استخدام نموذج انحدار كوكس لأوقات بقاء المرضى المصابين باللوكيميا في العراق

Using Cox Regression Model in Survival Time for leukemia patients in Iraq

محمد توفيق البلقيني/ استاذ الرياضيات والاحصاء الاكتواري كلية التجارة _جامعة المنصورة

بيومي عوض طاقية/ استاذ الاحصاء التطبيقي كلية التجارة _جامعة المنصورة

حياة سعيد رديعان/ باحثة ماجستير - كلية التجارة_ جامعة المنصورة.

الملخص

في هذا البحث تم استخدام نموذج انحدار كوكس وتطبيقه على مرضى سرطان الدم في العراق، ولقد تم الحصول على البيانات من مستشفى مدينة الطب في بغداد، حيث تم أخذ عينة مكونة من (62) مريض مصاب بسرطان الدم ولفترة محدودة مكونة من 4 شهور، حيث تمت متابعة المرضى من 5/1 حتى 1 /9 لسنة 2015 حتى الوفاة أو الفقد، ولقد تضمنت هذه العينة على معلومات ديموغرافية وهي العمر والجنس وعنوان المريض والمهنة، ومتغيرات طبية وهي حالة فقر الدم ونوع المرض ونوع العلاج. وقام الباحث بتقدِّر معلمات النموذج بطريقة الامكان الأعظم الجزئي (Partial Likelihood)، واختبار المتغيرات بواسطة اختبار والد (Wald Test)، حيث تبين أن متغير نوع العلاج والعنوان هما المتغيران المعنويان في النموذج والمؤثران في زمن البقاء، أستخدم الباحث اختبار نسبة الامكان الأعظم (Likelihood Ratio Test) وتبين ان النموذج الذي يحتوي متغير المعالجة ومتغير العنوان هو اكثر ملائمة لبيانات هذا البحث.

Abstract: In this paper we used Cox- Regression model as the best model to analysis data, and applied to patient with Leukemia in Iraq. The data which is used in this research was taken from the Medical city Hospital in Baghdad includes (62) patient during (1/5-1/9).we estimated models parameter by using (partial likelihood) method and then test the models parameter by using (Wald test) and the result shown that the variables (treatment type) and (address) has the effect on survival time. And we used (Likelihood ratio test) it was shown that the model which contains tow variables (treatment, address) is the best model to study.

المقدمة

نظرا لأهمية موضوع زمن البقاء وتأثرة بعدد من العوامل فقد ظهرت الوسائل والأساليب الأحصائية المتعلقه بزمن البقاء، حيث أن تحليل بيانات البقاء يمثل دراسة الوقت المنقضي ما بين وقوع حدث البداية (Start Event) حتى وقوع حدث النهاية (terminal event).

لذلك فأن نماذج الانحدار تكون ذات أهمية في هذه الدراسات ولكن بشكل مطور بحيث يتلائم مع حالة المتغير التابع ثنائي الاستجابة والذي يمثل المدة الزمنية التي تحتوي على الكثير من المشاكل مثل الرقابة (censoring) والتعقيدات التي تسببها.

عند ذلك تم اقتراح عدة طرق للدراسة الا ان اشهر وانسب طريقة لبناء نماذج انحدار لبيانات البقاء تتمثل في نموذج انحدار كوكس والذي وضع من قبل العالم الانكليزي Cox dived لاستخدامة في تحليل البقاء، وقد أستخدم في تحليل التجارب الطبية من خلال تأثير عدد من المتغيرات على مدة البقاء.

وفي هذا البحث سوف نتناول امكانية تطبيق نموذج كوكس للانحدار على مرضى سرطان الدم من خلال نمذجة تأثير عدد من العوامل والمتغيرات والاعتبارات على مخاطر حيوية ما، ويعتبر هذا المرض من الامراض التي تهدد هذا العصر ويسبب مشاكل كبيرة، لذلك فأن هذا الامر يعتبر من الامور المهمة يجب تسليط الضوء عليه.

مشكلة الدراسة:

أن حاجة الانسان الى الاستمرار بالحياة دفعته الى دراسات مهمة تتعلق بزمن البقاء والعوامل المؤثرة فيه، حيث يعتبر تحليل البقاء ضروري عند دراسة الأنظمة التي يكون فيها المتغير التابع هو الوقت حتى حدوث حدث معين وعليه فأن مشكلة الدراسة تتمثل في دراسة ومعرفة العوامل المؤثرة في أوقات البقاء لمرضى سرطان الدم في العراق وتحديد اهم المتغيرات المؤثرة في زمن البقاء وذلك عن طريق تطبيق نموذج أنحدار كوكس.

