

Article

Machine Learning-Based Risk Prediction of Critical Care Unit Admission for Advanced Stage High Grade Serous Ovarian Cancer Patients Undergoing Cytoreductive Surgery: The Leeds-Natal Score

Alexandros Laios ^{1,*}, Raissa Vanessa De Oliveira Silva ^{2,†}, Daniel Lucas Dantas De Freitas ², Yong Sheng Tan ¹, Gwendolyn Saalmink ¹, Albina Zubayraeva ¹, Racheal Johnson ¹, Angelika Kaufmann ¹, Mohammed Otify ¹, Richard Hutson ¹, Amudha Thangavelu ¹, Tim Broadhead ¹, David Nugent ¹, Georgios Theophilou ¹, Kassio Michell Gomes de Lima ² and Diederick De Jong ¹

- ¹ Department of Gynaecologic Oncology, St James's University Hospital, Leeds LS9 7TF, UK; yong.tan@nhs.net (Y.S.T.); gwendolyn.saalmink@nhs.net (G.S.); albinazubayraeva@gmail.com (A.Z.); racheal.johnson2@nhs.net (R.J.); angelika.kaufmann@lthtr.nhs.uk (A.K.); motify@nhs.net (M.O.); richard.hutson@nhs.net (R.H.); amudha.thangavelu@nhs.net (A.T.); tim.broadhead@nhs.net (T.B.); david.nugent@nhs.net (D.N.); georgios.theophilou@nhs.net (G.T.); diederick.dejong@nhs.net (D.D.J.)
- ² Department of Chemistry, Federal University of Rio Grande do Norte, Natal CEP 59078-970, Brazil; raissavanessa781@hotmail.com (R.V.D.O.S.); danielucas77@hotmail.com (D.L.D.D.F.); kassiolima@gmail.com (K.M.G.d.L.)
- * Correspondence: a.laios@nhs.net
- † These authors contributed equally to this work.



Citation: Laios, A.; De Oliveira Silva, R.V.; Dantas De Freitas, D.L.; Tan, Y.S.; Saalmink, G.; Zubayraeva, A.; Johnson, R.; Kaufmann, A.; Otify, M.; Hutson, R.; et al. Machine Learning-Based Risk Prediction of Critical Care Unit Admission for Advanced Stage High Grade Serous Ovarian Cancer Patients Undergoing Cytoreductive Surgery: The Leeds-Natal Score. *J. Clin. Med.* **2022**, *11*, 87. <https://doi.org/10.3390/jcm11010087>

Academic Editor: Satoru Kyo

Received: 29 October 2021

Accepted: 22 December 2021

Published: 24 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Achieving complete surgical cytoreduction in advanced stage high grade serous ovarian cancer (HGSOC) patients warrants an availability of Critical Care Unit (CCU) beds. Machine Learning (ML) could be helpful in monitoring CCU admissions to improve standards of care. We aimed to improve the accuracy of predicting CCU admission in HGSOC patients by ML algorithms and developed an ML-based predictive score. A cohort of 291 advanced stage HGSOC patients with fully curated data was selected. Several linear and non-linear distances, and quadratic discriminant ML methods, were employed to derive prediction information for CCU admission. When all the variables were included in the model, the prediction accuracies were higher for linear discriminant (0.90) and quadratic discriminant (0.93) methods compared with conventional logistic regression (0.84). Feature selection identified pre-treatment albumin, surgical complexity score, estimated blood loss, operative time, and bowel resection with stoma as the most significant prediction features. The real-time prediction accuracy of the Graphical User Interface CCU calculator reached 95%. Limited, potentially modifiable, mostly intra-operative factors contributing to CCU admission were identified and suggest areas for targeted interventions. The accurate quantification of CCU admission patterns is critical information when counseling patients about peri-operative risks related to their cytoreductive surgery.

Keywords: Machine Learning; surgical cytoreduction; Critical Care Unit; ovarian cancer; Graphical User Interface

1. Introduction

Cancer of the fallopian tube, ovary, or peritoneum (EOC) is the leading cause of death from gynaecological malignancy in the western world [1]. Over 70% of women diagnosed with EOC have advanced stage disease at presentation (International Federation Obstetrics and Gynaecology FIGO stage 3–4) [1]. The gynaecological oncology community is now recognizing high grade serous ovarian cancer (HGSOC), as yet the most prevalent, as a single clinical entity. For such an aggressive subtype, identification of novel, ideal biomarkers that strongly correlates with the stage and may be effective for early diagnosis, remains

critical [2]. The cornerstone of treatment includes a combination of surgical cytoreduction and platinum-based chemotherapy aiming at elimination of all visible disease, ideally reaching a total macroscopic tumor clearance. When the cancer is at an advanced stage, the debulking surgery often requires prolonged surgical times and possible multi-visceral resections, necessitating Critical Care Unit (CCU) support, and prolonged hospitalization [3]. Risk-prediction models of severe post-operative complications after ovarian cytoreductive surgery have been thoroughly proposed [4].

Routine post-operative CCU admission has been considered a standard of care for such patients at risk of increased peri-operative morbidity and mortality, and post-operative complications [5]. Patient-related pre-operative risk factors associated with significant peri-operative outcomes have been comprehensively studied [6]. There is currently a growing interest in monitoring CCU admissions, and performance has been compared across different trusts to improve the standards of care. In the realm of precision medicine, innovative data mining technologies, such as Machine Learning (ML), can be clearly used for monitoring quality improvement and effectiveness of delivery of modern ovarian cancer care. The ML approach for outcome prediction based on fully curated clinical data has been applied in a variety of hospital settings [7]. We previously developed ML algorithms to improve the prediction accuracy of complete cytoreduction in advanced stage HGSOc patients [8]. In addition, we highlighted the importance of feature selection for the ML-based two-year prognosis estimation in the same population [9]. The usefulness of ML as a prognostic tool in the ovarian cancer environment has been increasingly demonstrated [10].

