± 15.24 (18.00; 88.00) | 1,392   |
| Gender           | 567 males, 825 females      | 1,392   |
| **Biochemistry** |                            |         |
| RBCs ($10^{12}$/µl) | 04.76 ± 0.99 (02.50; 06.88) | 1,392   |
| Hematocrit (%)   | 42.31 ± 0.64 (27.20; 54.00) | 1,392   |
| Hemoglobin (g/dl) | 14.04 ± 0.38 (09.20; 18.90) | 1,392   |
| MCV (fl/cell)    | 89.16 ± 0.69 (54.60; 115.20) | 1,392   |
| MCH (pg/cell)    | 29.58 ± 0.20 (18.00; 37.60) | 1,392   |
| MCHC (g/dl)      | 33.16 ± 0.10 (27.50; 36.50) | 1,392   |
| CRP (mg/dl)      | 0.00 ± 0.67 (00.01; 08.74)  | 1,307   |

RBCs, red blood cells; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; CRP, C-reactive protein.
NN at 10%. The NN architecture was planned to be based on two hidden layers with Sigmoid activation function (real-valued arguments are transformed to the range 0, 1) and on Softmax activation for the output layer (the vector of real-valued arguments is transformed to a vector whose elements fall in the range 0, 1 and sum to 1). The Receiver Operating Characteristic (ROC) curve evaluated the Area Under the Curve (AUC), the sensitivity (true positive outcome), specificity (true negative outcome), and 1–specificity (false positive rate) of model-specific predictors in three successive run of NN, with the normalized importance of the independent predictors independent being reported for the most significant multivariate model.

**RESULTS**

**Preoperative Examination**

The demographic and preoperative biochemical variables are reported in Table 2. The study cohort showed a prevalence of about 41% of males and 59% of females, with men showing higher values than females of RBCs (05.02 ± 00.46 10⁶/µl vs. 04.58 ± 00.42 10⁶/µl; U = 104,835.00, Z = −17.389, p < 000001), Ht (44.53 ± 03.17% vs. 40.79 ± 03.12%; equal variances $t_{1390} = 21.833, p < 000001$), Hb (14.93 ± 01.20 g/dl vs. 13.43 ± 01.14 g/dl; equal variances $t_{1390} = 23.722, p < 000001$), MCH (29.84 ± 02.15 pg/cell vs. 29.40 ± 02.21 pg/cell; $U = 182,475.00, Z = −5.633, p < 000001$), and MCHC (33.53 ± 00.99 g/dl vs. 32.91 ± 00.96 g/dl; equal variances $t_{1388} = 11.909, p < 000001$). No sex differences for MCV (88.98 ± 05.67 fl/cell in males vs. 89.28 ± 05.71 fl/cell in females; $U = 219,466.00, Z = −0.742, p = 0.458$) and CRP (00.34 ± 00.70 mg/dl in males vs. 00.35 ± 00.65 mg/dl in females; $U = 173,124.00, Z = −0.665, p = 0.506$).

Males had a mean age of 55.15 ± 15.63 years old and females had 57.28 ± 14.92 years old. A negative association was found between age and RBCs ($Rho = −0.184, p < 000001$), Ht ($Rho = −0.116, p = 000001$), Hb ($Rho = −0.164, p < 000001$), and MCHC ($Rho = −0.207, p < 000001$). Conversely, age positively associated with MCV ($Rho = +0.203, p < 000001$), MCH ($Rho = +0.068, p = 0.012$), and CRP ($Rho = +0.198, p < 000001$). The variables concerning surgical and anesthetic parameters are reported in Table 3, indicating data for each of the three clusters of spine diagnosis.