هدف الدراسة:

أن الهدف من الدراسة هو تحديد وقياس أهم العوامل التي تؤثر على زمن البقاء لمرضى سرطان الدم في العراق، ومعرفة أهم الأسباب التي تؤثر في زمن البقاء بأستخدام نموذج الانحدار كوكس (Cox Regression Model).

نموذج الانحدار Cox:

يعرف نموذج الأنحدار كوكس بأنه أسلوب أحصائي لأكتشاف العلاقة بين مدة بقاء المريض وعدة متغيرات تفسيرية أخرى، وهو احد النماذج التي تهتم بالمخاطر النسبية والتي تتعامل مع الزمن في التحليلات الأحصائية، كما انه يعتبر من أهم الأساليب الأحصائية المستخدمة في تحليل البقاء (survival analysis) والذي يطبق أحصائيا على نطاق واسع (Fox 2002)، وأن نموذج الخطر النسبي

كوكس هو نموذج شبه معلمي لاحتواء النموذج على بعض المتغيرات التي تتبع العائلة المعلمية وهي في الجزء $[exp(Bx)]$ من النموذج، ويحتوي على متغيرات اخرى لاتتبع العائلة المعلمية وهي المتغيرات الموجودة في الجزء $[h_\circ(t)]$ من النموذج $(Crowder\ et\ al., 1991)$.

ولقد افترض العالم كوكس ان شكل النموذج يعتمد على معدل الخطورة (Hazard Rate) عند الزمن t، وتعرف صيغة النموذج عندما يكون الوقت مستمرا بالشكل الاتي (Cox 1972):

$$h(t,x) = h_\circ(t).exp(Bx) \quad \dots \ (1)$$

كما يتم تقدير متجه المعالم β بأستخدام دالة الامكان الجزئية وهي كالاتي (Cox 1975):

$$L = \prod_{k=1}^{r} \frac{exp(B\ x_k)}{\sum_{J\in R(t_i)} \exp(B\ x_J)} \qquad \dots (2)$$

و تتميز دالة الامكان التي قدمها (Cox 1975) بأنها تقدم تقديرات كفؤ لمعلمات المجتمع(Efron B., 1977) حيث أن:

$(t_k, k = 1,2, \dots \dots . r)$: هي أوقات البقاء غير المراقبة والمرتبة تصاعديا على الصورة التالية: $t_1 < t_2 < \dots < t_r$

r : هي عدد المشاهدات الغير مراقبة، وان عدد المشاهدات الكلية هو n.

$R_{(t_i)}$: تمثل مجموعة الخطر عند الزمن t_i.

وتمثل هذه دالة الامكان الجزئي ان في كل لحظة زمنية t_i هناك عنصر واحد فقط وقع له الحدث اي لايمكن ان يوجد حدثين في نفس الوقت.

وبأخذ اللوغارتيم للدالة :

$$LogL = \sum_{k=1}^{r} x_k\, B - Log \sum_{j\in R(t_i)} \exp(x_J\, B)$$

بأخذ المشتقة الأولى بالنسبة ل B

$$\frac{dLogL}{dB} = x_k - \frac{\sum_{j\in R(t_i)} \exp(x_J\, B)\ x_J}{\sum_{j\in R(t_i)} \exp(x_J\, B)}$$

وبعدها نساوي المشتقة بالصفر وبأستخدام طريقة نيوتن رافسون نحصل على تقديرات المعلمات $B_1, B_2, \dots, B_p$، وبأستخدام العلاقة الاتية:

$$\hat{B}^{k+1} = \hat{B}^k - [HE^k]^{-1}U^k \quad \dots \quad (3)$$

حيث أن:

$\hat{B}$: متجه المعالم المقدرة وأن HE و U عبارة عن مصفوفة المشتقة الثانية ومتجه المشتقة الاولى على التوالي لدالة الامكان الاعظم L (B).

