

تقدير معدل تكرار الإصابة بالاحتشاء العضلة القلبية باستخدام نموذج بواسون

م. انتصار مجيد جاسم **

أ. د. سلمى ثابت ذاك *

المستخلص:

تعد أمراض القلب والأوعية الدموية هي السبب الأول للوفيات بين الذكور والإناث على المستوى العالمي وعلى هذا الأساس فقد تركّز هذا البحث على تقدير معدل تكرار الإصابة بالاحتشاء العضلة القلبية خلال مدة زمنية أمدها عام للأشخاص فوق سن (62) وذلك من خلال بناء نموذج إحداد بواسون. حيث أن متغير الإستجابة الذي ينبغي تقديره ببيانات عددية (data a count) في ضوء عدد من المتغيرات التوضيحية المتمثلة بالعمر، الوزن، الجنس، الضغط، السكر، الكوليسترول. وحيث أن تأثير المتغيرات التوضيحية على متغير الاستجابة تأثيرات لا خطية فقد تم استخدام نموذج الإحداد اللوغارتمي الخطي (Log-Linear model) والذي هو حالة خاصة من النموذج الخطي العام (Generalized Linear model) إذ تم إيجاد النموذج المقدر وبيان جودة التوفيق (Goodness of fit) وتنفيذ كافة متطلبات التقدير والتحليل باستخدام برنامج (Statistic version 6). والتوصل أيضاً إلى أن أهم المتغيرات التي تؤدي دوراً أساسياً في معدل الإصابة هما العمر والضغط والسكر وقد تم أيضاً إيجاد الخطورة النسبية للإصابة حسب نوع الجنس والفئة العمرية.

Estimate the recurrence of myocardial infarction model using passion rate

Abstract :

The cardiovascular diseases are the main cause of death among males and females at the global level and on this basis this research focuses to estimate the recurrence of myocardial infarction rate was focused during the period of time for people over the age of 62, through the construction of Poisson regression model. Since the variable response which should be estimate the numerical data in the light of a number of explanatory variables of age, sex, weight, pressing, sugar, cholesterol as the impact of the explanatory variables on the response variable non-linear effects. The logarithmic linear regression model has been used which is a special case of the general linear model and the statement of goodness of fit and implementation of all the appreciation and analysis requirements using a program Statistic aversion 6 . It was also found that the most important variables that play a role in the infection rate are age and pressing, sugar has been also find a relative risk of injury by sex and age group.

* الجامعة المستنصرية / كلية الادارة والاقتصاد .

** الجامعة المستنصرية / كلية الادارة والاقتصاد .

تأريخ استلام البحث 2016/2/19

تأريخ قبول النشر 2016/5/10

1.1 المقدمة:

إن أمراض القلب والأوعية الدموية هي السبب الأول للوفيات بين الذكور والإناث على المستوى العالمي وهي التي من نتائجها الإصابة بنوبات الجلطة القلبية وعلى أقل تقدير فإن 80% من الوفيات ناجمة عن نوبات الجلطة القلبية، وتشمل أمراض شرايين القلب وأمراض شرايين الدماغ والأمراض الروماتيزمية لصمامات القلب وأمراض القلب الناجمة عن ارتفاع ضغط الدم والتهابات القلب. وللمقارنة فإن أعداد الوفيات الناجمة عن أمراض القلب هي ضعف أعداد الوفيات الناجمة عن أمراض السرطان بكافة أنواعه وفي أي مكان بالجسم. وخمسة أضعاف عدد الوفيات الناجمة عن أمراض الجهاز التنفسي وعشرة أضعاف الوفيات الناجمة عن أمراض الجهاز الهضمي وخمسة عشر ضعف الوفيات الناجمة عن أمراض الجهاز البولي.

وعلى هذا الأساس فقد تركّز هذا البحث على بناء نموذج إحداد لتقدير معدل تكرار الإصابة باحتشاء العضلة القلبية للأشخاص خلال فترة زمنية أمدها عام للأشخاص فوق سن (62) سنة، وذلك في ضوء عدد المتغيرات المفترضة التي يُعتقد أن لها تأثير على تكرار الإصابة باحتشاء العضلة القلبية والمتمثلة بالوزن، العمر، الجنس، الضغط، السكر، الكوليسترول، والتي تمثل المتغيرات التوضيحية، وحيث أن متغير الاستجابة أو المتغير المعتمد الذي تهدف إلى تقديره يعبر عنه بيانات عددية (Data a count) والتوزيع الخاص به وكما هو معروف توزيع بواسون لذلك سيتم استخدام نموذج إحداد بواسون لتقدير معدل تكرار احتشاء العضلة القلبية في ضوء المتغيرات التوضيحية المذكورة أعلاه، فضلاً عن بيان أهم المتغيرات التي تؤدي دوراً مهماً في معدل تكرار الإصابة وبيان الخطورة النسبية لها.

