HAL

archives-ouvertes.fr

Counterfactual approach with survival or time to event outcomes: An application to an exhaustive cohort of Epithelial Ovarian Carcinoma in the Rhône-Alps region of France

Marius Huguet, Lionel Perrier, Olivia Bally, Xavier Joutard, Nathalie Havet, Fadila Farsi, David Benayoun, Pierre De Saint Hilaire, Dominique Beal Ardisson, Magali Morelle, et al.

UMR 5824

93, chemin des Mouilles
69130 Ecully · France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 · France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

Counterfactual approach with survival or time to event outcomes: An application to an exhaustive cohort of Epithelial Ovarian Carcinoma in the Rhône-Alps region of France

Marius Huguet, Lionel Perrier, Olivia Ballyc, Xavier Joutard, Nathalie Havet, Fadila Farsi, David Benayoun, Pierre de Saint Hilaire, Dominique Beal Ardisson, Magali Morelle, Isabelle Ray-Coquard

Abstract:

Epithelial Ovarian Carcinoma (EOC) is a disease with poor prognosis, most often diagnosed at an advanced stage, thus necessitating aggressive and complex surgery. The aim of this study was to compare Progression Free Survival (PFS) at 1st line treatment of EOC patients treated in high vs low-volume hospitals. This retrospective study using prospectively implemented databases was conducted on an exhaustive cohort of 267 patients treated in first line during 2012 in the Rhone-Alps Region of France. In order to control for selection bias, a multivariate analysis and the Inverse Probability Weighting (IPW) using the propensity score were adopted. An Adjusted Kaplan Meier Estimator (AKME) and a univariate Cox model in the weighted sample were then applied in order to determine the impact of the centralization of care on EOC. Patients treated in lower volume hospitals had a probability of relapse (including death) that was 1.5 times higher than for patients treated in higher volume hospitals (p=0.02). As reported in other countries, the concentration of care for EOC has a significant positive impact on patient outcomes.

Working Paper

Counterfactual approach with survival or time to event outcomes: An application to an exhaustive cohort of Epithelial Ovarian Carcinoma in the Rhône-Alps region of France

Marius Huguet[a], Lionel Perrier[b], Olivia Bally[c], Xavier Joutard[d], Nathalie Havet[e], Fadila Farsi[f], David Benayoun[g], Pierre de Saint Hilaire[h], Dominique Beal Ardisson[i], Magali Morelle[b], Isabelle Ray-Coquard[j]

a Univ Lyon, Université Lyon 2, F-69007 Lyon, France.

b Univ Lyon, Leon Berard Cancer Centre , GATE L-SE UMR 5824, F-69008 Lyon, France.

c Leon Berard Cancer Centre, Lyon, France.

d Grequam-UMR 7316, Aix-Marseille University, Marseille, France.

e Univ Lyon, Université Lyon 2, GATE L-SE UMR 5824, F-69008 Lyon, France.

f Réseau Espace Sante Cancer Rhône-Alpes, Lyon, France.

g Hospital Lyon Sud, Pierre-Bénite, France.

h University Hospital of Lyon, Department of medicine. Lyon, France.

i Private Hospital Jean Mermoz, Lyon, France.

j Univ Lyon, Leon Berard Cancer Centre , EA7425 HESPER, F-69008 Lyon, France.

Abstract

Epithelial Ovarian Carcinoma (EOC) is a disease with poor prognosis, most often diagnosed at an advanced stage, thus necessitating aggressive and complex surgery. The aim of this study was to compare Progression Free Survival (PFS) at 1st line treatment of EOC patients treated in high vs low-volume hospitals. This retrospective study using prospectively implemented databases was conducted on an exhaustive cohort of 267 patients treated in first line during 2012 in the Rhone-Alps Region of France. In order to control for selection bias, a multivariate analysis and the Inverse Probability Weighting (IPW) using the propensity score were adopted. An Adjusted Kaplan Meier Estimator (AKME) and a univariate Cox model in the weighted sample were then applied in order to determine the impact of the centralization of care on EOC. Patients treated in lower volume hospitals had a probability of relapse (including death) that was 1.5 times higher than for patients treated in higher volume hospitals (p=0.02). As reported in other countries, the concentration of care for EOC has a significant positive impact on patient outcomes.