كما يمكن ايجاد دالة الخطر الاساسية $h_\circ(t)$ عندما تكون المتغيرات التفسيرية مساوية للصفر ولنفرض انتقال حدث من حالة حياة الى حالة وفاة وان كل أوقات البقاء لها وقت بداية وهو صفر، وان أوقات النهاية المرتبة هي $t_1 < t_2 <$

$t_z < \ldots$ ولنفرض أن Ei يمثل عدد الحوادث عند الزمن t_i وأن R_i تمثل مجموعة التعرض للخطر عند الزمن t_i وأن $\hat{B}$ تمثل مقدرات الأمكان الجزئي لمعلمات النموذج، وعند $t_\circ = 0$ يمكن أعتبار أن $\lambda_{i(t)}$ هي معدل الخطر الاساس للنموذج وهو ثابت خلال الفترة $(t_{i-1} - t_i)$ (عزيز 2011). وتظهر كدالة خطوة (Step function) بخطوات عند النقاط الزمنية المقترنة، ويمكن أن تمثل كما يأتي:

$$\boldsymbol{\lambda_i(t) = \frac{E_i}{(t_i - t_{i-1}) \sum_{j \in R(t_i)} \exp(x_j B)}} \quad \ldots . (4)$$

حيث أن: $\boldsymbol{i = 1, 2, .., z}$

●متغيرات الدراسة

تتمثل متغيرات الدراسة فى العوامل التى تؤثر فى سرطان الدم وهى كما يلى:
المتغير التابع : زمن البقاء ((Survival Time (ST) وهو زمن بقاء المرضى المصابين بمرض سرطان الدم حتى الوفاة أو الفقدان.
متغير الحاله : (Status) اى حالة خروج المريض من المستشفى ويأخذ القيم:
1 : تمثل حالة الوفاة ، 0 : تمثل حالة المراقبة

المتغيرات المستقلة:

X_1 :العمر بالسنين عند اصابته بالمرض.
X_2 :النوع ويأخذ القيمة (1) اذا كان المريض انثى ، (2) اذا كان المريض ذكر
X_3:العنوان ويأخذ القيمة (1) اذا كان المريض داخل بغداد، (0) اذا كان المريض خارج بغداد
X_4 :المهنة ويأخذ القيمة (1) اذا كان المريض ربة بيت ، (2) اذا كان المريض طالب، (3) اذا كان المريض كاسب، (4) اذا كان المريض متقاعد ، (5) اذا كان المريض موظف
X_5 :نوع المرض ويأخذ القيمة (1) اذا كان المرض سرطان الدم اللمفاوى الحاد ALL ، (2) اذا كان المرض سرطان الدم النخاعى الحاد AML، (3) اذا كان المرض سرطان الدم اللمفاوى المزمن CLL ، (4) اذا كان المرض سرطان الدم النخاعى المزمن CML
X_6 :نوع العلاج ويأخذ القيمة (1) اذا كان العلاج حيوى ، (2) اذا كان العلاج كيماوى
X_7 :النوع ويأخذ القيمة (1) اذا كان المريض يعانى من فقر الدم ، (0) اذا كان المريض يعانى من فقر الدم.

●تقدير نموذج انحدار كوكس:

يفترض نموذج كوكس ان دالة المخاطرة لمرضى سرطان الدم تتأثر ب(7) متغيرات وهي العمر والجنس والمهنة والعنوان وحالة فقر الدم ونوع المرض ونوع العلاج وعلى وقت البقاء والتي تاخذ الصيغة التالية:

$$\boldsymbol{h(t,x) = h_{\circ}(t).exp(Bx)}$$

حيث ان :

$h_o(t)$: وهي دالة الخطر الأساسية للمرضى المصابين بسرطان الدم عندما يكون جميع قيم $x = 0$.

B: تمثل متجه صفي (6x1) لمعلمات الانحدار غير المعلومة.

x: تمثل متجه عمودي (1x6) من المتغيرات المستقلة المصاحبة للمرضى قيد الدراسة.

$exp(B'x)$: هي دالة المخاطرة النسبية، وتصف تغير المخاطر مع تغير قيم المتغيرات المستقلة أي ليس لها علاقة بالزمن.