The association of sex, the years of age, and the biochemical markers with the ODI score at baseline are reported in Figure 1. There was found a positive association between the preoperative disability with the years of age ($B = +0.197, p < 000001, 95% CI = +1.142; +0.253$) and CRP ($B = +10.985, p < 000001, 95% CI = +6.162; +15.807$). An inverse association was observed for RBCs ($B = −4.117, p < 000001, 95% CI = −5.931; −2.302$), Ht ($B = −0.579, p < 000001, 95% CI = −0.814; −0.343$), Hb ($B = −1.736, p < 000001, 95% CI = −2.355; −1.117$), and
MCHC ($B = -1.829, p = 0.00003, 95\% CI = -2.686: -0.972$). No association were found with circulating MCV ($B = 0.121, p = 0.219, 95\% CI = -0.072: +0.314$) or MCH ($B = -0.176, p = 0.511, 95\% CI = -0.701: +0.349$). Men showed lower values of disability against women ($42.42 \pm 16.22\%$ vs. $49.88 \pm 15.86\%; U = 172,804.50, Z = -8.294, p < 0.00001$).

### Exploration of the Outcome

In the whole cohort, 1,022 patients recovered at least 30% from baseline ODI at 17 months, while the reduction of 12.7 points counted 1,019 successful outcomes. The recovery from disability considering 30% reduction showed an association with gender [$\chi^2(1) = 5.339, p = 0.021$], with about 76.72% of males and 71.15% of females encountering a recovery. However, the gender association was not confirmed [$\chi^2(1) = 0.017, p = 0.895$] after considering the reduction of 12.7 points. Patients with a unsuccessful recovery of at least 30% reduction were older at baseline ($U = 156,017.50, Z = -4.990, p < 0.00001$) with lower values of RBCs ($U = 165,980.50, Z = -3.385, p = 0.001$), Ht ($U = 170331.50, Z = -2.828, p = 0.005$), Hb ($U = 163,994.50, Z = -3.786, p = 0.0001$), MCHC ($U = 158,815.00, Z = -4.521, p < 0.00001$), but higher values of MCV ($U = 164,810.00, Z = -2.603, p = 0.009$). No differences for baseline MCH ($U = 178,818.50, Z = -0.051, p = 0.960$) and CRP ($U = 130,854.50, Z = -1.711, p = 0.087$). Considering the reduction of 12.7 points, the results were confirmed for age ($p = 0.002$), MCH ($p = 0.920$), MCHC ($p = 0.026$), and CRP ($p = 0.540$), but no differences were found for RBCs ($p = 0.098$), Ht ($p = 0.122$), Hb ($p = 0.059$), and MCV ($p = 0.147$). There was found a significant association between the 30% reduction and ASA [successful outcome in 81.07% with ASA 1, 72.55% with ASA 2, and 61.23% with ASA 3; $\chi^2(2) = 36.381, p = 0.00002$], diagnosis [successful
outcome in 61.11% of spine deformities, 80.57% of disc disease, and 70.07% of back surgeries; \( \chi^2 = 36.381, p < 0.00001 \), DS (209.83 ± 137.15 vs. 248.13 ± 152.25 min; \( U = 156,271.50, Z = -4.339, p = 0.00001 \)), and LOS (4.58 ± 3.35 vs. 5.53 ± 4.17 days; \( U = 155,719.00, Z = -4.077, p = 0.00004 \)). Considering the reduction of 12.7 points, the results were confirmed for ASA (\( p = 0.016 \)), diagnosis (\( p = 0.00001 \)), DS (\( p = 0.002 \)), and LOS (\( p = 0.005 \)), with baseline ODI resulting different between groups (\( p < 0.00001 \)).

### Predictive Model \( \text{PRE}_{\text{BIO}} \)

The values of RBCs, Ht, Hb, MCV, MCH, MCHC, and CRP were entered in the regression model to predict the variation of ODI at 17 months after surgery. The predictive model (\( N = 1,179 \)) with the highest correlation was RBCs (\( \text{standardized } B = -0.071, p = 0.015, 95\% \text{ CI} = -0.587: -0.622 \)). The model reported below significantly predicted the ODI variation [\( F(1,1177) = 5.895, p = 0.015, R^2 = 0.005 \)].

\[
\text{PRE}_{\text{BIO}} \text{(17 months)} = -9.121 - 3.239 \left( \frac{\text{RBCs} \times 10^6}{\mu l} \right)
\]

