تحديد أهم العوامل المؤثرة على مرض سرطان الكبد في العراق للعام 2017

م.د. غياث حميد مجيد
ghiath.hameed@gmail.com
الجامعة المستنصرية / كلية التربية الاساسية

الملخص

يعد الانحدار اللوجستي أحد أهم عناصر النمذجة الخطية وهو من أكثر الأساليب الاحصائية استخداماً في مجال تحليل البيانات المصنفة ولا يفترض أي شروط حول توزيع المتغيرات المستقلة. يستخدم الانحدار اللوجستي لأغراض التنبؤ باحتمالية وقوع حدث معين ويتم استخدامه في الكثير من المجالات المختلفة منها الطبية والاقتصادية والعلوم الاجتماعية وغيرها من المجالات المختلفة التي تتطلب توقع ميل الأشخاص في اتخاذ القرار حول الدراسة قيد البحث. ويتسم هذا النوع من أنواع الانحدار بالمرونة إذ يعطي تفسيراً واضحاً لوصف لعلاقة ما بين المتغير التابع والمتغيرات التوضيحية.

تناول البحث استخدام الانحدار اللوجستي لتحديد أهم العوامل والمتغيرات المؤثرة على المصابين بمرض سرطان الكبد في العراق للعام 2017، وذلك بتحليل العلاقة بين المتغيرات والوصول إلى المتغيرات التي تؤثر على الاورام السرطانية الخاصة بهذا النوع من المرض. ومن أهم الاستنتاجات التي تم التوصل إليها من الباحث التعرّف على أهم العوامل المؤثرة على مرض سرطان الكبد والمسببة للمرض لغرض تجنب الاصابة بمثل هكذا نوع من الأمراض.

المصطلحات الرئيسة / الانحدار الخطي، الانحدار اللوجستي، ثنائي الاستجابة، دالة الامكان الأعظم.

تحديد أهم العوامل المؤثرة على مرض سرطان الكبد في العراق للعام 2017

م.د. غياث حميد مجيد

مقدمة

إن الانحدار هو أسلوب إحصائي يُعنى بدراسة العلاقة بين متغير تابع ومتغير أو عدة متغيرات توضيحية ينتج عن ذلك معادلة رياضية تمثل تلك العلاقة بين المتغيرات أفضل تمثيل، ويعد الانحدار اللوجستي حالة خاصة من الانموذج الخطي العام والذي يعد أكثر أنواع النماذج شيوعاً في تحليل البيانات الوصفية.[4]

ويعد الارتقاء بالواقع الصحي في العراق هو أحد أهم مقومات قيام المجتمع، إذ لوحظ في المرحلة الأخيرة ارتفاع كبير في نسب الاصابة بمرض السرطان بشكل عام ومرضى سرطان الكبد بشكل خاص مما جعله وضعاً مثيراً للجدل لتحديد والوقوف على أهم العوامل المؤثرة والتي أدت إلى هذا الارتفاع. تعد عملية التعرف على أهم الاسباب والعوامل الرئيسة المؤثرة على المصابين بمرض سرطان الكبد أحد اهـم المؤشرات المهمة للنهوض بالواقع الصحي في العراق لما يمثله من ثروة وطنية يساهم بالنهوض بالواقع الصحي للفرد العراقي. ومـن أجل ذلك تطرق الباحث إلى التعرف على أهم المحددات المعنوية المؤثرة بأستخدام تقنـية الانحدار اللوجستي ثنائي الاستـجابة (Binary logistic regression) والذي يستخدم في حالة المتغير التابع والمؤلف من مستويين.[2,3]

إن من أهم ما يميز هذا البحث هو التركيز على استخدام تقنية الانحدار اللوجستي في المجال الصحي، إذ تعد هذه التقنية من أكثر التقنيات شيوعاً وتستخدم بشكل كبير في كثير من المجالات المختلفة منها الطبية والاجتماعية والاقتصادية بالنظر إلى خصوصية البيانات في هذه المجالات والتي غالباً ما تكون بياناتها من النوع الثنائي، فضلاً عن أن تقديرات معالم الانموذج اللوجستي تعد مقبولة في ظل غياب بعض القيود المفروضة على الأنواع الأخرى من نماذج الانحدار مثل الانحدار الخطي والانحدار اللوغاريتمي.