وبأستخدام البرنامج الاحصائي SPSS يتم تقدير معلمات نموذج الانحدار كوكس بطريقة الامكان الاعظم "Partial Likelihood"، وبأستخدام طريقة نيوتن رافسون والاعتماد على معادلة رقم (3) نحصل على النتائج التالية:

جدول رقم (1)

تقديرات نموذج انحدار كوكس (COX)

المتغيرات المستقلة	B	SE	Wald	Df	Sig.	Exp(B)
العمر	.005	.025	.041	1	.839	1.005
النوع	.082	.785	.011	1	.917	1.086
العنوان	-2.272	1.220	3.470	1	.062	.103
المهنة	.266	.269	.981	1	.322	1.305
نوع المرض	.023	.462	.003	1	.960	1.023
نوع العلاج	1.553	.732	4.506	1	.034	4.727
فقر الدم	-.786	.715	1.207	1	.272	.456

يعرض جدول رقم (1) تقديرات نموذج انحدار كوكس (COX)، ويتضح من الجدول معنوية نوع العلاج عند مستوى معنوية 5% حيث بلغت قيمة احصاء (Wald = 4.506) و (Sig. =.034) وهى أقل من 5% ، وكذلك معنوية العنوان عند مستوى معنوية 10% حيث بلغت قيمة احصاء (Wald = 3.470) و (Sig. =.062) وهى أقل من 10% ، ويتضح من الجدول عدم معنوية باقى المتغيرات (العمر، النوع، المهنة، نوع المرض، فقر الدم) عند مستوى معنوية 5% حيث ان قيمة .Sig أكبر من مستوى المعنوية، اى نستنتج ان العنوان ونوع العلاج يؤثران فى زمن البقاء لمرضى سرطان الدم.

وحيث أن إشارة معامل نوع العلاج موجوبة (1.553) فإن زيادة المتغير المستقل بمقدار وحده واحده والتى تعنى من الانتقال الحيوى الى العلاج الكيماوى

سوف يؤدى إلى زيادة المخاطرة وان حالة المريض تتجه نحو السوء بمقدار (exp (1.553))، واشارة العنوان سالبة (2.272-) فإن انخفاض المتغير المستقل وحده واحده والتى تعنى من الانتقال داخل بغداد الى خارج بغداد سوف يؤدى إلى زيادة المخاطرة وان حالة المريض تتجه نحو السوء بمقدار ((2.272) exp).

- **تقدير الدالة التجميعية لمعدل المخاطرة لنموذج كوكس:**

وبأستخدام معادلة رقم (4)، يتم ايجاد تقدير قيم معدل المخاطرة الاساس والحصول على النتائج التالية:

جدول رقم (2)

تقدير الدالة التجميعية لمعدل المخاطرة لنموذج انحدار كوكس

At mean of covariates			Baseline Cum Hazard	Time
Cum Hazard	SE	Survival		
.007	.008	.993	.001	10
.014	.012	.986	.002	13
.022	.017	.978	.003	14
.030	.021	.970	.004	16
.040	.026	.961	.006	20
.054	.032	.948	.008	50
.070	.039	.933	.010	55
.090	.047	.914	.013	57
.112	.056	.894	.016	68
.160	.074	.852	.023	96

يعرض جدول رقم (2) تقدير الدالة التجميعية لمعدل المخاطرة لنموذج انحدار كوكس، ويتضح من الجدول ان احتمال البقاء حتى (57) يوما تبلغ (0.914) مع خطورة تراكمية (0.090) بينما عند (68) يوما فإن احتمال البقاء تنخفض إلى (.894) مع زيادة الخطورة التراكمية الى (0.112).

- **تحديد أهم المتغيرات التى تؤثر فى زمن البقاء على قيد الحياة:**

يتم هذا الاختبار بطريقة Backward وبالاعتماد على أحصاءة (Wald)، ويعرض جدول رقم (3) تقديرات نماذج انحدار كوكس بطريقة Backward للمتغيرات التى تؤثر فى زمن البقاء كالاتي:

جدول رقم (3)