The estimated values calculated using the regression equation (\(-24.63 \pm 1.53\); \( \min = -30.75; \max = -18.48 \)) showed no difference from real values (\(-24.37 \pm 20.81; \min = -96.00; \max = +58.00 \)) (\( Z = -0.079, p = 0.937 \)). Using the 17-month reduction of 30% from baseline ODI score, increasing RBCs was associated with an increased likelihood of exhibiting a positive clinical outcome (\( B = -0.431, p = 0.001 \)). The OR for having a positive outcome resulted in 1.539 (95\% CI = 1.188: 1.992). The logistic regression model with RBCs correctly classified 73.34\% of cases and was able to explain 1.14\% (\( \text{Nagelkerke } R^2 \)) of the variance in the clinical outcome, with very low misspecification in its predictive capacity [Hosmer-Lemeshow test \( \chi^2(8) = 13.207, p = 0.105 \)]. The RBCs-based NN analysis using the progress of 30\% showed 34.75, 20.00, 24.48\% of incorrect predictions of the holdout phase. The models resulted to be a poor diagnostic instrument (first run AUC = 0.553; second run AUC = 0.565; third run AUC = 0.559). Using the reduction of 12.7 from baseline, increasing RBCs was not significantly associated with an increased likelihood of exhibiting a positive clinical outcome (\( B = 0.185, p = 0.152 \)). The OR for having a positive outcome resulted in 1.203 (95\% CI = 0.934: 1.550). The logistic regression model correctly classified 73.19\% of cases and was able to explain 0.2\% of the variance in the clinical outcome. The NN analyses using the variation of −12.7 points showed 27.70, 29.14, and 24.27\% of incorrect predictions of the holdout phase. In conclusion, the RBCs-based \( \text{PRE}_{\text{BIO}} \)(NN) resulted to be a poor diagnostic instrument (first run AUC = 0.527; second run AUC = 0.533; third run AUC = 0.528).

### Predictive Model \( \text{PERI}_{\text{BIODEMCLI}} \)

The preoperative biomarkers were used to build the model at 17 months together with the two demographic variables of age and gender and the five clinical variables of diagnosis, ASA, baseline ODI, DS, and LOS. The model that resulted statistically significant [\( F(5,1365) = 93.487, p < 0.00001, R^2 = 0.255 \)] included the following predictor-specific coefficients: Hb (\( \text{standardized } B = -0.067, p = 0.006, 95\% \text{ CI} = -1.710: -0.281 \)), age (\( \text{standardized } B = +0.100, p = 0.0003, 95\% \text{ CI} = +0.063: +0.207 \)), ASA (\( \text{standardized } B = +0.063, p = 0.024, 95\% \text{ CI} = +0.300: +4.230 \)), baseline ODI score (\( \text{standardized } B = -0.500, p < 0.00001, 95\% \text{ CI} = -0.688: -0.570 \)), and LOS (\( \text{standardized } B = +0.129, p < 0.00001, 95\% \text{ CI} = +0.698: +1.498 \)). Gender (\( p = 0.088 \)), RBCs (\( p = 0.698 \)), Ht (\( p = 0.100 \)), MCV (\( p = 0.839 \)), MCH (\( p = 0.992 \)), MCHC (\( p = 0.140 \)), CRP (\( p = 0.083 \)), diagnosis (\( p = 0.820 \)), and DS (\( p = 0.265 \)) were excepted.

\[
\text{PERI}_{\text{BIODEMCLI}} \text{(17 months)} = +2.021 - \left[ 0.995 \times \left( \frac{\text{Hb} \times 10^6}{\mu l} \right) \right] + \left[ 0.135 \times \text{ (years of age)} \right] + \left[ 2.265 \times \text{(ASA)} \right] - \left[ 0.629 \times \text{ (preoperative ODI)} \right] + \left[ 1.098 \times \text{ (days of hospital stay)} \right]
\]