Accurate prediction is difficult with conventional statistics, because patient characteristics show a multidimensional and non-linear relationship. Machine Learning methods could be helpful in monitoring CCU admissions to improve standards of care. In an era of stringent economics, CCU admission is often arbitrary, driven by local practice or bed availability. It remains heavily at the surgeon's discretion, as post-operatively, the gynaecologic oncologist is mainly responsible for daily monitoring and decision making with regards to the patient's recovery [5]. International comparison studies have suggested that patterns of CCU admission after surgery are inconsistent, and the value of such admissions routinely may be nuanced [6]. While focusing on the cost-effectiveness of CCU utilization after major surgeries, timely identification of those patients, who will benefit from CCU support, remains a major challenge [11]. Lately, the introduction of maximal surgical effort to achieve complete cytoreduction without compromising peri-operative management and subsequent recovery requires the availability of CCU beds [12]. This paradigm shift has prompted the development of tools to accurately predict CCU admission following cytoreductive surgery. To inform consent and enhance decision making, we developed and validated an ML-based risk score to predict CCU admission of advanced stage HGSOc patients undergoing surgical cytoreduction. We hypothesized that CCU admission of HGSOc patients is multifactorial and could be accurately predicted by using ML algorithms. The primary outcome was the prediction accuracy of several ML prediction methods, based on a set of performance metrics, and the performance comparison with the conventional logistic regression (LR) analysis. The secondary objective was factor refinement for CCU admission and exploration of the associations between CCU admission, severe post-operative complications, hospital stay, and 30-day hospital re-admission.

2. Materials and Methods

Prospectively registered data in the hospital-wide Patient Pathway Manager (PPM) database from 291 consecutive HGSOc women, who underwent cytoreductive surgery at St James's University Hospital, Leeds by a certified Gynaecologic Oncology Surgeon from January 2014 to December 2019 was analyzed. This database was developed internally for clinical trials and integrated with an electronic patient record system [13]. Our hospital is a tertiary centre, recently accredited by the European Society of Gynaecologic Oncology (ESGO) as a centre of excellence for ovarian cancer surgery. The staging was reported by the 2014 International Federation of Gynaecology and Obstetrics classification [14].

Women underwent either primary debulking surgery (PDS) or 3–4 cycles of neoadjuvant chemotherapy (NACT) followed by interval debulking surgery (IDS) if they had: FIGO stage 4 disease; poor performance status (PS); uncertainty about the possibility of optimal tumor removal. Only HGSOc patients with at least one pre-treatment CA125 measurement were included in the study. Patients aged < 18 years, as well as those with progressive disease, synchronous primary malignancy or secondary cytoreduction, were excluded. Post-operative CCU admission was considered for high-risk HGSOc patients, who were scheduled to undergo complex major surgery, including multi-visceral resections or undertook pre-operative cardiopulmonary exercise (CPEX) fitness testing, if they were deemed high risk at pre-assessment [15].

2.1. Candidate Variables for Prediction Model

Candidate predictors were selected a priori from two domains:

- Patient: age, year of diagnosis, body mass index (BMI), pre-operative co-morbidities (Charlson Co-morbidity Index (CCI)), Eastern Co-operative Oncology Group (ECOG) performance status (PS), and pre-operative albumin.
- Operative factors: timing of surgery (PDS or IDS), operative time (OT), surgical complexity score (SCS), bowel resection with stoma formation (yes/no), residual disease (RD), and estimated blood loss (EBL).

All these variables are widely available at tertiary centres. Based on clinical experience, they have been previously shown to be independent predictors of post-operative morbidity and mortality for ovarian cancer patients [4]. Severe post-operative complications were reported according to the Clavien-Dindo morbidity score (grade 3–5) [16]. Only several surgical site infections (SSIs) were recorded. The CCI was categorized as: 0, 1, 2 with higher scores indicating greater co-morbidity [17]. The SCS was assigned based on the Aletti classification as low, intermediate, and high [18]. Outcomes also included rate of CCU admission, length of hospital stays, and hospital re-admission, defined as admission to any hospital for longer than 24 h within 30 days of discharge.

Descriptive statistics were displayed by frequency and percentages for binary and categorical variables and by means and standard deviation (SD) or medians (with lower or upper quartiles for continuous variables). The Chi-square test was performed for categorical variables. The Fischer's exact tests were used for binary variables. All analyses were performed with IBM® SPSS® 26.0 Statistics Software for Windows (IBM, New York, NY, USA).

2.2. Feature Reduction and Selection

Our objective was to use all the available data to maximize the generalizability of our results. Feature selection from the candidate predictors aimed to optimize the predictor group by finding the smallest independent set of variables with the greatest association to the outcome in question, in order to ensure best performance and to minimize over-fitting. We used standard feature selection strategies, including dropping-O variance and highly correlated variables. All clinical variables were evaluated individually, one by one, with the vector of CCU classification responses. Forward selection and backward stepwise regression were then employed to select independent variables for inclusion in the final models. Variables with a threshold p -value < 0.05 and 95% Confidence Intervals (CI) were considered to be statistically significant and were included in the predictive model. To re-value two variables as more relevant, and to optimize the latent collinearity and avert variable over-fitting, a correlation analysis was evaluated by assessing the determination coefficient (R^2) and the root mean square error of cross validation (RMSECV). A poor correlation between variables was observed if the R^2 was low or the RMSECV was elevated. Variables with a threshold of >0.6 were eliminated to determine the final input variables for the model. Highly inter-correlated variables were grouped, and variables with no mutual optimization were dropped. The resulting rankings were used to select the set of features that led to the highest prediction accuracy. Subsequently, the optimal number of important

features resulting in the highest prediction accuracy was identified. For feature importance, we employed resources for assigning weight to system variables. That is, for any chosen ML method, we assigned different weights to different variables.

2.3. Model Development and Training

The model development followed a standard data science approach. The raw clinical data were imported into MATLAB environment (MathWorks Inc, MA, USA) for pre-processing and subsequent multivariate analysis. The data set was randomly split into training and test cohorts (60:40% ratio) with repeated random sampling (bootstrap), until there was no significant difference ($p = 0.20$) between the two cohorts with respect to all variables. The training set was used for model construction. A specific type of cross-validation was employed—the CV LOO, in which each sample from the training set was evaluated using the model built by all the other samples in the same training set. This was used to check and define important parameters. The remaining samples assigned to the test set were used to evaluate the model performance. For the Artificial Neural Network (ANN) models in the regression mode, all the data were used, and sample selection differed by using the Kennard-Stine method with 60% for training and 40% for testing. Subjects with missing values were omitted. Following the pre-processing stage, all the quantitative variables were normalized. Categorical variables were transformed into binary dummy variables. Next, different subsets of data were labeled to solve the classification prediction problem, namely CCU admission, as positive or negative. This was then formulated as a binary classification problem. For all methods and to avoid over-fitting or selection bias, a leave-one-out cross-validation (VB-CV) was applied to test the model's ability to predict new data by computing a statistic on the left-out sample. Only subjects with fully curated data were eligible for the prediction analysis.