1. مشكلة البحث

تم رصد العديد من حالات الاصابة بالأورام السرطانية في العراق خلال العقد الأخير، والتي توحي بالتدهور الصحي الكبير للأسرة العراقية، الأمر الذي أدى إلى انتشار هذا النوع من الأمراض بين تلك الأسر، وتناولت العديد من الدراسات اتجاهات المستويات الاقتصادية والاجتماعية التي أثرت على واقع الاصابة بهذا النوع من الأمراض. لذا ارتأى الباحث أن يقوم بإجراء هذه الدراسة من أجل تحديد أهم العوامل المؤثرة على مرض سرطان الكبد في العراق في ظل تدهور الواقع الصحي في العراق.

2. هدف البحث

يهدف البحث إلى تسليط الضوء على أهم العوامل المؤثرة للاصابة بمرض سرطان الكبد والتطرق إلى شرح موجز الى ماهو معروف عن طبيعة هذا المرض بشكل عام من خلال استعمال أنموذج الانحدار اللوجستي الثنائي الاستجابة.

3. الانحدار اللوجستي Logistic Regression

يعد تحليل الانحدار أحد أهم الأساليب الاحصائية التي تستخدم في تحليل البيانات الفئوية في مجالات البحوث المختلفة إذ تستخدم لوصف العلاقة بين متغير معتمد ومتغير أو أكثر من المتغيرات التوضيحية. وتقسم نماذج الانحدار إلى نوعين رئيسين الأول نماذج الانحدار الخطية والثاني اللاخطية، إذ يتم استخدام النوع الثاني من نماذج الانحدار في العديد من الظواهر الطبيعية والتي عند دراسة سلوكها نجد أنها تسلك سلوكاً غير خطي لوصف وتحليل تلك الظواهر. ولكن وبسبب صعوبة استخدام نماذج الانحدار اللاخطية فأن غالباً ما يلجأ الباحثون إلى استخدام نماذج الانحدار الخطية

وتحديداً انموذج الانحدار اللوجستي الذي يعد واحداً من أهم النماذج التي تستخدم لوصف وتحليل تلك الظاهر.[1,5]

بالنظر لكون انموذج الانحدار اللوجستي يهتم بدراسة وتحليل البيانات ذات الاستجابة الثنائية والتي يكون فيها المتغير التابع ثنائياً، يأخذ متغير الاستجابة القيمة (1) في حالة النجاح والقيمة (0) في حالة الفشل، أي أن يتبع هذا المتغير توزيع برنولي. يستخدم أنموذج الانحدار اللوجستي في وصف العلاقة بين متغير الاستجابة (التابع) ومتغير أو مجموعة من المتغيرات التوضيحية (المستقلة)، ويتم تمثيل تلك العلاقة بالصيغة الآتية: [5]

$$P(x_i) = \frac{1}{1+e^{-\alpha-\beta x_i}} \qquad -\infty \le \alpha \le \infty \ , \quad \beta > 0 \qquad \dots (1)$$

إذ أن: $P(x_i)$ تمثل احتمال الاستجابة للمفردة i

و x_i يمثل المتغير المستقل

α , β معلمات الانموذج

تُعرف الصيغة أعلاه $P(x_i)$ بأنها دالة الاستجابة اللوجستية والتي تفترض جملة من الافتراضات منها أن تكون قيم المتغير المعتمد ما بين (1,0) وأن يتوزع توزيع برنولي، وأن تكون حجوم العينات كبيرة، ولا يشترط أن تكون العلاقة خطية بين المتغير التابع والمستقل، وأن لا يوجد ارتباط قوي بين المتغيرات التوضيحية (المستقلة)، وأن تكون معلمتا الانموذج (α , β) غير مقيدتين.