However, the extreme distributions of the regression predicted values (\(-24.65 \pm 10.42\%\); \( \min = -64.55; \max = +62.62 \)) differed from real values (\(-24.37 \pm 20.81\%\); \( \min = -96.00; \max = +58.00 \)). Using the reduction of 30\% from baseline ODI score, a greater likelihood of exhibiting positive outcome was again associated with increasing Hb (\( B = 0.100, p = 0.035, OR = 1.106, 95\% \text{ CI} = 1.007: 1.213 \)), decreasing age (\( B = -0.016, p = 0.001, OR = 0.984, 95\% \text{ CI} = 0.975: 0.994 \)), and shorter LOS (\( B = -0.093, p = 0.0003, OR = 0.911, 95\% \text{ CI} = 0.866: 0.958 \)). However, no association was found with the preoperative ODI score (\( B = 0.008, p = 0.051, OR = 1.008, 95\% \text{ CI} = 1.000: 1.013 \)) and ASA (\( B = -0.196, p = 0.136, OR = 0.822, 95\% \text{ CI} = 0.636: 1.064 \)). The logistic regression model correctly classified 73.67\% of cases and was able to explain 5.39\% (\( \text{Nagelkerke } R^2 \)) of the variance in the clinical outcome, with very low misspecification in its predictive capacity [Hosmer-Lemeshow test \( \chi^2(8) = 3.896, p = 0.866 \)]. In Table 5 are the percentages of clinical success based on ranges of RBCs and Hb. There were expected 96.5\% of true positives and 92.7\% of false positives after setting a value of RBCs to 4 \( 10^6/\mu l \). Similarly, setting a value of Hb to 12 g/dl exhibited about 93.8\% of the positive outcomes correctly classified as positive, but 88.9\% of the negative outcomes incorrectly specified as positive.

Considering RBCs, Hb, the two demographic variables, and the five clinical variables, the \( \text{PERI}_{\text{BIODEMCLI}} \)(NN) with 30\% reduction showed 29.46, 22.63, 25.71\% of incorrect predictions in the holdout phase. The model resulted to be a fair diagnostic instrument (first run AUC = 0.628; second run AUC = 0.632; third run AUC = 0.626), with normalized importance of 100.00\% given by age in the first and third run and LOS in the second run. Using the 17-month progress of −12.7 from baseline ODI score, the odds were confirmed for age (\( p = 0.002 \)), ASA (\( p = 0.154 \)), and LOS (\( p = 0.001 \)), but Hb was no more a significant predictor (\( p = 0.088 \)) whereas baseline ODI significantly predicted the outcome (\( p < 0.00001 \)). The logistic regression model correctly classified 74.91\% of cases and was able to explain 10.81\% of the variance in the clinical outcome. The \( \text{PERI}_{\text{BIODEMCLI}} \)(NN) with 12.7 point reduction showed 34.69, 30.22, 29.17\% of incorrect...
predictions in the holdout phase. The model resulted to be a fair diagnostic instrument (first run AUC = 0.720; second run AUC = 0.723), with normalized importance of 100.00% given by baseline ODI.