2.4. Multivariate Analysis

Initially, the model was evaluated variable by variable in a univariate fashion, but there was no trend observed. Various state-of-the-art supervised classifiers, suitable for the type and size of the data set, were tested in a multivariate fashion. The algorithms were chosen because of their relative simplicity for interpretation and overall popularity within the data science community and through our own testing of several ML algorithms during model development, focusing on different types of interactions [19]. Algorithms included the non-linear distance models using neural networks (K-Nearest Neighbors (KNN), ANN in Backpropagation Neural Network (BPN) mode, and the models using linear and quadratic discriminant methods (LDA and QDA)). A popular ML model is the ANN; roughly modeled around the way human brain works, it is useful for supervised prediction tasks [20]. Its variation, the BPN, is a type of neural network training that fine-tunes the weights of a neural network based on the error rate obtained from the previous iteration. Proper tuning of the weights allows for reduction of error rates and makes the model reliable, thus increasing its generalizability. This method helps to calculate the gradient of a loss of function with respect to all the weights in the network. A basic prediction technique, such as KNN is employed to verify the complexity of the prediction problem. Linear logistic regression (LR) was used as a benchmark. The conceptually simple approach of LDA and its sibling, QDA, remain among the most effective procedures in the domain of high-dimensional prediction [21,22].

2.5. Model Performance

To fully evaluate the discriminatory effectiveness of the model, we calculated several performance metrics including accuracy, sensitivity, specificity, F-scores, and the area under the receiver operating characteristic curve (AUC) for the test set. The accuracy represents the total number of samples correctly classified considering true and false negatives. We examined the model performance including all the variables as well as the simplest model with the highest and more robust performance (less features). The sensitivity represented

the proportion of positive samples correctly classified. The specificity represented the proportion of negative samples correctly classified and finally, the F-score measured the model performance considering imbalanced data. A Receiver Operating Curve (ROC) was generated as a single metric that is not dependent on the discrimination threshold.

2.6. Development of the CCU Risk Calculator

For clinical utility, and to generate real-time predictions, numerous CCU prediction tests at different model probability output thresholds were performed in the validation set, and checked in the derivation set. Because most of the data were classified into one group or another with no intervals, we developed a similar method, using the ANN(BPN) model in calibration mode to turn discrete values into continuous values. A total of five features were selected as threshold, given their significance, and offering the best compromise of computing time required and possible information loss. A certain range of values, minimum and maximum were set for each feature, as below:

Surgical Complexity Score: 2–18;
Albumin levels: 17–49;
Estimated Blood Loss: 100–4000;
Operative Time: 65–445;
Bowel Resection: 0–1;

A software tool (Calibration Tools Hall, UFRN, Natal, Brazil, upon request) by using the ANN(BPN) regression algorithm was developed based on five selected predictor features. The algorithm is a method of fine-tuning the weights of a neural network based on the error rate obtained in the preceding iteration. Proper weight tuning reduces the error rates rendering the model reliable, and increasing its generalizability. Entering values above or below made the Graphical User Interface reset to the closest minimum/maximum value in one or all input variables. All values were numerically entered, and the 'Estimate Risk' button was pressed to obtain the estimate in three different formats that complemented each other as a result. The combination of these five features would yield 30 multivariate sets of predictors to evaluate in three different modeling thresholds. These thresholds were influenced by the data distribution and clinical importance of various cut-offs; the entire data set, merging both training and test sets, was then used to report the CCU admission probability percentages. The development of the predictive models and the report of the findings were used in accordance with the EQUATOR Guideline for reporting ML predictive models [23], in support of the STROBE statement for reporting of observational studies [24].

3. Results

A general architecture of our proposed methodology is showing in Figure 1. A cohort of 291 patients was eligible for the study. The mean age and BMI of the patients were 64 ± 10.5 yrs, respectively. The mean SCS was 3 ± 2 (2–11). Of these patients, 69/291 (23.7%) and 222/291 (76.3%) underwent PDS and IDS, respectively. Complete and optimal cytoreduction was achieved in 190/291 (65.3%) and 78/291 (26.8%) patients. The details of baseline characteristics and operative factors are displayed in Table 1 to enable the better understanding of this group of patients undergoing cytoreductive surgery and potentially requiring CCU admission. There were 56/291 (19.2%) admissions; of those, 49 admissions were elective whilst seven were unplanned admissions. From the variables selected to predict CCU admission, the following variables were significantly different: pre-treatment albumin, SCS, EBL, OT, and performance of bowel resection with stoma formation (Figure 2). The correlation structure of pre-processed features was demonstrated by a correlation heat map, as well as the most significant variables in Figure 3.

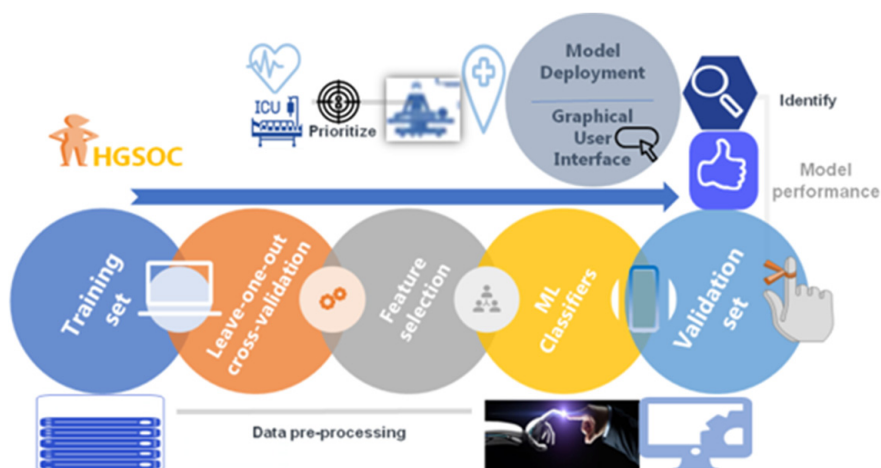


Figure 1. Machine Learning (ML) Model derivation and validation flow from a fully curated source of advanced stage high grade serous ovarian cancer patient and surgical data. The framework comprised data pre-processing with feature selection and model training. The model was evaluated using performance metrics prior to the development of an ML-based nomogram and a web application to help with risk assessment.