1. Introduction

Observational studies are increasingly being used in the health care sector, for several reasons. It is sometimes impossible to randomize the affectation to treatment for practical or ethical reasons. It is often the case in retrospective studies analyzing the care pathway. In this context of observational data, we must take into account the selection bias due to the sample heterogeneity [1]. Indeed, a selection bias, or recruitment bias, could emerge since participation to the treatment is also not random - some kinds of patients have a higher probability of being treated than others. Several well-known methods can be used to solve this issue, such as stratification or multivariate analysis, and more sophisticated method are more and more used such as matching methods, instrumental variable, regression by discontinuity and difference in difference.

Epithelial ovarian carcinoma (EOC) is a serious disease which extend is still underestimated due to a rather low incidence and a significant mortality. The EOC remains the eighth cancer sorted by frequency order for women, with an incidence of 7.9 for 100 000 persons/year. EOC remains the main gynecologic cancer death cause in industrialized countries with a mortality rate in France estimated to 4/100 000 persons per year [2]. Relapse-free survival and overall survival of patients are related to the characteristics of the disease, the patient herself and the disease management. Several recent retrospective studies investigated the relationship between outcomes of ovarian cancer treatment and type of care provider [3]. The results showed a better quality of surgery when performed by "gynaecologic oncologists" in specialized hospitals (referent centers) and only small differences in chemotherapy regimens between the settings. Some studies [4-14] also investigated the impact of centralization of care, in terms of volume, in patients' outcomes. Patients are more likely to be optimally debulked in a high volume hospital or in a specialized provider. These studies also showed that patients have better survival outcomes in high volume hospital. But most of these studies focus on high stage disease, and none have been done in France. The majority of patients with ovarian cancer do not receive care in specialized settings.

The aim of this study was to compared Progression Free Survival (PFS) in 1st line treatment of EOC patients treated in high (HVH) versus low (LVH) volume hospitals taking into account all prognostic factors. To that end, we used observational data, and thus have to control for cofounder factor.

All Counterfactual methods are not always applicable depending on the context, database and are based on different hypotheses. For difference in difference method, we need to observe patients before and after treatment (longitudinal data) in order to compare the outcome of each patient before and after treatment, and between treated and untreated patients. This method was not applicable in our case because we had cross sectional data. With regression by discontinuity, patients on either side of the threshold to define high volume hospitals are considered almost identical, and thus, can be compared in

order to estimate the treatment effect. This method requires a large dataset which was not the case in our study. The Instrumental Variable method treats the endogeneity of the treatment variable directly in the regression. But this method requires to find a good and valid instrument, which is a variable strongly correlated to the treatment variable and uncorrelated to the outcome. This method was not applicable because such a variable was not available in our database. Finally, we used multivariate analysis and matching methods in order to treat the selection bias because they are less restrictive in terms of sample size and are applicable with cross sectional data.

2. Counterfactual methods

2.1. Different types of treatment effect: marginal versus conditional and relative versus absolute

The multivariate analysis allows one to estimate a conditional effect; it is the average effect of being treated in higher volume hospital at the individual level, as if a patient in a low volume hospital was treated in a higher volume hospital. It differs from the marginal treatment effect, which is the difference in outcome between the groups of patient in high versus low volume hospital [15]. But, to estimate the marginal treatment effect, we needs that the two groups are identical in all aspects, except of being treated in a high-low volume hospital. In a RCT, it is possible to estimate the marginal treatment effect by simple difference in outcome because the randomization balances the covariates, but it is impossible in observational studies with time to event outcome because the conditional differs from the marginal treatment effect. Moreover, Austin PC (2014) [15] made some recommendation in the estimation of a treatment effect in an observational study. He recommends to estimate both absolute and relative measure of treatment, as the CONSORT statement recommendation for a RCT with a dichotomous outcome. Again, multivariate analysis allows one to estimate a relative effect, while it is possible to estimate both absolute and relative treatment effect with a propensity score approach. Indeed, the estimation of the treatment effect in observational studies should mimic treatment effect estimation in a RCT. And in a RCT, the conditional effect equal the marginal effect, so we should find similar results if the selection into treatment is based only on observable characteristics. If not, it could be due to unobserved characteristics interacting in the estimation of the treatment effect, and lead to different treatment effect.