بشكل عام هناك نوعان من نماذج الانحدار اللوجستي، النوع الأول يعرف بإنموذج الانحدار اللوجستي ثنائي الاستجابة والثاني يعرف بإنموذج الانحدار اللوجستي متعدد الاستجابة.

1.4 الانحدار اللوجستي ثنائي الاستجابة Binary Logistic Regression [4,6]

يتم بناء إنموذج الانحدار اللوجستي ثنائي الاستجابة على افتراض أساسي ألا وهو أن يكون متغير الاستجابة (المتغير التابع) من النوع الثنائي والذي يتبع توزيع برنولي، إذ يأخذ هذ المتغير القيمة (1)، أي حدوث الاستجابة، بإحتمال (p)، والقيمة (0)، أي عدم حدوث الاستجابة، بإحتمال (q=1-p). إن انموذج الانحدار اللوجستي عبرة عن تحويل لوغاريتمي للانحدار الخطي وباستخدام خصائص التوزيع اللوجستي نقوم بتقييد الاحتمالات المقدرة لتكون محصورة بين (1,0) بدلاً عن القيم ($-\infty$, ∞) في حالة الانحدار الخطي. أي أنه يتم تحويل دالة الانحدار اللوجستي إلى دالة خطية بأخذ اللوغاريتم للطرفين.

تعد طريقة الإمكان الأعظم من أكثر الطرق شيوعاً واستخداماً في تقدير معلمات الانموذج اللوجستي لتقدير معلمات النماذج غير الخطية، وتعد أفضل من طريقة المربعات الصغرى العادية المستخدمة لتقدير معلمات النماذج الخطية كونها لا تفترض قيوداً كثيرة مثل الخطية وتجانس التباين وهي تقيس الاحتمالات المشاهدة لعدد من المتغيرات المستقلة في العينة.[3]

ولغرض تقدير معلمات انموذج الانحدار اللوجستي ثنائي الاستجابة باستخدام طريقة الإمكان الأعظم التي تتضمن ايجاد قيم معلمات الانموذج التي تعظم دالة الإمكان الأعظم، نفترض لدينا k من المتغيرات العشوائية المستقلة ($X_1, X_2, \dots, X_k$) والتي تتوزع بدورها توزيعاً ثنائي الحدين بالمعالم (n_i , p_i)، إذ أن (X_i) تمثل مجموع حالات النجاح في كل محاولة من (n_i) وان هناك (r) من المتغيرات المستقلة في كل مجموعة من المجاميع، لذا فإن دالة الكثافة الاحتمالية لـ (X_i) تمثل بالشكل:

$$P_i(X_i = x_i) = C_{x_i}^{n_i} P_i^{x_i} (1-P_i)^{(n_i - x_i)} \quad , \quad i = 1,2,...,k \qquad \dots (2)$$

إذ أن P_i تمثل نسبة الاستجابة (احتمال النجاح) لتوزيع ثنائي الحدين والتي يتم تقديرها على النحو الآتي:

$$P_i = \frac{x_i}{n_i} \qquad ...(3)$$

أما بالنسبة إلى $Q_i=1-P_i$ والتي تمثل احتمال عدم حدوث الاستجابة (احتمال الفشل) لنفس التوزيع المذكور فيتم تقديرها على النحو الآتي:

$$Q_i = 1 - \frac{x_i}{n_i} = \frac{n_i - x_i}{n_i} \qquad \dots(4)$$

وبالتالي يتم تمثيل دالة الإمكان الأعظم للتوزيع المشترك للمفردات (x_i) بالشكل الآتي:

$$L_{(p)} = \prod_{i=1}^{k} C_{x_i}^{n_i} \; P_i^{x_i} \; (1-P_i)^{(n_i - x_i)} \qquad \dots(5)$$

وعند أخذ اللوغاريتم للأساس الطبيعي لدالة الإمكان الأعظم والموضحة في الصيغة رقم (5) يكون لدينا:

$$Ln\, L_{(p)} = \sum_{i=1}^{k} \left[Ln\, C_{x_i}^{n_i} + x_i Ln\left(\frac{p_i}{1-p_i}\right) + n_i \; Ln\,(1-p_i) \right] \qquad \dots(6)$$