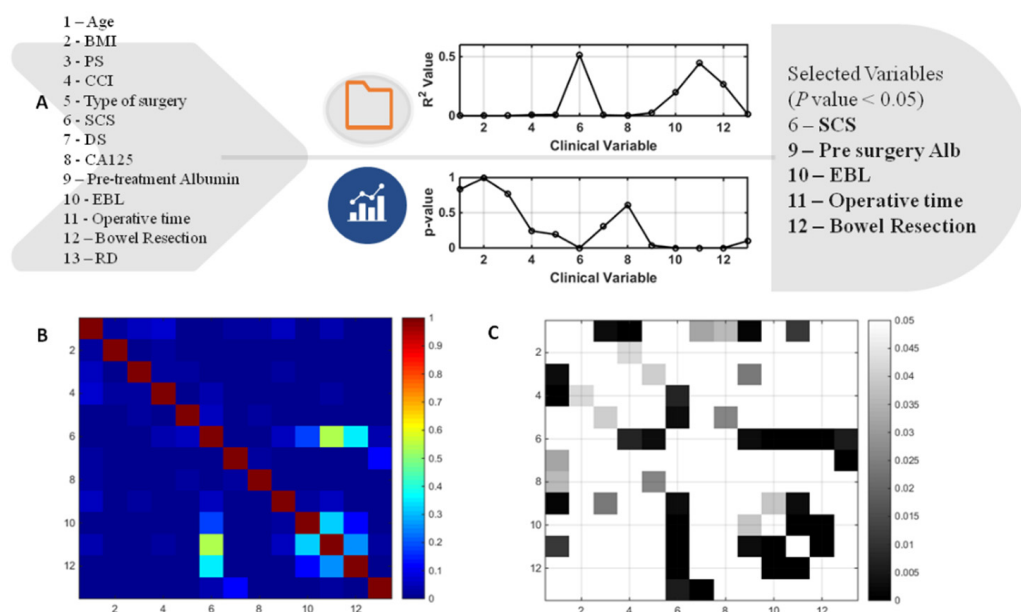


Figure 2. (A) Results of multivariate logistic regression leading to feature selection in the development of the predictive Machine Learning model. (B) Correlation heatmap representing the values of the determination co-efficient (R^2) between the original variables entered for model development. Not surprisingly, the highest correlation was observed between surgical complexity and operative time. (C) Correlation heatmap representing the P values between the original system variables. In areas where there was no representation (square in any key), there was no significant correlation amongst variables ($p > 0.05$). BMI; body mass index, PS; performance status, CCI; Charlson co-morbidity index, SCS; surgical complexity score, DS; disease score, EBL; estimated blood loss, RD; residual disease, Alb; albumin.

Table 1. Pre-operative, intra-operative and post-operative parameters of the study cohort.

Variables (<i>n</i> = 291)	Levels	Frequency	Percentage
Age, year, SD (range)	N/A	64.2 ± 10.5 (41–90)	N/A
BMI, mean, SD (range)	N/A	27 ± 5.8 (18.5–58)	N/A
Pre-Treatment CA125, mean, SD (min-max)	N/A	1777 ± 3125 (12–28,000) U	N/A
Pre-Treatment Albumin, mean, SD (min-max)	N/A	38.3 ± 3.8 (17–49) U	N/A
Surgical Complexity Score (SCS)	Low (1–3)	166	57%
	Moderate (4–7)	108	37.1%
	High (8–12)	17	5.9%
Disease score (DS)	Pelvis	17	5.8%
	Lower abdomen	220	75.6%
	Upper abdomen	54	18.6%
Residual Disease (RD)	R0	190	65.3%
	R1 (<1 cm)	78	26.8%
	R2 (≥1 cm)	23	7.9%
CCU admission	Yes	56	19.2%
	No	235	80.8%
Bowel resection with stoma	Yes	21	7.2%
	No	270	92.8%
ECOG Performance Status (PS)	0	127	43.6%
	1	119	40.9%
	2	38	13.1%
	3	7	2.4%
Charlson Co-morbidity Index (CCI)	Low (0–2)	146	50.2%
	High (≥3)	145	49.8%
Timing of cytoreduction	PDS	69	23.7%
	IDS	222	76.3%
Operation time, mean SD (min-max)	N/A	182 ± 75 (65–480) min	N/A
Length of stay, mean, SD (min-max)	N/A	9 ± 13 (3–172) days	N/A
Clavien-Dindo complications (3a–5)	Yes	16	5.5%
Admission within 30 days	Yes	Data	Data

ECOG; Eastern Co-operative Oncology Group; PDS; primary debulking surgery; IDS; interval debulking surgery.

One sample from the training set was omitted due to missing values, totaling 290 patients with complete data, aiming to predict CCU admission. When all the variables were employed, the training accuracy from the derivation set ranged from 0.93 to 0.97. The corresponding internally validated (test) predictive accuracies for the KNN, ANN, LDA, and QDA algorithms were 0.80, 0.82, 0.90 and 0.93, respectively. All but one F-score were convincingly >0.80 (Table 2). Following feature reduction, we observed a slightly worse, yet satisfactory performance with the mean accuracy across classifiers between 0.76 and 0.89 for the CCU admission prediction (Table 2). In that same table, the predictive comparison between ML algorithms and LR suggested a significant outperformance for LDA and KNN over LR. To guide model choice, a visual summary of all the operational algorithms performed can be seen in Figure 3 for both the training and testing partitions. Model performance remained robust between the training and test sets. When all the

variables were included, the QDA model performed best in the training set (AUC: 0.98) (Figure 3A). With feature selection, the smallest, highest-performing, and most robust model was the LDA model in the training set (AUC: 0.97) (Figure 3C), yet it performed best in the test set (AUC: 0.91) (Figure 3D).

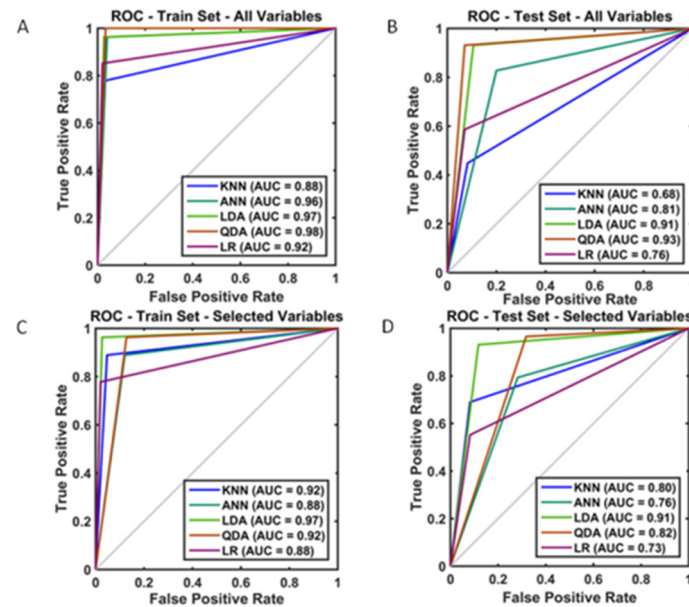


Figure 3. Visual summary of the model performance and evaluation based on the comparison of Receiver Operator Curve (ROC) and Area Under (ROC) (AUROC) for Critical Care Unit admission (A) training set; all variables included (B) test set; all variables included (C) training set; selected variables (D) test set; selected variables included.