2.2. Multivariate analysis

One common approach when dealing with cofounding factor is to use multivariate regression [1]. The principle was to regress the survival time on the treatment variable (an indicator variable denoting high or low volume hospital) and a set of patients' characteristics. In practice, we first ran a Cox proportional hazard model of the Progression Free Survival (PFS) on the set of covariates, and then tested whether the hazard is proportional or not by the Schoenfeld residual test. Then, if the proportional hazard assumption holds, the model to prefer is a semi parametric Cox proportional hazard regression. If not, we turned out on a parametric estimation of an accelerated failure time (AFT) model. With AFT model, we had to choose a parametric distribution of the hazard. A common practice was to estimate, at first, a Generalized Gamma model. The density function is given in equation (1).

$$f(t) = \frac{\gamma\sigma(\gamma t)^{\sigma k-1}e^{-(\gamma t)^{\sigma}}}{\Gamma(k)} \quad (1)$$

With $\gamma_i = e^{-(X_i\beta)}$, and k and σ are parameters linked to the hazard's function. The Generalized Gamma distribution includes the Exponential ($k = \sigma = 1$), Weibull ($k = 1$), Lognormal ($k = 0$) and Gamma ($\sigma = 1$) distributions. Then, it was possible to test for these parameters in order to choose between these distributions by a likelihood ratio test.

2.3. Propensity score matching using the inverse probability weighting (IPW)

In order to estimate the marginal effect of being treated in a high volume hospital, matching methods can be used. There are many different matching methods, but two simulations studies (Handouyahia A. et al [16], Austin PC (2013) [17]) had shown that the IPW seems to performs better in estimating the marginal hazard ratio of the treatment effect, compared with other matching methods using the propensity score.

The IPW method balances the covariate of the two groups by weighting all patients of the data base by the inverse of the propensity score. The propensity score was the conditional probability for a patient to be treated in a high volume hospital, conditionally to observables characteristics. This probability is estimated by a logit model of an indicator variable denoting high or low volume hospital on a set of covariates. The standardized difference in mean is recommended to compare the baseline characteristics, instead of t-test [15, 18].

The stabilized weighted of the IPW proposed by J.M. Robins [19] can be used in order to reduce the volatility of the weights. Note T_i the treatment variable, p_i the propensity score and $f(T)$ the distribution

of the treatment estimate by another logit model. In order to estimate the Average Treatment effect on the Treaties (ATT), weighted can be calculated with the formula in equation (2).

$$w_i^{ATT} = f(T) * [\ T_i + \frac{p_i(1-T_i)}{(1-p_i)}]\quad (2)$$

An Adjusted Kaplan Meier Estimator (AKME), proposed by Xie and Liu [20] and a univariate Cox model in the weighted sample, as described by Cole and Hernan [21], were then applied in order to determine respectively the absolute and relative impact of the concentration of care on EOC. We used the robust variance estimator of Lin and Wei [22] for the weighted Cox model and the Cox test of equality for the AKME, to take into account the within matched set correlation due to matching. Patients in high versus low volume hospital are not anymore independent after matching using the IPW, because they are linked by their propensity score. Indeed, two patients with similar propensity score will have similar characteristics, and patients' characteristics are strongly correlated with outcome.

3. Application

3.1 Study design

This retrospective study using prospectively implemented databases was conducted on an exhaustive cohort of patients treated in first line during 2012 in the Rhone-Alps Region of France. The inclusion criteria were to have been in 1st line treatment of Epithelial Ovarian Cancer (EOC), diagnosed in 2012, incident case, more than 18 years old, residency in France. The exclusion criteria were non-epithelial disease, relapse disease or being less than 18 years old. The study was conducted in accordance with the ethical principles for medical research involving human subjects developed in the Declaration of Helsinki by the World Medical Association (WMA). The study received approval in France from the National Ethics Committee (N°909226) and the National Committee for Protection of Personal Data (N°09-203).