وحيث أن كلاً من (p_i) و($q_i=1-p_i$) اللذين يمثلان كلاً من احتمال حدوث الاستجابة واحتمال عدم حدوثها على التوالي لأنموذج اللوجستي واللذين يتأثران بالمتغيرات المستقلة بحسب الإنموذج المدروس، فأن كلاً من هذين الاحتمالين توضح صيغهما الحسابية على النحو الآتي:

$$p_i = \frac{\exp\,(X_i'B)}{1+\exp\,(X_i'B)} \quad , \quad q_i = 1 - p_i = \frac{1}{1+\exp\,(X_i'B)}$$

وعند تعويض قيمتي هذين الاحتمالين في الصيغة رقم (6) نحصل على:

$$Ln\, L_{(p)} = \sum_{i=1}^{k} \left[Ln\, C_{x_i}^{n_i} + x_i \; Ln \left(\frac{\dfrac{\exp\,(X_i'B)}{1+\exp\,(X_i'B)}}{\dfrac{1}{1+\exp\,(X_i'B)}} \right) + n_i \; Ln \left(\frac{1}{1+\exp\,(X_i'B)} \right) \right] \qquad \dots(7)$$

وبتبسيط الصيغة رقم (7) نحصل على الصيغة الآتية:

$$Ln\, L_{(p)} = \sum_{i=1}^{k} \left[Ln\, C_{x_i}^{n_i} + x_i \; (X_i'B) + n_i \; Ln \left(\frac{1}{1+\exp\,(X_i'B)} \right) \right] \qquad \dots(8)$$

وبالنظر إلى صعوبة حل الصيغتين أعلاه بشكل آني ولغرض الحصول على تقدير معلمات الأنموذج نلجأ الى استخدام إحدى طرائق التحليل العددي وهي طريقة نيوتن رافسون (Newton -Raphson). وباستعمال هذه الطريقة وتقدير معلمات الأنموذج نحصل على الأنموذج بصيغته النهائية والذي يكون على النحور الآتي:[2,5]

$$\underline{m}_{s+1} = \underline{m}_s - D^{-1}\,\underline{d}_{(s)} \qquad \ldots(9)$$

إذ أن: $\underline{m}_{s+1}$ تمثل متجه معلمات الأنموذج المراد تقديرها B

$\underline{m}_s$ تمثل متجه القيم الأولية لمعلمات الأنموذج

D تمثل مصفوفة المشتقة الثانية للوغاريتم دالة الإمكان الأعظم

$\underline{d}_{(s)}$ تمثل متجه المشتة الأولى للوغاريتم دالة الإمكان الأعظم.

ولغرض بيان أهية العوامل المؤثرة في إنموذج الانحدار اللوجستي نستخدم مجموعة من الاختبارات منها اختبار Wald-Wolfowitz runs والذي يتوزع توزيع مربع كاي Chi-Square والذي يكون بدوره معنوياً إذا كانت القيمة الاحتمالية لإحصاءة هذا الاختبار أقل من مستوى المعنوية الذي يتم تحديده مسبقاً من قبل الباحث. وأيضاً يمكن استخدام اختبار Hosmer & Lemeshow والذي يستخدم لمعرفة مدى تمثيل الإنموذج للبيانات المستخدمة في البحث، وهو أحد الاختبارات التي تعتمد على اختبار مربع كاي لحسن المطابقة إذ يتم تقييم الفروق بين القيم المشاهدة والمتوقعة فإذا كانت قيمة إحصاءة مربع كاي الاحتمالية أكبر من مستوى المعنوية المحدد فإن الإنموذج لا يمثل البيانات بصورة جيدة.[4]

4. الجانب التطبيقي

تم أخذ البيانات الخاصة بالمرض المذكور من شعبة الإحصاء في مركز السرطان والتابع إلى وزارة الصحة العراقية للعام 2017 إذ بلغ عدد المصابين بالمرض (393) مصاباً بواقع (202) مصاب من الذكور و(191) من الاناث، وقد تم استخدام برنامج التحليل الإحصائي SPSS الاصدار 23 منه. الجدول الآتي يبين وصف المتغيرات الداخلة في عملية التحليل الإحصائي:

جدول (1) يوضح المتغيرات الداخلة في التحليل

اسم المتغير	وصف المتغير	نوع المتغير
Gender	متغير الجنس (ذكر، أنثى)	مستقل
Age	عمر المريض	مستقل
AddrCode	المحافظة التي يسكنها المريض	مستقل
Occup	مهنة المريض	مستقل
Extent	طبيعة انتشار المرض	مستقل
Radio	العلاج الإشعاعي	مستقل
Chemo	العلاج الكيمياوي	مستقل
Hormon	نوع الورم (عادي، هرموني)	مستقل
Immuno	مناعة الجسم	مستقل
Status	حالة المصاب بالمرض (حي، ميت)	معتمد

وبعد عملية ادخال ابيانات في النظام تم استخلاص النتائج الآتية موضحة في الجداول الآتية:

جدول (2) يبين نتائج التحليل بالنسبة إلى متغير الجنس

المتغير	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة
الجنس	0.299	0.518	0.333	1	0.564	1.349

من الجدول أعلاه نلاحظ أن القيمة الاحتمالية أكبر من مستوى الدلالة للاختبار والبالغة (0.05) أي أنه لا يوجد تأثير لمتغير الجنس على حالة المصاب بالمرض.

جدول (3) يبين نتائج التحليل بالنسبة إلى متغير العمر

المتغير	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة
العمر	-0.024	0.009	6.514	1	0.011	0.976

أما هنا فنلاحظ أن القيمة الاحتمالية أصغر من (0.05) وهذا يعني وجود تأثير للعمر على حالة المصاب بالمرض.

جدول (4) يبين نتائج التحليل بالنسبة إلى متغير المحافظة

متغير المحافظة	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
بغداد	0.498	1.218	0.167	1	0.683	1.646	0.007
نينوى	36.784	8121.326	0.000	1	0.996	9441666101069880	
البصرة	-0.444	1.436	0.096	1	0.757	0.641	
ذي قار	21.413	15515.862	0.000	1	0.999	1994104099.059	
بابل	-2.478	1.415	3.067	1	0.080	0.084	
أربيل	22.265	40192.970	0.000	1	1.000	4670555805.421	
الأنبار	1.116	1.617	0.476	1	0.490	3.052	
ديالى	3.016	1.963	2.360	1	0.125	20.401	
صلاح الدين	0.093	1.813	0.003	1	0.959	1.097	
النجف	0.736	1.274	0.334	1	0.563	2.088	
واسط	3.000	1.439	4.345	1	0.037	20.094	
القادسية	22.345	10616.536	0.000	1	0.998	5062465974.502	
كربلاء	0.285	1.810	0.025	1	0.875	1.330	
دهوك	17.987	28292.836	0.000	1	0.999	64837088.934	
ميسان	-1.206	1.640	0.540	1	0.462	0.299	
المثنى	2.683	1.561	2.956	1	0.086	14.634	

نلاحظ في هذا الجدول بأن القيمة الاحتمالية الكلية لمتغير المحافظة بشكل عام قد بلغت (0.007) والتي بدورها تكون أقل من (0.05) وبالتالي فإن هناك دور كبير لهذا المتغير في احتمال وفاة المصاب من عدمه. ونلاحظ أيضاً أن المحافظات التي كان لها تأثير مباشر هما محافظتا (واسط والمثنى).

جدول (5) يبين نتائج التحليل بالنسبة إلى متغير المهنة

متغير المهنة	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
ربة بيت	-1.247	.492	6.416	1	.011	.287	
عامل في مكتب	-3.056	1.266	5.824	1	.016	.047	
مزارع	17.388	40192.970	.000	1	1.000	35600115.659	
موظف	-1.431	.809	3.131	1	.077	.239	0.004
معلم	15.523	40192.970	.000	1	1.000	5515510.235	
مهني	-1.521	2.181	.486	1	.486	.219	
كاسب	-2.293	.650	12.446	1	.000	.101	
أخرون	-2.523	.633	15.899	1	.000	.080	

وفي ما يتعلق بمتغير المهنة نلاحظ أن القيمة الاحتمالية الكلية لمتغير المهنة بشكل عام قد بلغت (0.004) والتي بدورها تكون أقل من (0.05) وبالتالي فإن هناك دوراً كبيراً لهذا المتغير في حالة الاصابة بالمرض. ونلاحظ أيضاً أن المهن التي كان لها تأثير مباشر هي كل من (ربة البيت، عامل في مكتب، كاسب، آخرون).