3.1. Model Deployment

To assess the clinical utility of our model, a user-friendly Graphical User Interface of the CCU admission (the Leeds-Natal) risk score software macro was developed to guide the observers through the analysis. The interface is available upon request. In brief, the interface provided probability results and visualization of the risk position for CCU admission for clinical implementation and easy interpretation by the physicians (Figure 4A). The model performance using the ANN(BPN) regression principle was akin to the ML algorithms used (Figure 4B). We describe the implementation of our method in two HGSOC patients, who underwent cytoreductive surgery (Figure 4C). Patient 1 underwent a standard total abdominal hysterectomy, bilateral salpingo-oophorectomy, and omentectomy following primary chemotherapy. This was a difficult procedure, completed in three hours with an EBL of 1lt. Her probability of getting admitted to CCU was only 1.49%, and she was considered low risk. Patient 2 underwent a standard primary surgical cytoreduction, in addition to lymph node dissection and pelvic peritoneal stripping. The EBL was 600 mL, and the procedure was completed in 220 min. Her risk for CCU admission was 32.56% and she was still considered low risk. If a rectosigmoid resection was anticipated intra-operatively in that same patient, resulting in stoma formation, her SCS would increase from 5 to 8, and her risk for CCU admission would rise to 84.27%.

3.2. Association between CCU Admission and Length of Stay, Post-Operative Complications, and Hospital Re-Admission

Admission to CCU was associated with increased length of stay ($p = 0.000$; OR 8.0, CI 7.7–8.35) and reduced risk for post-operative CD (3a-5) complications, ($p = 0.001$; OR 0.16, CI 0.057–0.45) but no risk for re-admission within 30 days ($p = 0.63$).

Table 2. Performance metrics of the Machine Learning models and comparisons with conventional Logistic Regression for the prediction of Critical Care Unit admission following cytoreductive surgery for advanced high grade serous ovarian cancer.

Predictors	Model	Set	Accuracy	Sensitivity	Specificity	F-Score
All variables (n = 13)	KNN (K = 4)	Train	0.94	0.78	0.97	0.86
		CV LOO	0.94	0.78	0.97	0.86
		Test	0.80	0.45	0.92	0.60
	ANN	Train	0.97	0.96	0.97	0.96
		CV LOO Test	0.88 0.82	0.85 0.86	0.88 0.81	0.86 0.83
LDA	Train Test	0.97 0.90	0.96 0.93	0.97 0.89	0.96 0.91	
QDA	Train Test	0.97 0.93	1.00 0.93	0.97 0.93	0.98 0.93	
LR	Train Test	0.96 0.84	0.85 0.59	0.98 0.93	0.91 0.72	
Selected * Variables (p < 0.05)	KNN (K = 6)	Train	0.94	0.89	0.95	0.92
		CV LOO	0.94	0.89	0.95	0.92
		Test	0.86	0.69	0.92	0.79
	ANN	Train	0.90	0.89	0.90	0.99
		CV LOO Test	0.89 0.76	0.89 0.79	0.89 0.74	0.89 0.76
LDA	Train Test	0.97 0.89	0.96 0.93	0.97 0.88	0.96 0.90	
QDA	Train Test	0.89 0.75	0.96 0.97	0.87 0.68	0.91 0.80	
LR	Train Test	0.95 0.82	0.78 0.55	0.98 0.92	0.87 0.69	

* Surgical complexity score; pre-surgery albumin; blood loss; operative time; bowel resection with stoma. KNN; k-Nearest Neighbors, ANN; Artificial Neural Network, CV-LOO; Leave-one-out-cross-validation; LDA; Linear Discriminant Analysis, QDA; Quadratic Discriminant Analysis, LR; Logistic Regression.

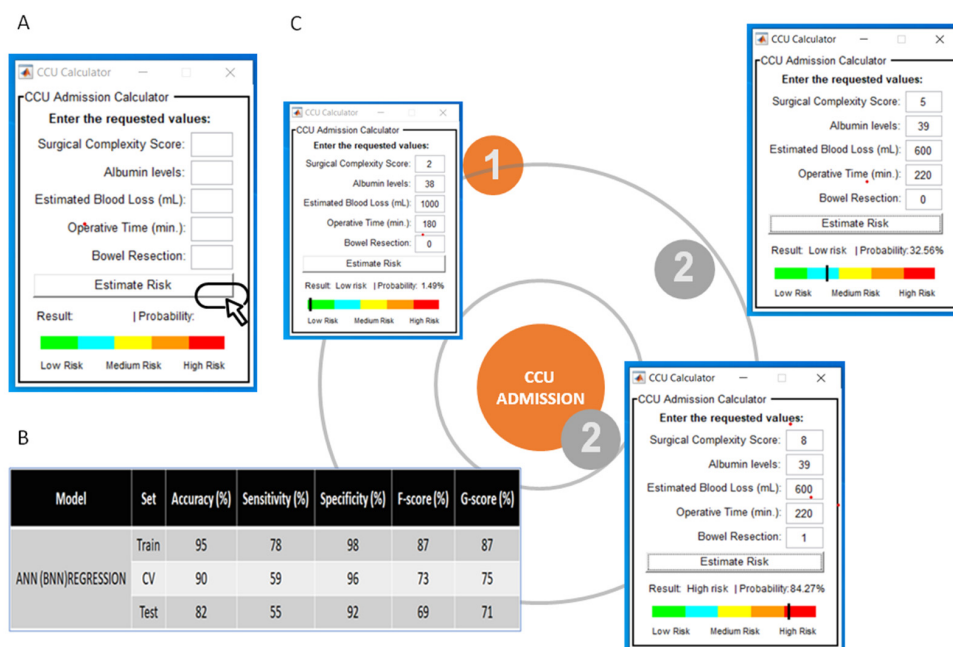


Figure 4. (A) Screenshot from the Critical Care Unit (CCU) calculator Graphical User Interface for the prediction of CCU admission. (B) Performance of the Artificial Neural Network (ANN(BPN)) algorithm used for the development of the Graphical User Interface based on the regression principles. (C) Examples of clinical application. Note that in the same patient, intra-operative decision making for rectosigmoid resection and stoma formation would massively increase the risk for CCU admission.