Several international studies [4-14] have used a threshold to define a high volume hospital between 10-20 cases per year for EOC. But, it appears that the median of the volume of patients treated in a hospital per year in the Rhone-Alpes region of France is 4. And on average, hospital in the region considered in this study had treated less than 6 patients per year in 2012. Therefore, we considered a threshold of 10 patients treated for EOC per year in primary care, which is strict enough in the case of France and which has also been used by Mercado C et al (2010) [8], Bristow RE et al (2015) [6]. Thus, 90% of hospital had treated less than 10 patients per year in the region in 2012. Different countries need different threshold to define a high volume hospital, according to its prevalence of the disease. For example, the

mean volume activity of high volume hospital in the study of Ioka A et al (2004) [9] on a Japanese dataset was 8.8, which can seems low compared to American studies.

3.2 Patient's characteristics

In 2012, 267 patients were identified with an EOC in the region, but only 231 were used in modelization because of missing data. Patients were treated by 49 different hospitals in the region. The median volume activity by hospital is 4 patients treated for EOC per year. Volume activity is disparate among hospital, from a minimum of 1 patient to a maximum of 42 patients in 2012. N=98 (37%) patients in 1st line setting for EOC in the region have been treated in a high volume hospital.

Table 1 : Characteristics of the patients

	Low volume hospital (n=98)		High volume hospital (n=169)			
	m	σ	m	σ	P-value	% bias
Age	63.68	13.79	65.80	12.09	0.205	16.4
Cancer history	13.60	0.34	17.34	0.38	0.411	10.03
Ascites	59.76	0.49	67.34	0.47	0.218	15.7
Histology :						
- HGSC & undiff	52.25	0.50	59.14	0.49	0.294	13.8
- SLG	7.09	0.25	3.22	0.17	0.203	-17.5
- Muccinous	9.67	0.29	4.30	0.20	0.124	-21.1
- Endometrioid	15.48	0.36	10.75	0.31	0.296	-14.0
- Clear cell	5.91	0.23	5.10	0.22	0.781	-3.6
- Unknown	9.03	0.28	17.20	0.37	0.056	24.3
FIGO Stage :						
- I	25.44	0.43	15.46	0.36	0.058	-24.8
- II	5.32	0.22	7.21	0.26	0.534	7.8
- III	55.62	0.49	65.97	0.47	0.098	21.3
- IV	13.60	0.34	11.34	0.31	0.595	-6.8
Tumor Grade :						
- 1	18.05	0.38	4.54	0.20	0.002	-43.5
- 2	20.83	0.40	16.63	0.34	0.168	-19.1
- 3	61.11	0.48	81.81	0.38	0.000	46.9
					mean	19.2
					median	16.9

m : mean (frequency) if the covariate is continuous (binary) / σ : standard deviation
SLG : Serous of Low Grade ; HGSC & undiff : High Grade Serous Carcinoma & undifferentiated

At baseline (i.e. before matching), patients treated in higher volume hospital seem to have a higher tumor grade at a 5% level of significance, and higher FIGO stage at a 10% level of significance, compare with patients treated in lower volume hospital (table 1). On the other hand, there are no significant difference between patients in high versus low volume hospital in terms of age, ascites and history of cancer. More details are provided in table 1.

3.3 Standard Kaplan Meier estimator

In the health care sector, not all studies take into account the selection bias in statistical analysis. In our case, a simple way to estimate the treatment effect without solving the recruitment bias issue is to estimate a standard Kaplan Meier estimator for each group of patient. Then, we can compare the two survivals curves with a log rank test in order to see if the difference is significant. This naïve estimation of the treatment effect is given by figure 1.
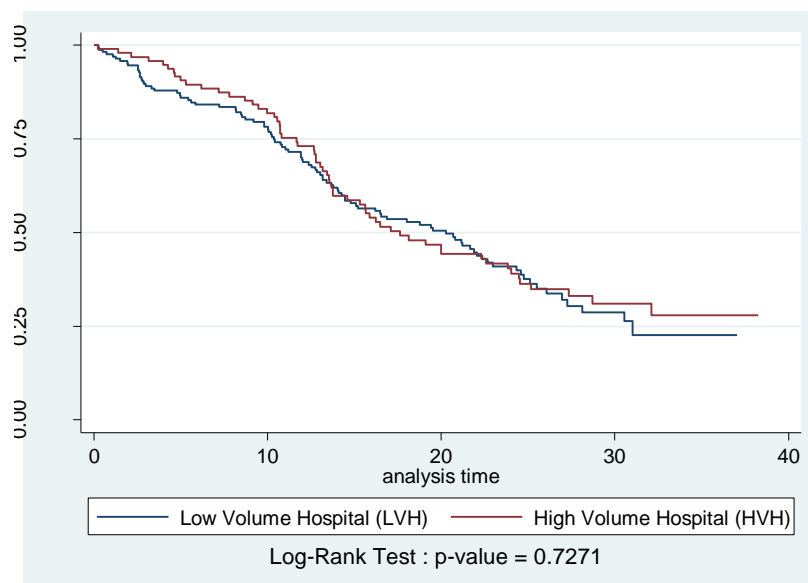