تقديرات نماذج انحدار كوكس بطريقة Backward للمتغيرات التى تؤثر فى زمن البقاء

		B	SE	Wald	Df	Sig.	Exp(B)
Step 1	العمر	.005	.025	.041	1	.839	1.005
	النوع	.082	.785	.011	1	.917	1.086
	العنوان	-2.272	1.220	3.470	1	.062	.103
	المهنة	.266	.269	.981	1	.322	1.305
	نوع المرض	.023	.462	.003	1	.960	1.023
	نوع العلاج	1.553	.732	4.506	1	.034	4.727
	فقر الدم	-.786	.715	1.207	1	.272	.456
Step 2	العمر	.005	.025	.047	1	.828	1.005
	النوع	.084	.786	.011	1	.915	1.087
	العنوان	-2.263	1.206	3.522	1	.061	.104
	المهنة	.266	.269	.977	1	.323	1.304
	نوع العلاج	1.553	.732	4.507	1	.034	4.727
	فقر الدم	-.788	.715	1.214	1	.271	.455
Step 3	العمر	.006	.024	.056	1	.812	1.006
	العنوان	-2.298	1.160	3.924	1	.048	.100
	المهنة	.281	.226	1.540	1	.215	1.324
	نوع العلاج	1.551	.731	4.503	1	.034	4.718
	فقر الدم	-.794	.712	1.242	1	.265	.452
Step 4	العنوان	-2.268	1.156	3.851	1	.050	.104
	المهنة	.284	.227	1.564	1	.211	1.329
	نوع العلاج	1.519	.713	4.536	1	.033	4.568
	فقر الدم	-.790	.715	1.219	1	.269	.454
Step 5	العنوان	-2.536	1.150	4.864	1	.027	.079
	المهنة	.348	.228	2.332	1	.127	1.417
	نوع العلاج	1.512	.718	4.432	1	.035	4.535
Step 6	العنوان	-1.976	1.061	3.470	1	.062	.139
	نوع العلاج	1.355	.695	3.800	1	.051	3.878

ويتضح من الجدول بمقارنة قيم عمود (Wald) بعمود (.Sig) فى المرحلة الأولى يتضح معنوية كل من العنوان ونوع العلاج وعدم معنوية باقى المتغيرات، فى المرحلة الثانية تم حذف نوع المرض، وفى الخطوة الثالثة تم حذف متغير النوع، وفى الخطوة الرابع تم حذف العمر، وفى الخطوة الخامسة تم حذف فقر الدم، وفى الخطوة السادسة والأخير تم حذف المهنة، ويتضح معنوية العنوان

ونوع العلاج وهما أهم متغيرين يؤثران قى زمن البقاء بالايام لمرضى سرطان الدم.وبذلك يكون النموذج الملائم هو:

$$Y = -1.976\ x_3 + 1.355\ x_6$$

- **تحديد أفضل نموذج:**

يمكن تحديد أفضل نموذج بأستخدام أختبار نسبة دالة الامكان" Likelihood Ratio Test"، حيث ان اختبار نسبة الامكان ليس بالضرورة أن يكون الاختبار المنظم الاكثر قوة، لكن ثبت في الؤلفات أن مثل هذا الاختبار في كثير من الاحيان له خصائص مرغوب فيها (البلقيني 2016).
وتكون أحصاءة الاختبار بالشكل الاتي:

$$G = 2\left[\ln L\left(\hat{B}\right) - \ln L\left(\hat{B}^*\right)\right] \quad \ldots\ldots.(3.10)$$

علما ان:

$$\hat{B}^* = \left(\hat{B}_1, \hat{B}_2, \ldots\ldots, \hat{B}_q\right);\ 1 \le q \le p$$

$L\left(\hat{B}\right)$:هو دالة الامكان الاعظم عندما تكون دالة المخاطرة الاساسية موجودة فقط في النموذج.

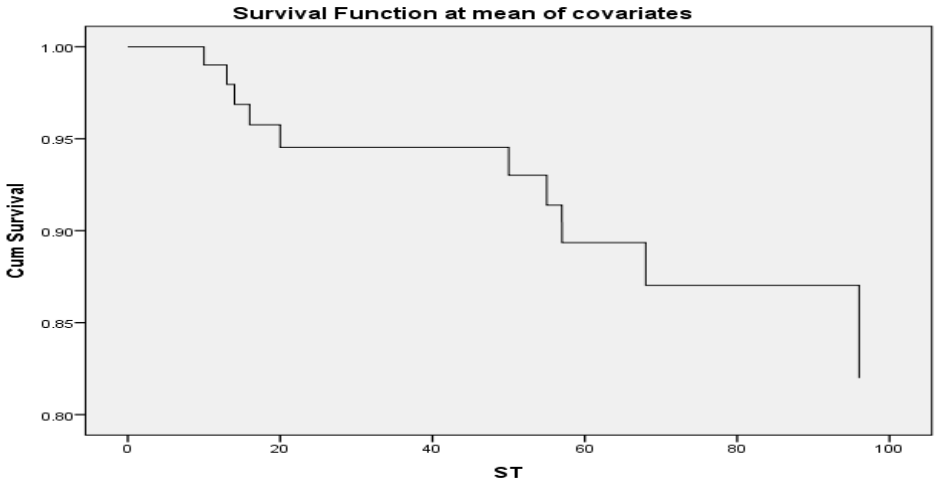
$L\left(\hat{B}^*\right)$: هودالة الامكان الاعظم عند وجود عدد من المتغيرات داخل النموذج.
وكما يتضح من شكل الاختبار، فانه يقوم على مقارنة الدالتين وتحديد اذا كان هناك فرق معنوي بين الدالتين، وان نتيجة أحصاءة الاختبار تتبع توزيع مربع كاي بدرجة حرية مساوية لعدد المعلمات**:**