4. Discussion

In HGSOE, the recent paradigm shift towards achieving complete surgical cytoreduction has certain ramifications to prevent an anticipated increased peri-operative morbidity and mortality [25]. For such a heterogeneous disease, the accuracy of any related outcome can be limited, prompting the development of new tools to improve prediction. In this cohort study, we developed and fine-tuned several ML prediction algorithms that accurately predicted CCU admission. Equally, we determined the importance of non-inherent, hence potentially modifiable, prediction features. To our knowledge, this is the first study to quantify patterns of CCU admission following cytoreductive surgery for HGSOE by using ML methods.

Our study allowed for the comparison of several learning algorithms to identify the approach with the most favorable performance. We demonstrated the utility of linear discriminants (LDA, QDA) of clinical characteristics in model performance with maximum accuracy reaching 98% and the lowest boundary of performance being no worse than 68%. The overall model was rendered stable, owing to the quality of the fully curated training data set. Although the study was not powered to the outcome prediction, the data sample was large enough to capture any potential variance. Collinearity was successfully addressed and reduced by feature selection, when required. In addition, we selected supervised ML algorithms such as LDA, QDA, and KNN, which outperformed conventional LR with respect to prediction accuracy indices, which is potentially more clinically relevant than AUC. Such a comparison modeling system has already been shown to test intra-institutional performances [24]. We are familiar with the performance of the KNN algorithm, which is very much reflective of “previous clinical experience”, for the accurate prediction of R0 resection during HGSOE surgery [8]. Where the data classes were unbalanced, the tested methods performed similarly to LR.

Confirming findings from prior empirical research, in addition to the premises of this present work, our study identified those features with the maximal relevance to the CCU admission prediction accuracy. All models placed importance on variables with known predictive implications. All but one was intra-operative parameters. We acknowledge that for planning purposes, it would have been more useful to predict CCU admission based on more pre-operative variables; this was not possible due to the feature selection process. Nevertheless, similar performance metrics were observed when all the variables were included in the model. This highlights the value of the relative importance of each predictor in developing the model. The significant variables may be differentially weighted in the model analysis to reflect estimates of their contribution to relevant similarity. Some discriminants have been previously shown to be predictors of short-term outcomes in EOC patients [13]. Concurrent bowel resection with stoma formation, as part of a frequently extensive cytoreductive surgery, was a critical predictor of CCU admission. This agrees with findings from other studies [26]. It can be explained by the increased risk of severe post-operative complications in patients who undergo rectosigmoid surgery during pelvic exenteration [27]. Advanced stage HGSOE is commonly associated with weight loss, severe malnutrition, and impaired gastrointestinal function, hence low pre-treatment albumin, resulting in a protracted surgical recovery [28]. Primary chemotherapy could potentially impact some of these factors, as shown to decrease intra-operative blood loss, earlier resumption to ambulation, and return of intestinal function [29,30]. The effect of NACT on post-operative admission to CCU remains controversial. Although primary chemotherapy is associated with lower peri-operative morbidity and post-surgical mortality [31], this was not correlated to CCU admission. New systemic treatment algorithms for newly diagnosed advanced EOC could serve as tools to help with clinical decision making and patient stratification on a daily basis [32].

Accurate identification of HGSOE patients at a quantified risk for CCU admission following their surgery is important in gynaecologic oncology. Our findings demonstrated that ML algorithms are able to predict a patient’s risk for CCU admission with excellent discrimination. The main strength of the model was versatility, allowing for case mix when

auditing the CCU admission as a proxy indicator of the quality of care. As specificity was high for most models, this would make it attractive in a cancer screening program; if diagnostic prediction is the goal, whereas both sensitivity (true positives) and specificity (true negatives) are important, our highly performed linear discriminant algorithms (LDA and QDA) would primarily apply. Nevertheless, such a tool could be useful in aiding a clinician's risk assessments in personalized management plans for patients. The development of the first to our knowledge CCU risk calculator—the Leeds-Natal score—provided a fine balance between high model accuracy and model simplicity that makes it easily interpretable to the gynaecologic oncology surgeon. Based on five significant features, it can evaluate which HGSOC subpopulation has the best prediction or which subpopulation is most difficult to predict. In addition, it gives off certain advantages, such as the percentage of those belonging to the “closest to admission” group, which is a ‘true idea’ of the real probability of fitting in the group of women who should be admitted to CCU (Figure 4).

The rate of CCU admission was <20%. The length of CCU stays has been always topical as extensive fluid resuscitation during surgery, poor nutritional status, and older age are associated with a prolonged CCU stay [33]. A breakdown of planned and unplanned admissions reflecting towards objective criteria with respect to valid peri-operative goal-directed haemodynamic therapy protocols is an important topic, and research is currently under way. Herein, although the CCU admission rate increased transiently from 2014 to 2019, the rate of unplanned admissions remained steady (data not shown). In HGSOC, unplanned admissions are associated with reportedly poor survival outcomes [34]. In high-output tertiary centres, interdisciplinary meetings to allocate and prioritize CCU beds prior to the surgical cases may have translated into the observed increased planned and decreased unplanned post-operative CCU admissions.

Prolonged hospitalizations, CCU stays, and hospital re-admissions, have all come under scrutiny due to increasing emphasis on improving palliative care and quality of life for patients near their end of life [35]. Admission to CCU contributes to increased costs for both patients and hospitals [36]. In our post-hoc analysis, CCU admission has been independently associated with increased hospitalization, fewer post-operative complications, but no effect on the risk of re-admission within 30 days, in contrast to other studies [34]. If extensive surgery is associated with morbidity-related infections and wound complications, it may be prudent to provide nutrition support ensuring the peri-operative period in these women [37]. In that same report, CCU admission increased the odds of severe surgical infections, compared to routine ward admission. For such an inherently high-risk population, this information is critical when counseling women about peri-operative risks ahead of their surgery, and potential adverse impacts in their overall survival.

This was a single-centre study with a limited rate of heterogeneity in the study population, which may differ from other tertiary unit settings. The retrospective nature of the study was the main limitation, despite a proprietary, prospectively maintained database. A very slight discordance of model performance between the training and test sets may not be generalized to other data sources. Nevertheless, ML retains the strength of the structural model used for the prediction, even when applied to other populations with different prediction features. Although model robustness was strongly supported by internal validation, further external validation in an independent cohort may be required to confirm the ability of the models as guidance tools.