Figure 1: Standard Kaplan Meier estimator of the PFS

This estimation of the treatment effect is based on a sample homogeneity assumption. It appears in figure 1 that the PFS is not statistically different in high than in low volume hospital. But it is clear that this estimation is biased by the selection bias. Indeed, survival in high and in low volume hospital seems to be comparable, while high volume hospitals are treated patients of higher stage and grade of the disease.

3.4 Multivariate analysis approach

The Schoenefeld residual test showed that the proportional hazard assumption did not holds (p=0.0804). Thus, the model is not applicable and we turn out in a parametric estimation of an AFT model. It appears that the Weibull distribution fitted best our data. We chose Weibull instead of Gompertz and Loglogistic, which are not a particular case of the generalized gamma, because the AFT model with a Weibull distribution minimized the Akaike Information Criterion (AIC). Table 2 shows that patients treated in higher volume hospital have a higher PFS at a 10% level of significance than patients in lower volume hospital, on average and all other things being equal.

Table 2 : A Weibull accelerated failure time model on PFS

	Coefficient	σ	P-value
High volume hospital	0.23	0.130	0.066
Age	-0.01	0.005	0.012
Cancer history	-0.23	0.155	0.132
Ascites	-0.33	0.148	0.026
Histology :			
- HGSC & undiff	Ref	Ref	Ref
- SLG	0.20	0.499	0.678
- Muccinous	0.17	0.466	0.713
- Endometrioid	0.25	0.258	0.316
- Clear cell	-0.07	0.285	0.793
- Unknown	-0.19	0.215	0.362
FIGO Stage :			
- I	1.02	0.314	0.001
- II	0.63	0.336	0.060
- III	0.31	0.166	0.057
- IV	Ref	Ref	Ref
Tumor Grade :			
- 1	0.01	0.423	0.964
- 2	-0.09	0.157	0.527
- 3	Ref	Ref	Ref
σ : standard deviation / Ref : modality in reference SLG : Serous of Low Grade ; HGSC & undiff : High Grade Serous Carcinoma & undifferentiated			

3.5 Propensity score approach

In order to control for the selection bias, to estimate both a relative and an absolute treatment effect and that this treatment effect is a marginal effect, we used the Inverse probability weighting using the propensity score (IPW). Table 3 shows a good quality of the matching by the IPW. Indeed, there is no significant difference for all covariates between the two groups, while there were significant differences

prior matching for stage, grade and histology. The mean standardized difference in mean is 1.7 for the matched sample (Cf table 3) instead of 19.2 for the unmatched sample (Cf table 1). Moreover, the common support of the distribution of the propensity score is sufficient to valid the overlap assumption (figure 2). Indeed, the distribution of the propensity score for treated and untreated patients showed that none treated patients were excluded of the matching because of their propensity score.
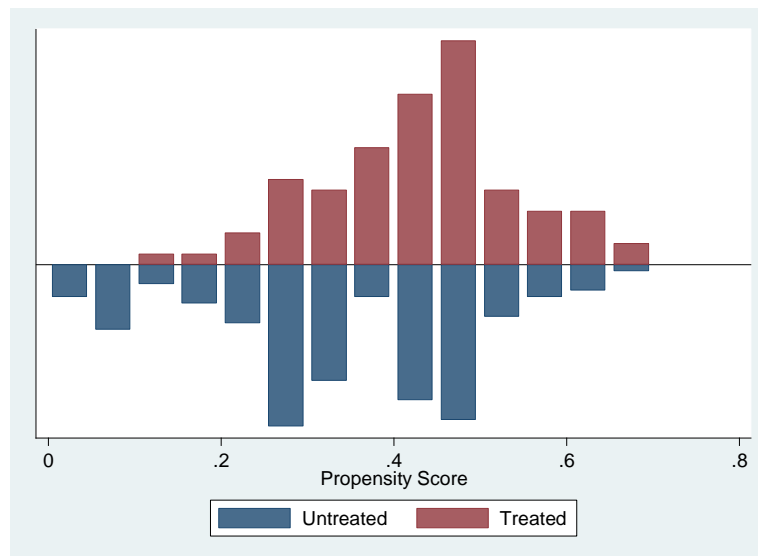