**DISCUSSION**

It is foreseeable that the workload of spine surgery centers will intensify in the coming years as the population is getting older and debilitating polymorbid conditions are becoming not uncommon (19, 20). Technological advances in spine treatments help maintain short-term patient satisfaction high (21). However, it is necessary to plan patient-centered care pathways to achieve long-term results, thus revising the determinants of clinical success that simultaneously capture the perspective of the surgeon, the anesthesiologist, and the patient. In the present study, we analyzed the predictive potential of the preoperative biochemical markers on the ODI score at 17 months after surgery in patients enrolled in the institutional SpineReg of IRCCS Orthopedic Institute Galeazzi. The study cohort involved 1,392 patients undergoing surgery for deformity, disc disease, or other back spine disorders, and consisted of a majority of female older adults (Table 2). In absolute terms, an improvement in disability at the last follow-up was observed in over 88% of patients. Considering the more restrictive MCID of the ODI, about 73% of patients reported a successful recovery. Similar rates have already been observed in spine patients (5). There were no differences in recovery between males and females, but individuals who did not experience a clinical improvement were older at the time of surgery, had higher ASA, and lower ODI. Equally, these trends based on clinical determinants are in line with previous studies (22). Patients with spinal deformities experienced lower recovery rates than the other clusters of diagnosis, conceivably due to the greater surgical complexity that requires longer operative times and prolonged hospitalization (Table 3). Analyses of laboratory values confirmed that males generally have higher levels of RBCs and Hb than females and that there is a significant depletion with increasing age, feasibly mirroring iron supply discrepancies common in older adults. Similarly, the positive association of MCV and MCH with age would suggest an etiology from cobalamin or folate deficiency, which are known to cause macrocytic anemia in older individuals with poor strength of constitution (23, 24). This consideration was corroborated by increased disability and inflammation found in older patients (Figure 1). Based on available laboratory parameters, predictive modeling demonstrated that RBCs and Hb levels prior to surgery were the strongest determinants of clinical success at 17 months in all types of spine surgery. In particular, the univariate linear model explains 0.5 of the postoperative change in disability, with each unit increase in RBCs being associated up to 1.539 times the probability of clinical success. However, the corresponding neural network models showed poor diagnostic performance, having an erroneous prediction rate of up to 34.75% and an AUC of 0.565, making it unreliable in terms of sensitivity and specificity. Furthermore, only slight reductions in disability scores (−17 to −31 for RBCs) could be predicted. The prediction accuracy for poor outcomes did not improve after setting lower blood values, failing to identify both highly successful outcomes and worsening observed in 140 patients at 17 months. Therefore, it can be reasonably argued that stratification of patients based on univariate cutoffs may not be recommended and that studying laboratory biomarkers as continuous variables might be preferable (25, 26). In fact, the results in Table 5 showing a comparable trend between blood parameters and success rates give both RBCs and Hb a strong connotation that is also relevant for the patient. With the inclusion of clinical parameters, the variables in the multivariate linear models were able to explain only slight reductions in disability scores (−17 to −31 for RBCs) could be predicted. The prediction accuracy for poor outcomes did not improve after setting lower blood values, failing to identify both highly successful outcomes and worsening observed in 140 patients at 17 months. Therefore, it can be reasonably argued that stratification of patients based on univariate cutoffs may not be recommended and that studying laboratory biomarkers as continuous variables might be preferable (25, 26). In fact, the results in Table 5 showing a comparable trend between blood parameters and success rates give both RBCs and Hb a strong connotation that is also relevant for the patient. With the inclusion of clinical parameters, the variables in the multivariate linear models were able to explain only slight reductions in disability scores (−17 to −31 for RBCs) could be predicted. The prediction accuracy for poor outcomes did not improve after setting lower blood values, failing to identify both highly successful outcomes and worsening observed in 140 patients at 17 months. Therefore, it can be reasonably argued that stratification of patients based on univariate cutoffs may not be recommended and that studying laboratory biomarkers as continuous variables might be preferable (25, 26). In fact, the results in Table 5 showing a comparable trend between blood parameters and success rates give both RBCs and Hb a strong connotation that is also relevant for the patient. With the inclusion of clinical parameters, the variables in the multivariate linear models were able to explain only slight reductions in disability scores (−17 to −31 for RBCs) could be predicted. The prediction accuracy for poor outcomes did not improve after setting lower blood values, failing to identify both highly successful outcomes and worsening observed in 140 patients at 17 months. Therefore, it can be reasonably argued that stratification of patients based on univariate cutoffs may not be recommended and that studying laboratory biomarkers as continuous variables might be preferable (25, 26). In fact, the results in Table 5 showing a comparable trend between blood parameters and success rates give both RBCs and Hb a strong connotation that is also relevant for the patient. With the inclusion of clinical parameters, the variables in the multivariate linear models were able to explain only slight reductions in disability scores (−17 to −31 for RBCs) could be predicted. The prediction accuracy for poor outcomes did not improve after setting lower blood values, failing to identify both highly successful outcomes and worsening observed in 140 patients at 17 months. Therefore, it can be reasonably argued that stratification of patients based on univariate cutoffs may not be recommended and that studying laboratory biomarkers as continuous variables might be preferable (25, 26).
FIGURE 2 | Multiplayer perceptron network predicting 17-month clinical recovery from disability using baseline biochemical, demographic, and clinical parameters. The Multiplayer Perceptron Network (MPN) shows the prediction of 17-month recovery set as a raw decrease of 12.7 points from baseline ODI. (A) The architecture has two hidden layers, with the Sigmoid and Softmax being the hidden and output functions, respectively (criteria training = batch; optimization = scaled conjugate). (B) The network ROC curve (Receiver Operating Characteristic) shows a fair diagnostic performance of the parameters in predicting the last ODI score at 17 months after spine surgery. (C) The independent variables were rescaled (standardized) to show their normalized contribution in the model: gender (8.25%), diagnosis (10.20%), ASA (11.18%), age (34.72%), RBCs (41.95%), Hb (49.22%), DS (49.67%), LOS (70.20%), baseline ODI (100.00%). (D) The classification table of predicted vs. observed values was set with a training sample at 70%, a testing sample to track prediction at 20%, and a holdout sample to assess the final model at 10%. The final percent correct shows a high performance in predicting the clinically successful outcomes, but poor reliability to forecast negative outcomes. ODI, Oswestry Disability Index, ranging from 0 (no disability) to 100 (maximum disability); Hb, hemoglobin; RBCs, red blood cell count; ASA, American Society of Anesthesiologists physical status classification system (1, healthy; 2, mild; 3, severe; 4, life threatening; 5, moribund; 6, brain-dead); DS, duration of surgery; LOS, length of hospital stay.