5. Conclusions

In high-output centres, CCU admission is a measurable outcome that can be used as a benchmark of surgical care. We developed ML models that predicted CCU admission following cytoreductive surgery for HGSOC with an accuracy as high as 98%. Equally, we refined risk-adjusted predictors of CCU admission using pertinent pre-operative and intra-operative variables. Most importantly, we demonstrated clinical feasibility by the development of a Graphical User-friendly Interface to quantify patterns of CCU admission.

Predictive ML algorithms may facilitate quality improvement of modern care by improving the CCU admission prediction accuracy.

Author Contributions: Conceptualization, A.L.; methodology, A.L., R.V.D.O.S. and D.L.D.D.F.; software, A.L., R.V.D.O.S., D.L.D.D.F. and K.M.G.d.L.; validation, R.V.D.O.S. and D.L.D.D.F.; formal analysis, A.L. and A.Z.; investigation, A.L. and M.O.; data curation, A.L., Y.S.T., A.Z. and A.K.; writing—original draft preparation, A.L.; writing—review and editing, A.L., R.V.D.O.S., D.L.D.D.F., T.B., G.S., M.O., R.H., A.T., R.J. and G.T.; supervision, D.N. and D.D.J.; project administration, G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Leeds Teaching Hospitals Ethics Committee of St James’s Institute of Oncology (MO20/133163/18.06.20).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We would like to thank all the members of the Gynaecologic Oncology team at St James’s University Hospital, Leeds who managed the patients enrolled in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer Statistics. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [[CrossRef](#)]
2. Giusti, I.; D’Ascenzo, S.; Dolo, V. Microvesicles as Potential Ovarian Cancer Biomarkers. *BioMed Res. Int.* **2013**, *2013*, 703048. [[CrossRef](#)] [[PubMed](#)]
3. Pelosi, P.; Ball, L.; Schultz, M.J. How to optimize critical care resources in surgical patients: Intensive care without physical borders. *Curr. Opin. Crit. Care* **2018**, *24*, 581–587. [[CrossRef](#)]
4. Aletti, G.; Santillan, A.; Eisenhauer, E.L.; Hu, J.; Aletti, G.; Podratz, K.C.; Bristow, R.E.; Chi, D.S.; Cliby, W.A. A new frontier for quality of care in gynecologic oncology surgery: Multi-institutional assessment of short-term outcomes for ovarian cancer using a risk-adjusted model. *Gynecol. Oncol.* **2007**, *107*, 99–106. [[CrossRef](#)] [[PubMed](#)]
5. Davidovic-Grigoraki, M.; Thomakos, N.; Haidopoulos, D.; Vlahos, G.; Rodolakis, A. Do critical care units play a role in the management of gynaecological oncology patients? The contribution of gynaecologic oncologist in running critical care units. *Eur. J. Cancer Care* **2017**, *26*, e12438. [[CrossRef](#)]
6. Pearse, R.M.; Moreno, R.P.; Bauer, P.; Pelosi, P.; Metnitz, P.; Spies, C.; Vallet, B.; Vincent, J.-L.; Hoeft, A.; Rhodes, A. Mortality after surgery in Europe: A 7 day cohort study. *Lancet* **2012**, *380*, 1059–1065. [[CrossRef](#)]
7. Rowan, M.; Ryan, T.; Hegarty, F.; O’Hare, N. The use of artificial neural networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. *Artif. Intell. Med.* **2007**, *40*, 211–221. [[CrossRef](#)] [[PubMed](#)]
8. Laios, A.; Gryparis, A.; DeJong, D.; Hutson, R.; Theophilou, G.; Leach, C. Predicting complete cytoreduction for advanced ovarian cancer patients using nearest-neighbor models. *J. Ovarian Res.* **2020**, *13*, 117. [[CrossRef](#)] [[PubMed](#)]
9. Laios, A.; Katsenou, A.; Tan, Y.S.; Johnson, R.; Otify, M.; Kaufmann, A.; Munot, S.; Thangavelu, A.; Hutson, R.; Broadhead, T.; et al. Feature Selection is Critical for 2-Year Prognosis in Advanced Stage High Grade Serous Ovarian Cancer by Using Machine Learning. *Cancer Control.* **2021**, *28*, 10732748211044678. [[CrossRef](#)]
10. Enshaei, A.; Robson, C.N.; Edmondson, R.J. Artificial Intelligence Systems as Prognostic and Predictive Tools in Ovarian Cancer. *Ann. Surg. Oncol.* **2015**, *22*, 3970–3975. [[CrossRef](#)]
11. Ghaffar, S.; Pearse, R.M.; Gillies, M.A. ICU admission after surgery: Who benefits? *Curr. Opin. Crit. Care* **2017**, *23*, 424–429. [[CrossRef](#)] [[PubMed](#)]
12. Fotopoulou, C.; Concin, N.; Planchamp, F.; Morice, P.; Vergote, I.; Du Bois, A.; Querleu, D. Quality indicators for advanced ovarian cancer surgery from the European Society of Gynaecological Oncology (ESGO): 2020 update. *Int. J. Gynecol. Cancer* **2020**, *30*, 436–440. [[CrossRef](#)] [[PubMed](#)]
13. Newsham, A.C.; Johnston, C.; Hall, G.; Leahy, M.G.; Smith, A.B.; Vikram, A.; Donnelly, A.M.; Velikova, G.; Selby, P.J.; Fisher, S.E. Development of an advanced database for clinical trials integrated with an electronic patient record system. *Comput. Biol. Med.* **2011**, *41*, 575–586. [[CrossRef](#)] [[PubMed](#)]
14. Mutch, D.G.; Prat, J. 2014 FIGO staging for ovarian, fallopian tube and peritoneal cancer. *Gynecol. Oncol.* **2014**, *133*, 401–404. [[CrossRef](#)]
15. Older, P.; Hall, A.; Hader, R. Cardiopulmonary Exercise Testing as a Screening Test for Perioperative Management of Major Surgery in the Elderly. *Chest* **1999**, *116*, 355–362. [[CrossRef](#)]