Figure 2 : Distribution of the propensity score

Matching using the IPW allow to estimate both the absolute treatment effect, with the AKME, and the relative reduction of an event occurring by the univariate weight cox model. Figure 3, based on the AKME, shows that patients in high volume hospital have a significant higher PFS (p=0.020) than patients in lower volume hospital. The median of the volume activity is 19.1 for high volume hospital instead of 14.1 for low volume hospital.

Figure 3: AKME of PFS after matching using the inverse probability weighting (ATT weight)
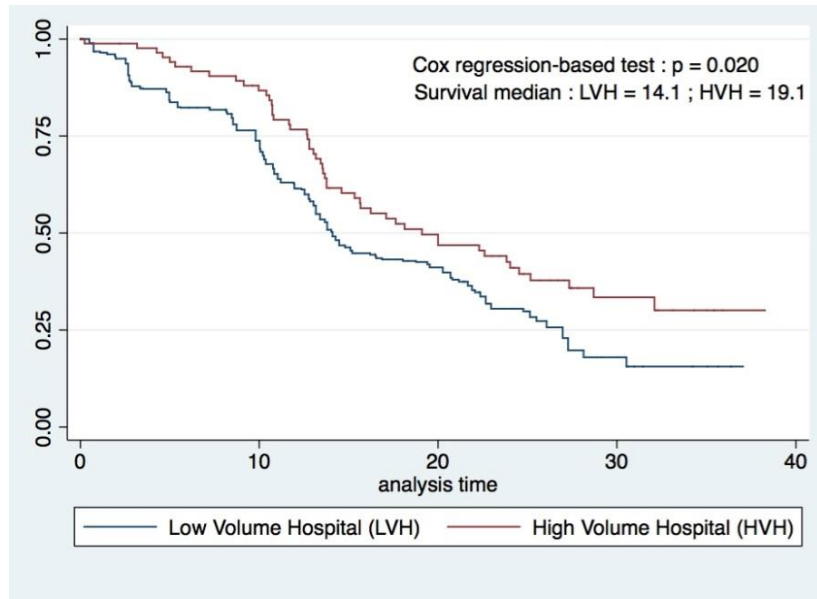
Table 3: Characteristics of the patients after using IPW matching

	High volume hospital (n=144)	Low volume hospital (n=87)		
	m	m	Pvalue	% bias
Age	65.97	65.39	0.822	-4.5
Cancer history	18.50	18.39	0.989	-0.3
Ascites	66.50	66.66	0.986	0.3
Histology :				
- HGSC & undiff	62.62	63.21	0.953	1.2
- SLG	3.17	3.44	0.941	1.3
- Muccinous	3.52	3.44	0.984	-0.3
- Endometrioid	10.82	11.49	0.919	2.0
- Clear cell	6.45	5.74	0.885	-3.1
- Unknown	13.39	12.64	0.914	-2.2
FIGO Stage :				
- I	15.77	16.09	0.967	0.8
- II	8.53	8.04	0.932	-2.0
- III	65.24	64.36	0.930	-1.8
- IV	10.44	11.49	0.873	3.2
Tumor Grade :				
- 1	4.23	4.59	0.933	1.2
- 2	13.18	13.79	0.932	1.6
- 3	82.58	81.60	0.903	-2.2
			Mean :	1.7
			Median :	1.7

m : mean (frequency) if the covariate is continuous (binary) / σ : standard deviation / %bias, also call standardized difference in mean

Furthermore, the univariate Cox model of the PFS, weighted by the inverse of the propensity score, reveals that the hazard ratio of being treated in a high volume hospital is 0.65 (p=0.020). The Schoenfeld residual test reveals that the proportional hazard assumption is valid for the univariate weighted Cox model, and confirms the robustness of the result. Thus, patients treated in lower volume hospitals have a probability of relapse (including death) 1.5 times higher compared with patients treated in higher volume hospitals (p=0.02).