model at the last follow-ups showed the highest diagnostic performance even for the more restrictive MCID (AUC between 0.720 and 0.726 with up to 15.15% of correct prediction of negative outcomes), providing a decreasing order of importance of the preoperative determinants: ODI, LOS, DS, Hb, RBCs, age, ASA, diagnosis, gender (Figure 2). Thus, Hb seems to have a high predictive potential even greater than variables of demographic or clinical nature. The importance of preoperative Hb has already been studied in relation to complications in children, adults, and older adults undergoing spinal surgery (27–29), but it is unclear how it affects patients’ long-term daily activities. It is plausible to think that the blood concentration reflects not only the strength of the patient’s constitution (e.g., nutritional status) (8, 9, 30), but also the disease-specific weaknesses whose complications might consequently affect the daily activities of the patients (31). For example, there was found an inverse association between Hb levels and the number of patients reporting fatigue and shortness of breath (32). Whatever the connection, it is undeniable the recognition of the predictive potential that Hb has in the many surgical fields (33, 34). This study has
limitations. Although patients admitted for complications were excluded from this research, information on intraoperative (e.g., transfusion-associated complications) or postoperative events that did not require access in our hospital was not accessible, thus possibly explaining the inability of the models to predict worsening. Furthermore, the study cohort might not represent the population of patients undergoing spinal surgery in our hospital, being indicative only of those who have agreed to participate in the registry over the years. While the completeness of the registry was satisfactory, missing data at predefined follow-ups reached 40% and may have provided some bias to the results. However, the models at 17 months were built on the scores at the last controls, thus making the results consistent (35). Lastly, although they can be estimated on the basis of surgical plan, both operative times and days of hospitalization are information available only after the intervention, which could undermine the preoperative nature of models.

In conclusion, our study sheds light on the role of preoperative Hb and RBCs in predicting long-term recovery reported by patients. Based on this research, values of RBCs < 4 x 10^6/µl and of Hb < 12 g/dl in both genders may be associated with excessive rates of long-term failure after spine surgery from a patient’s perspective. However, the model is not reliable in its current form and should be integrated with multidimensional variables of demographic, laboratory, and clinical nature to investigate further recovery determinants, such as body weight (36), the psychological distress (37), or the propensity for postoperative movement (38) and social participation (39). The ideal predictive model should have both high sensitivity and low false-positive rates. This is especially relevant when the consequence of not identifying patients at risk for negative outcomes could affect long-term daily activities. The performance of predictive models also varies according to the extent of recovery considered clinically relevant (17, 18), a concept that places the need to involve the patient in planning the treatment path, thus making the pre-admission visit more patient-centered. In the future, the correct stratification of individuals at risk will ensure opportunities to optimize patient’s health in time for surgery, more affordable clinical care, and greater patient’s satisfaction (16, 40, 41).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

MB, PP, FL, TC, and ED conceived and designed the research. MB, FL, TC, and ED collected the data and managed the database. MB analyzed the data and wrote the first draft of the manuscript. PP, FL, TC, ED, PR, MP, LS, RB, MB-B, GB, and PB revised the first draft and contributed to the manuscript sections. GB and PB supervised the study. All authors contributed to the manuscript revision and approved the submitted version.