16. Clavien, P.A.; Barkun, J.; de Oliveira, M.L.; Vauthey, J.N.; Dindo, D.; Schulick, R.D.; de Santibanes, E.; Pekolj, J.; Slankamenac, K.; Bassi, C.; et al. The Clavien-Dindo classification of surgical complications: Five-year experience. *Ann. Surg.* **2009**, *250*, 187–196. [[CrossRef](#)]
17. Armitage, J.N.; van der Meulen, J.H. Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score. *BJS* **2010**, *97*, 772–781. [[CrossRef](#)]
18. Aletti, G.; Dowdy, S.C.; Podratz, K.C.; Cliby, W.A. Relationship among surgical complexity, short-term morbidity, and overall survival in primary surgery for advanced ovarian cancer. *Am. J. Obstet. Gynecol.* **2007**, *197*, 676–e1. [[CrossRef](#)]
19. Richter, A.N.; Khoshgoftaar, T.M. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artif. Intell. Med.* **2018**, *90*, 1–14. [[CrossRef](#)]
20. Dihge, L.; Ohlsson, M.; Edén, P.; Bendahl, P.O.; Rydén, L. Artificial neural network models to predict nodal status in clinically node-negative breast cancer. *BMC Cancer* **2019**, *19*, 610. [[CrossRef](#)]
21. Hand, D.J. Statistical methods in diagnosis. *Stat. Methods Med. Res.* **1992**, *1*, 49–67. [[CrossRef](#)] [[PubMed](#)]
22. Guo, Y.; Hastie, T.; Tibshirani, R. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics* **2006**, *8*, 86–100. [[CrossRef](#)]
23. Luo, W.; Phung, Q.-D.; Tran, T.; Gupta, S.; Rana, S.; Karmakar, C.; Shilton, A.; Yearwood, J.L.; Dimitrova, N.; Ho, T.B.; et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *J. Med. Internet Res.* **2016**, *18*, e323. [[CrossRef](#)]
24. Khuri, S.F. The NSQIP: A new frontier in surgery. *Surgery* **2005**, *138*, 837–843. [[CrossRef](#)] [[PubMed](#)]
25. Tseng, J.H.; Cowan, R.A.; Zhou, Q.; Iasonos, A.; Byrne, M.; Polcino, T.; Polen-De, C.; Gardner, G.J.; Sonoda, Y.; Zivanovic, O.; et al. Continuous improvement in primary debulking surgery for advanced ovarian cancer: Do increased complete gross resection rates independently lead to increased progression-free and overall survival? *Gynecol. Oncol.* **2018**, *151*, 24–31. [[CrossRef](#)] [[PubMed](#)]
26. Pepin, K.; Bregar, A.; Davis, M.; Melamed, A.; Hinchcliff, E.; Gockley, A.; Horowitz, N.; del Carmen, M.G. Intensive care admissions among ovarian cancer patients treated with primary debulking surgery and neoadjuvant chemotherapy-interval debulking surgery. *Gynecol. Oncol.* **2017**, *147*, 612–616. [[CrossRef](#)]
27. Fournier, M.; Huchon, C.; Ngoa, C.; Bensaid, C.; Batsa, A.; Combeb, P.; Belda, M.L.F.; Berger, A.; Lecurua, F. Morbidity of rectosigmoid resection in cytoreductive surgery for ovarian cancer. Risk factor analysis. *Eur. J. Surg. Oncol.* **2018**, *44*, 750–753. [[CrossRef](#)]
28. Pin, F.; Barreto, R.; Kitase, Y.; Mitra, S.; Erne, C.E.; Novinger, L.J.; Zimmers, T.A.; Couch, M.E.; Bonewald, L.F.; Bonetto, A. Growth of ovarian cancer xenografts causes loss of muscle and bone mass: A new model for the study of cancer cachexia. *J. Cachex Sarcopenia Muscle* **2018**, *9*, 685–700. [[CrossRef](#)]
29. O’Shea, A.; McCool, K.; Harrison, R.; Sampene, E.; Connor, J.; Barroilhet, L. Neoadjuvant chemotherapy is associated with more anemia and perioperative blood transfusions than primary debulking surgery in women with advanced stage ovarian cancer. *Gynecol. Oncol.* **2018**, *150*, 19–22. [[CrossRef](#)] [[PubMed](#)]
30. Fagotti, A.; Ferrandina, M.G.; Vizzielli, G.; Fanfani, F.; Gallotta, V.; Chiantera, V.; Costantini, B.; Margariti, P.A.; Alletti, S.G.; Cosentino, F.; et al. Phase III randomised clinical trial comparing primary surgery versus neoadjuvant chemotherapy in advanced epithelial ovarian cancer with high tumour load (SCORPION trial): Final analysis of peri-operative outcome. *Eur. J. Cancer* **2016**, *59*, 22–33. [[CrossRef](#)]
31. Yang, L.; Zhang, B.; Xing, G.; Du, J.; Yang, B.; Yuan, Q.; Yang, Y. Neoadjuvant chemotherapy versus primary debulking surgery in advanced epithelial ovarian cancer: A meta-analysis of peri-operative outcome. *PLoS ONE* **2017**, *12*, e0186725. [[CrossRef](#)] [[PubMed](#)]
32. DiSilvestro, P.; Colombo, N.; Harter, P.; González-Martín, A.; Ray-Coquard, I.; Coleman, R.L. Maintenance Treatment of Newly Diagnosed Advanced Ovarian Cancer: Time for a Paradigm Shift? *Cancers* **2021**, *13*, 5756. [[CrossRef](#)]
33. Díaz-Montes, T.P.; Zahurak, M.L.; Bristow, R.E. Predictors of extended intensive care unit resource utilization following surgery for ovarian cancer. *Gynecol. Oncol.* **2007**, *107*, 464–468. [[CrossRef](#)] [[PubMed](#)]
34. Ross, M.S.; Burriss, M.E.; Winger, D.G.; Edwards, R.P.; Courtney-Brooks, M.; Boisen, M.M. Unplanned postoperative intensive care unit admission for ovarian cancer cytoreduction is associated with significant decrease in overall survival. *Gynecol. Oncol.* **2018**, *150*, 306–310. [[CrossRef](#)] [[PubMed](#)]
35. Earle, C.C.; Park, E.R.; Lai, B.; Weeks, J.C.; Ayanian, J.Z.; Block, S. Identifying Potential Indicators of the Quality of End-of-Life Cancer Care From Administrative Data. *J. Clin. Oncol.* **2003**, *21*, 1133–1138. [[CrossRef](#)]
36. Birkmeyer, J.D.; Dimick, J.B.; Birkmeyer, N.J. Measuring the quality of surgical care: Structure, process, or outcomes? *J. Am. Coll. Surg.* **2004**, *198*, 626–632. [[CrossRef](#)]
37. Mujagic, E.; Marti, W.R.; Coslovsky, M.; Soysal, S.D.; Mechera, R.; Von Strauss, M.; Zeindler, J.; Saxer, F.; Mueller, A.; Fux, C.A.; et al. Associations of Hospital Length of Stay with Surgical Site Infections. *World J. Surg.* **2018**, *42*, 3888–3896. [[CrossRef](#)]