4. Conclusion

Patients treated in high versus low volume hospital aren't similar (table 1). Thus, we expect the presence of selection bias, which means that the participation into treatment (i.e. being treated in a high volume hospital) is not random [1]. In other word, some types of patients are more likely to be treated in high volume hospital than other, and this must be taking into account in order to estimate a proper treatment effect.

Propensity score approach is based on less constraining assumption than multivariate analysis [23]. Indeed, propensity score and multivariate analysis are based on the conditional independence assumption (CIA), which tells that conditionally on observed covariates, patients are treated in a high or low volume hospital randomly. But multivariate analysis requires stronger assumption about the distribution of the covariates and their relation with the relapse free survival. In our case, we also had to choose a distribution of the hazard in order to fit a parametric AFT model of the progression free survival on a variable denoting treatment and on a set of covariates because the proportional hazard assumption was violated.

Therefore, the combination of a multivariate analysis and a matching method allow one to estimate both conditional and marginal effect of being treated in a high volume hospital and to prove the robustness of our findings. The conditional effect tells us that if a patient treated in a lower hospital was treated in a higher volume hospital, this would on average improve her progression free survival (p=0.066). However, the marginal treatment effect tells us that patients treated in higher volume hospital have a probability of relapse (including death) 1.5 times inferior to patients treated in lower volume hospital (p=0.020) and that the absolute difference of survival is significant (p=0.020) (See figure 3). We can be confident about the robustness of our result, because both the parametric (AFT model) and non-parametric (propensity score) approach led to similar results.

Many countries already require a minimum activity for a hospital to have the authorization for cancer treatment. In France, the minimum cutoff in order to have an authorization to treat gynecologic cancer

is defined by the French ministerial order of the 27 march 2007 at 20 surgeries per year. Below this volume of activity, a hospital cannot treat patients with gynecologic cancer. But this threshold takes into account all different kind of gynecologic cancer, such as cervical, ovarian, vaginal, uterine and vulvar cancer. Our findings would suggest that there is a need of a specific minimum activity cutoff only for ovarian cancer. Indeed, the global threshold of 20 cases per year does not assure a sufficient activity for each disease include in gynecologic cancer. In the Rhone-Alps Region of France, 63% of patients treated in primary care in 2012 were treated in a hospital with less than 10 cases per year, and 24% in a hospital with less than 5 cases per year. Our findings would encourage the use of a specific cutoff for ovarian cancer, and more research need to be done for other rare cancers in order to see if a specific minimum activity cutoff is needed too.

As reported in other countries [4-14], the concentration of care for EOC has a significant positive impact on patient relapse free survival. High volume hospitals are mostly treated advanced stage EOC, while it is clear that the concentration of care improve patient survival for both advanced and early EOC. More research needs to be done with dynamic treatment methods in order to take into account the entire care pathway. Indeed, in our study, we only considered the hospital of 1st line treatment which is the one of primary interest because the 1st line surgery is strongly correlated with survival. But it would be interesting to see how the effect of volume goes on outcomes if we considered hospitals which are treating relapse disease. More research also need to be done on the cost associated with the concentration of care and on the likely simultaneous causality between volume and survival.

Acknowledgments

References

1. Michael L. Johnson, William Crown, Bradley C. Martin, Colin R. Dormuth, Uwe Siebert (2009). Good Research Practices for Comparative Effectiveness Research:Analytic Methods to Improve Causal Inference from Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part III. *Value in Health*.

2. Réseau FRANCIM (2010). Projection de l'incidence et de la mortalité par cancer en France en 2010. Rapport Technique avril 2010.