يعرض جدول رقم (4) نتائج اختبار أفضل نموذج، ويتضح من الجدول ان النموذج السادس هو الاكثر معنوية بين النماذج المقدرة حيث انه ذو اقل قيمةٍ Sig. (0.019) والذى يحتوى على متغيرين مستقليين (العنوان ونوع العلاج).

جدول رقم (4)
نتائج اختبار أفضل نموذج

Sig.	Df	Chi-square	-2 Log Likelihood	Step
.154	7	10.666	60.337	Step 1
.104	6	10.537	60.339	Step 2
.069	5	10.214	60.351	Step 3
.038	4	10.125	60.406	Step 4
.031	3	8.889	61.712	Step 5
.019	2	7.882	64.008	Step 6

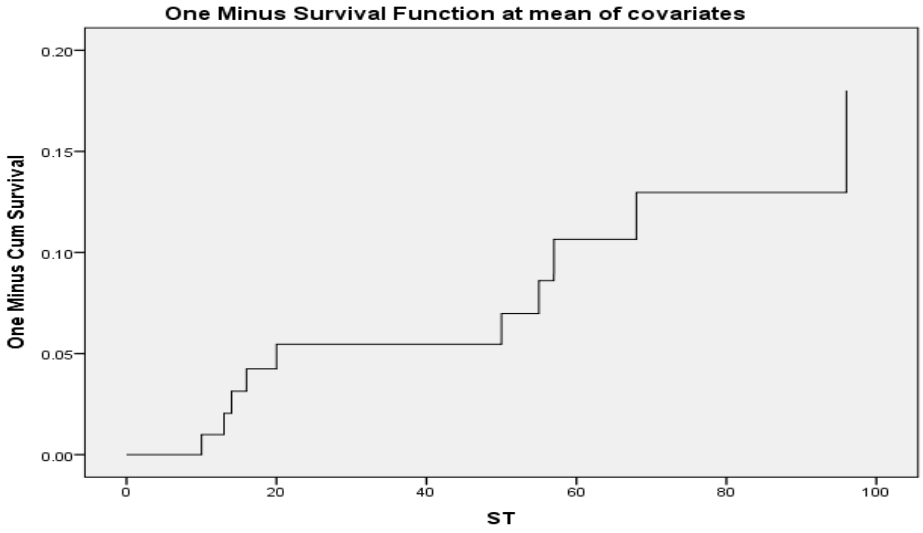
يعرض شكل رقم (1) دالة البقاء عند متوسط المتغيرات التفسيرية، ويتضح من الجدول ان دالة البقاء تبدء متناقصة الى ان تستقر عند الفترة (40 - 20) وتبدء فى التناقص وحت تستقر مرة أخرى عند الفترة (85-65).



شكل رقم (1)

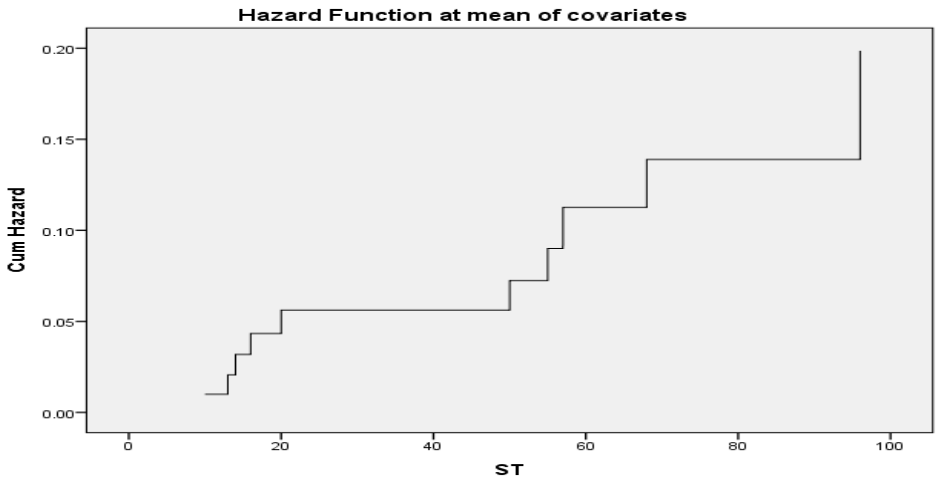
دالة البقاء عند متوسط المتغيرات التفسيرية