جدول (6) يبين نتائج التحليل بالنسبة إلى متغير مستوى انتشار المرض

متغير مستوى انتشار المرض	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
انتشار بعيد للورم	1.838	.794	5.360	1	.021	6.283	
موضعي	.945	.697	1.839	1	.175	2.572	
حول الورم فقط	.795	1.300	.374	1	.541	2.215	0.006
وصل للعقد اللمفاوية	1.298	.381	11.580	1	.001	3.661	

أما هنا فمن الواضح إن لهذا المتغير دوراً كبيراً في تحديد حالة المريض، إذ بلغت القيمة الاحتمالية للمتغير ككل (0.006) وهي كما هو واضح أصغر من القيمة الاحتمالية (0.05). وبالتحديد كانت لدينا حالتان أثرت وبشكل واضح وهما (انتشار بعيد للورم، وصول الورم للعقد اللمفاوية).

أما فيما يخص متغيري خضوع المريض لكل من العلاج الاشعاعي والكيماوي فقد كانت نتائجهما موضحة في الجدولين الآتيين:

جدول (7) يبين نتائج التحليل بالنسبة إلى متغير خضوع المريض للعلاج الاشعاعي

متغير الخضوع للعلاج الاشعاعي	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
نعم	16.531	6033.589	.000	1	.998	15114362.172	0.003
كلا	19.697	6033.589	.000	1	.997	358216347.257	

جدول (8) يبين نتائج التحليل بالنسبة إلى متغير خضوع المريض للعلاج الكيماوي

متغير الخضوع للعلاج الكيماوي	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
نعم	3.959	1.280	9.568	1	.002	52.407	0.006
كلا	2.914	1.445	4.069	1	.044	18.434	

إذ نلاحظ بأن هذين المتغيرين كان لهما وبشكل عام تأثيراً فعالاً على حالة المريض كون القيمة الاحتمالية لكل منهما أقل من قيمة الاختبار الاحتمالية.

جدول (9) يبين نتائج التحليل بالنسبة إلى متغير طبيعة المرض

متغير طبيعة المرض	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
هرموني	-19.117	27329.498	.000	1	.999	.000	1.000
عادي	-38.649	11228.845	.000	1	.997	.000	

جدول (10) يبين نتائج التحليل بالنسبة إلى متغير العلاج المناعي

متغير العلاج المناعي	قيمة معامل الانحدار	الخطأ المعياري	إحصاءة Wald	درجة الحرية	القيمة الاحتمالية	قيمة معامل الانحدار المتوقعة	القيمة الاحتمالية الكلية لمتغير المحافظة
نجاح	19.335	7946.366	.000	1	.998	249410809.019	0.835
فشل	20.302	7946.366	.000	1	.998	656190382.038	

أما بالنسبة للمتغيرين الأخيرين والموضحين في الجدولين (9) و (10) على التوالي، فنلاحظ أن القيمة الاحتمالية أكبر بكثير من (0.05) وبالتالي فإن هذين المتغيرين ليس لهما أي تأثير يذكر على حالة المصاب بهذا المرض.