ACKNOWLEDGMENTS

This study was part of the project Ricerca Corrente del Ministero della Salute (Italian Ministry of Health).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsurg.2022.850342/full#supplementary-material

REFERENCES

1. Parenteau CS, Lau EC, Campbell IC, Courtney A. Prevalence of spine degeneration diagnosis by type, age, gender, and obesity using Medicare data. Sci Rep. (2021) 11:5389. doi: 10.1038/s41598-021-84724-6
2. Zanotto BS, Etges A, Marcolino MAZ, Polanczyk CA. Value-based healthcare initiatives in practice: a systematic review. J Healthc Manag. (2021) 66:340–65. doi: 10.1097/JHM-D-20-00283
3. Finkelstein JA, Schwartz CE. Patient-reported outcomes in spine surgery: past, current, future directions. J Neurosurg Spine. (2019) 31:155–64. doi:10.3171/2019.1.SPINE18770
4. Fairbank JC, Pynsent PB. The Oswestry disability index. Spine (Phila Pa 1976). (2000) 25:2940–52. discussion 2952. doi:10.1097/00007632-200011150-00017
5. Nielsen CJ, Lewis SJ, Otminc C, Martin AR, Lenke LG, Qiu Y, et al. Stratifying outcome based on the Oswestry Disability Index for operative treatment of adult spinal deformity on patients 60 years of age or older: a multicenter, multi-continenal study on Prospective Evaluation of Elderly Deformity Surgery (PEEDS). Spine J. (2021) 21:1775–83. doi:10.1016/j.spinee.2021.07.007
6. Prodinger B, Taylor P. Improving quality of care through patient-reported outcome measures (PROMs): expert interviews using the NHS PROMs Programme and the Swedish quality registers for knee and hip arthroplasty as examples. BMC Health Serv Res. (2018) 18:87. doi: 10.1186/s12913-018-2898-2
7. Briguglio M, Gianola S, Aguirre MF, Sirtori P, Perazzo P, Pennestri F, et al. Nutritional support for enhanced recovery programs in orthopedics: Future perspectives for implementing clinical practice. Nutr Clin Metab. (2019) 33:190–8. doi:10.1016/j.nupar.2019.04.002
8. Xu B, Xu WX, Lao YJ, Ding WG, Lu D, Sheng HF. Multimodal nutritional management in primary lumbar spine surgery: a randomized controlled trial. Spine (Phila Pa 1976). (2019) 44:967–74. doi:10.1097/BRS.0000000000002992
9. Sachdeva I, Carmouche JJ. Postoperative anemia predicts length of stay for geriatric patients undergoing minimally invasive lumbar spine fusion surgery. Geriatr Orthop Surg Rehabil. (2020) 11:2151459320911874. doi:10.1177/2151459320911874
10. Suresh KV, Wang K, Sethi I, Zhang B, Margalit A, Puvanesarajah V, et al. Spine surgery and preoperative hemoglobin, hematocrit, and
hemoglobin A1c: a systematic review. *Global Spine J.* (2022) 12:155–65. doi: 10.1177/2192568220979821.

11. Briguglio M, Crespi T, Pino F, Mazzochi M, Porta M, De Vecchi E, et al. Clinical characteristics of severe COVID-19 patients admitted to an intensive care unit in Lombardy during the Italian pandemic. *Front Med.* (2021) 8:582896. doi: 10.3389/fmed.2021.582896.

12. Briguglio M, Hrelias S, Malaguti M, De Vecchi E, Lombardi G, Ranfi G, et al. Oral supplementation with sucrosomial ferric pyrophosphate plus L-ascorbic acid toameliorate the marital status: a randomized controlled trial. *Nutrients.* (2020) 12:386. doi: 10.3390/nu12020386.

13. Barrie U, Youssef CA, Pernik MN, Adeyemo E, Elguindy M, Johnson ZD, et al. Transfusion guidelines in adult spine surgery: a systematic review critical summary of currently available evidence. *Spine J.* (2022) 22:238–48. doi: 10.1016/j.spinee.2021.07.018.