3. M. Fung-Kee-Fung, E.B. Kennedy, J. Biagi, T. Colgan, D. D'Souza, L.M. Elit et al (2015). The optimal organization of gynecologic oncology services: a systematic review. Current Oncology. Vol 22, No 4.

4. Bristow RE, Palis BE, Chi DS, Cliby WA. (2010). The National Cancer Database report on advanced-stage epithelial ovarian cancer: impact of hospital surgical case volume on overall survival and surgical paradigm. *Gynecol Oncol.* 118:262-7.

5. Bristow RE, Chang J, Ziogas A, Randall LM, Anton-Culver H (2014). High-volume ovarian cancer care: Survival impact and disparities in access for advanced-stage disease. *Gynecol Oncol.* 132:403-10.

6. Bristow RE, Chang J, Ziogas A, Campos B, Chavez LR, Anton-Culver H (2015). Impact of National Cancer Institute Comprehensive Cancer Centers on ovarian cancer treatment and survival. *J Am Coll Surg.* 220:940-50.

7. Reade C, Elit L (2012). Trends in gynecologic cancer care in North America. *Obstet Gynecol Clin N Am.* 39:107-29.

8. Mercado C, Zingmond D, Karlan BY, Sekaris E, Gross J, Maggard-Gibbons M et al (2010). Quality of care in advanced ovarian cancer: the importance of provider speciality. *Gynecol Oncol.* 117:18-22.

9. Ioka A, Tsukuma H, Ajiki W, Oshima A (2004). Influence oh hospital procedure volume on ovarian cancer survival in Japan, a country with low incidence of ovarian cancer. *Cancer Sci.* 95:233-

10. Woo YL, Kyrgiou M, Bryant A, Everett T, Dickinson HO (2012). Centralization of services for gynaecological cancers: A Cochrane systematic review. *Gynecol Oncol.* 126:286-90.

11. Phippen NT, Barnett JC, Lowery WJ, Miller CR, Leath CA 3rd (2013). Surgical outcomes and national comprehensive cancer network compliance in advanced ovarian cancer surgery in a low volume military treatment facility. *Gynecol Oncol.* 131:158-62.

12. Cowan RA, O'Cearbhail E, Gardner GJ, Levine DA, Roche KL, Sonoda Y et al (2015). Is it time to Centralize Ovarian Cancer Care in the United States? *Surgical Oncology.* 23:998-93.

13. Vernooij F, Heintz AP, Coebergh JW, Massurger LF, Witteveen PO, Van der Graaf Y (2009). Specialized and high volume care leads to better outcomes of ovarian cancer treatment in the Netherlands. *Gynecol. Oncol.* 112:455-61.

14. Aune G, Torp SH, Syversen U, Hagen B, Tingulstad S (2012). Ten Years' Experience With Centralized Surgery of Ovarian Cancer in One Helath Region in Norway. *International Journal of Gynecological Cancer.* 22: 226-231.

15. Austin PC (2014). The use of propensity score methods with survival or time-to-event outcomes : reporting measures of effect similar to those used in randomized experiments. *Statistics in Medicine, 33 1242-1258.*

16.   Handouyahia A., Haddad T., Eaton F. (2013). Kernel Matching versus Inverse Probability Weighting: A comparative Study. *International Science Index, Mathematical and Computational Sciences,* 7(8).

17.   Austin PC (2013). The performance of different propensity score methods for estimating marginal hazard ratios. *Statistics in Medicine.* 32(16):2837-2849.

18. Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical Science : A Review Journal of the Institute of Mathematical Statistics. 25(1), 1–21.

19.   Robins JM (1998), Marginal structural models, in: 1997 Proceedings of the American Statistical Association, Section on Bayesian Statistical Science, pp. 1-10.

20.   Xie J, Liu C (2005). Adjusted Kaplan-Meier estimator and log rank test with inverse probability of treatment weighting for survival data. *Statistics in Medicine.* 24(20):3089-3110.

21. Cole SR, Hernan MA (2004). Adjusted survival curves with inverse probability weights. *Computer Methods and programs in Biomedicine*; 75:45-49.

22.   Lin DY, Wei LJ (1989). The robust inference for the proportional hazards model. *Journal of the American Statistical Association*; 84:1074-1078.

23.   Smith JA, Todd PE (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics.* 125: 305-353.