يعرض شكل رقم (2) احتمال عدم البقاء بدلالة متوسط المتغيرات التفسيرية، ويتضح من الشكل ان دالة عدم البقاء تزايدية ولكنها تستقر على فترات ، وهذا مايتوافق مع دالة الخطر الذى يوضحها شكل رقم (3).



شكل رقم (2)

احتمال عدم البقاء بدلالة متوسط المتغيرات التفسيرية



شكل رقم (3)
دالة المخاطرة بدلالة متوسط المتغيرات التفسيرية

النتائج:

يمكن تلخيص النتائج النهائية التي تم التوصل اليها من خلال هذا البحث في النقاط التالية:

1- من خلال أختبار Backward يظهر لنا أن متغير العنوان ونوع العلاج هما المتغيران المؤثران في وقت البقاء، وهما المرشحان للبقاء في النموذج، بينما المتغيرات الاخرى وهي النوع والجنس والمهنة وحالة فقر الدم ونوع السرطان لم تظهر اي تأثير معنوي على وقت البقاء.

2- اشارة نوع العلاج موجبة فأن زيادة المتغير المستقل بوحدة واحدة والتي تعني الانتقال من العلاج الحيوي الى الكيميائي سيؤدي الى زيادة المخاطرة .

3- زيادة الخطورة مع تقدم الزمن.

التوصيات:

1. اجراء بحث شامل يضم بيانات لمدة لاتقل عن 2 سنة وأضافة معلومات جديدة الى بيانات المصابين بهذا المرض مثل اذا كان لديه امراض مزمنة اخرى، وهل المريض مدخن ام لا، لاجراء دراسات حول هذا المرض ومعرفة التأثير الحقيقي للمتغيرات حول هذا المرض.

2. اجراء دراسات متعلقة بالبقاء وتطبيق نموذج كوكس على بيانات باقي أنواع السرطانات الاخرى، ومعرفة العوامل المؤثرة على كل نوع من تلك الانواع.

3. وضع كوادر أحصائية متخصصة وانشاء قاعدة بيانات في المستشفيات والمراكز الصحية لتسجيل البيانات بشكل دقيق، لان البيانات هي الاساس في الدراسات والبحوث.

4. متابعة مخاطر الوفيات بسبب مرض السرطان، وفتح مراكز جديدة بهدف رصد و تسجيل حالات مرضية للسرطان في جميع المناطق ومعالجة المرض والحد من أنتشاره.
5. اجراء حملات توعية حول خطورة هذا المرض، والعوامل الرئيسية المسببة لهذا المرض.

المصادر:

1 .البلقيني، محمد توفيق. (2016). دراسات متقدمة في الاحصاء الرياضي. كتاب. كلية التجارة. جامعة المنصورة.

2 .عزيز، ئـــةڕى أكرم. (2011). أستخدام نموذج cox للانحدار في تحليل البقاء لمرضى اللوكيميا في مستشفى نانكلي- اربيل. رسالة ماجستير في الاحصاء التطبيقي. جامعة السليمانية.

3 .Cox, D. R. (1972). Regression model and life table, with discussion, J. R. S. S., B,34,2: 187-220.

4 .Cox, D. R. (1975). Partial likelihood, biometric, 62, 2: 269-276.

5 .Crowder, M. J., Kimber, A. C., Smith, R. L., & Sweeting, T. J. (1991). Statistical analysis of reliability data, chapman & Hall, New York.

6 .Efron, B. (1977). The efficiency of Coxs likelihood function for censored data. J. Amer. Statist. Assoc. 72: 557-565.

7 .Fox, Y. (2002). Cox proportional hazard regression for survival data. University Sheffield.