14. Mehta VA, Van Belleghem F, Price M, Jaykel M, Ramirez L, Goodwin J, et al. Hematocrit as a predictor of preoperative transfusion-associated complications in spine surgery: a NSQIP study. *Clin Neurol Neurosurg.* (2021) 200:106322. doi: 10.1016/j.clineuro.2020.106322.

15. Murphy MF, Goodnough LT. The scientific basis for patient blood management. *Transfus Clin Biol.* (2015) 22:90–96. doi: 10.1016/j.tracl.2015.04.001.

16. Debono B, Wainwright TW, Wang MY, Sigmundsson FG, & Debono B, Wainwright TW, Wang MY, Sigmundsson FG, et al. Impact of obesity on outcomes following lumbar disc herniation surgery: a systematic review and meta-analysis. *Clin Neurol Neurosurg.* (2021) 145:1–14. doi: 10.1016/j.clineuro.2021.112128.

17. Vanti C, Ferrari S, Villafane JH, Berjano P, Monticone M. Responsiveness and minimum important change of the Oswestry Disability Index in Italian subjects with symptomatic lumbar spondylolisthesis. *J Orthop Traumatol.* (2017) 18:145–50. doi: 10.1007/s10195-017-0446-y.

18. Monticone M, Biaardi P, Vanti C, Ferrari S, Pillastri P, Mugnai R, et al. Responsiveness of the Oswestry Disability Index and the Roland Morris Disability Questionnaire in Italian subjects with sub-acute and chronic low back pain. *Eur Spine J.* (2012) 21:122–9. doi: 10.1007/s00592-011-1959-3.

19. Sethi R, Bohl M, Vitale S. State-of-the-art reviews: safety in complex spine surgery. *Spine Deform.* (2019) 7:657–68. doi: 10.1016/j.jspd.2019.04.002.

20. Briguglio M. The burdens of orthopedic patients and the value of the HPES approach (healthy eating, physical activity, sleep hygiene). *Front Med.* (2021) 8:650947. doi: 10.3389/fmed.2021.650947.

21. Mooney J, Michalopoulos GD, Alvi MA, Geary E, Elguindy M, Johnson ZD, et al. Transfusion guidelines in adult spine surgery: a systematic review critical summary of currently available evidence. *Spine J.* (2022) 22:238–48. doi: 10.1016/j.spinee.2021.07.018.

22. Aghayev E, Mannion AF, Fekete TF, Janssen S, Goodwin J, et al. Impact of obesity on outcomes following lumbar disc herniation surgery: a systematic review and meta-analysis. *Clin Neurol Neurosurg.* (2021) 145:1–14. doi: 10.1016/j.clineuro.2021.112128.

23. Goyal A, Elminawy M, Kerezoudis P, Alvi MA, et al. Impact of obesity on outcomes following lumbar spine surgery: a systematic review and meta-analysis. *Clin Neurol Neurosurg.* (2019) 177:27–36. doi: 10.1016/j.clineuro.2018.12.012.

24. Goyal A, Elminawy M, Kerezoudis P, Alvi MA, et al. Impact of obesity on outcomes following lumbar spine surgery: a systematic review and meta-analysis. *Clin Neurol Neurosurg.* (2019) 177:27–36. doi: 10.1016/j.clineuro.2018.12.012.

25. Sivaganesan A, Khan I, Pennings JS, Roth SG, Nolan ER, Oleisky ER, et al. Correction of hypovitaminosis D improved global longitudinal strain earlier than left ventricular ejection fraction in cardiovascular elderly adults after orthopaedic surgery. *J Geriatr Cardiol.* (2018) 15:519–22. doi: 10.11097/jisst.1671-5411.2018.08005.
41. Briguglio M, Wainwright T. Nutritional and physical prehabilitation in elective orthopedic surgery: rationale and proposal for implementation. Ther Clin Risk Manag. (2022) 18:21–30. doi: 10.2147/TCRM.S341953

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Briguglio, Perazzo, Langella, Crespi, De Vecchi, Riso, Porrini, Scaramuzzo, Bassani, Brayda-Bruno, Banfi and Berjano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.