5. الاستنتاجات

على ضوء ما توصل إليه الباحث من نتائج في الجانب التطبيقي، توصل إلى أهم الاستنتاجات التي يمكن استخلاصها على النحو الآتي:

1. من خلال تحليل البيانات تبين أن الاصابة بمرض سرطان الكبد لا يعتمد على ثلاثة متغيرات وهي (جنس المريض، طبيعة المرض (عادي، هرموني)، نجاح أو فشل العلاج المناعي). بينما يعتمد على مجموعة من المتغيرات هي (العمر، المحافظة التي يسكن فيها المريض، المهنة التي يزاولها المريض، طبيعة انتشار المرض في الجسم، خضوع المريض للعلاج الاشعاعي، خضوع المريض للعلاج الكيماوي).
2. ضرورة استخدام إنموذج الانحدار اللوجستي الثاني الاستجابة في عملية تحليل البيانات التي يكون فيها المتغير التابع متغيراً تصنيفياً مكوناً من مستويين فقط.
3. استخدام طريقة الإمكان الأعظم في تقدير معلمات إنموذج الانحدار الذي يحوي متغيرات تابعة نوعية لا يؤدي إلى نتائج دقيقة كونه يعاني من بعض المشاكل الإحصائية.
4. الوقوف عند أهم العوامل التي تؤثر على مرضى سرطان الكبد والمسببة للمرض لتجنب الاصابة بمثل هكذا نوع من الأمراض من خلال الوقاية.

6. التوصيات

1. ضرورة تشخيص المرض بشكل مبكر واتخاذ الاجراءات اللازمة للحيلولة دون تفاقمه.
2. ضرورة استحداث وتطوير مراكز الأبحاث الاحصائية في المستشفيات والمراكز الصحية.
3. استخدام الانحدار اللوجستي للمتغيرات ثنائية الاستجابة لمثل هذا النوع من الدراسات، وزيادة عدد المتغيرات لمعرفة تأثيرها على المرض.
4. عمل خطة لتثقيف افراد المجتمع من الناحية الصحية وخصوصاً في القرى ذات المستوى الاجتماعي والاقتصادي المتدني للحد من خطر الاصابة بمثل هذا النوع من الأمراض.

7. المصادر

1. Al-Baldawi, Tansim Hassan, (1996), "***Analytical Comparison Between Logistic Models and Discriminatory Function Models***", PhD thesis, College of Administration and Economics, Baghdad University.

2.Jawad, Abbas Naji, (2010), "***Superlative Between Methods of Estimating Economic Functions with Specific Dependent Variables***", Tikrit Journal of Administrative and Economic Sciences, Volume (6), Issue (18), University of Tikrit.

3. Al-Ta'ee, Abdulhusain Hassan, (2000), "***Estimation and analysis of the multiple regression equation in case the adopted variables are descriptive and specific***", PhD thesis, College of Administration and Economics, Baghdad University.

4. Filia Vonta, Mikhail Nikulin, Nikolaos Limnios, Catherine Huber-Carol, (2008), "***Statistical Models and Methods for Biomedical and Technical System***", Berlin: Birkhauser.

5. Jos Dessens, Wim Janson, (2005), "***Logistic Regression Analysis***", Wiley, New York.

6. Menard, Scott, (2002), "***Applied Logistic Regression Analysis (Quantitative Applications in the Social Sciences)***". 2nd Edition, Volume 106, Beverly Halls, CA: Sage.

تحديد أهم العوامل المؤثرة على مرض سرطان الكبد في العراق للعام 2017

م.د. غياث حميد مجيد

Determine the most important factors affecting liver cancer i n Iraq for the year 2017

Abstract

The logistic regression is one of the most important elements of linear modeling. It is one of the most statistical methods used in the analysis of classified data and does not assume any conditions for the distribution of independent variables. Logistic regression is used to predict the probability of a particular event in many different fields, including medical, economic, social sciences and other fields that require predicting the tendency of people to make decisions about current study. This type of regression is flexible, giving a clear explanation for the relationship between the dependent variable and the explanatory variables.

The study examined the use of logistic regression to determine the most important factors and variables affecting patients with liver cancer in Iraq for the year 2017, by analyzing the relationship between variables and access to variables that affect of this type of cancer. One of the most important conclusions reached by the researcher is to identify the most important factors affecting liver cancer and the cause of the disease in order to avoid the incidence of such a type of disease.

Keywords: Linear Regression, Logistic Regression, Binary Response, Maximum Likelihood Function.