1.2 هدف البحث :

يستهدف هذا البحث بناء نموذج إحداد بواسون لتقدير معدل تكرار الإصابة باحتشاء العضلة القلبية للأشخاص بعمر 62 فأكثر وذلك في ضوء عدد من المتغيرات المفترضة والتي تتمثل بالعمر، الضغط، الكوليسترول، السكر، الوزن، الجنس، الوراثة والتدخين، مع بيان المتغيرات الأكثر معنوية. وتأثير تكرار حدوث احتشاء العضلة القلبية عند المصابين (*).

1.3 توزيع بواسون Poisson Distribution

يعد توزيع بواسون من التوزيعات المتقطعة المهمة جداً في الكثير من التطبيقات الإحصائية ويسمى في بعض الأحيان توزيع الحوادث نادرة الوقوع. إن أول من اشتق هذا التوزيع هو العالم الرياضي الفيزيائي الفرنسي (Simon Poisson) وأسس ذلك الاشتقاق على أن التوزيع يمثل حالة تقاربية لتوزيع ثنائي الحدين، ونشر اشتقاقه عام 1873 م كما أطلق اسمه على ذلك التوزيع^[4].

لتفرض أن Y_i متغير عشوائي يشير إلى عدد الأوقات لحصول حدث معين خلال مدة من الزمن، فإن Y_i يتوزع تبعاً لتوزيع بواسون بمعلمة قدرها μ وأن دالة الكثافة الاحتمالية للتوزيع :

$$P_r(Y|\mu) = \frac{e^{-\mu} \mu^y}{y!}, Y_i = 0, 1, 2, \dots$$

إذ إن :-

μ : تمثل معلمة التوزيع وهي ذات قيمة أكبر من الصفر ($\mu > 0$).

1.4 خصائص توزيع بواسون^{[1],[2]}:

إن المعدل أو التوقع لحصول حدث معين خلال مدة من الزمن تمثل معلمة التوزيع وهي بذات الوقت وسطه الحسابي^[4]. أي أن:

$$E(Y) = \mu$$

إن خصائص التوزيع تتمثل بما يأتي :

1. إن قيمة التباين لتوزيع بواسون تكون مساوية إلى قيمة الوسط الحسابي (μ) أي أن:

$$\text{Var}(Y) = E(Y) = \mu$$

(*) هنالك العديد من التطبيقات والدراسات لنموذج إحداد بواسون راجع المصدر (دراسة مقارنة طرائق تقدير نموذج بواسون في حالة وجود مشكلة التعدد الخطي) رسالة دكتوراه للطالب حسام رشيد 2013 ، جامعة بغداد / الإدارة والاقتصاد .

- إذ تعرف هذه الخاصية بـ (Equi dispersion)، وفي التطبيق غالباً ما يكون التباين للمتغيرات المعدودة (Count variables) أكبر من الوسط الحسابي وفي هذه الحالة تدعى الخاصية أعلاه بخاصية فوق التششت (Over dispersion).
2. إن توزيع بواسون يتصف بأنه من التوزيعات الملتوية باتجاه اليمين.
 3. كلما إزدادت قيمة معلمة التوزيع تبدأ قيمة احتمال الصفر بالإخفاض، فإذا كانت قيمة $\mu = 0.8$ فإن احتمال الصفر $P(0) = 0.45$ وإن كانت قيمة $\mu = 2.9$ فإن $P(0) = 0.05$ ، وإن كانت قيمة $\mu = 10.5$ فإن $P(0) = 0.00002$ عندها تنجح كتلة التوزيع لتكون بهيئة الناقوس وهو شكل التوزيع الطبيعي^[3].
 4. في حالة خاصة عندما تكون قيمة المعلمة والتي تمثل الوسط الحسابي والتباين μ مساوية إلى (10.5) فإن توزيع بواسون سيقتررب إلى التوزيع الطبيعي .

1.5 النموذج الخطي المعمم (Generalized Linear model) [5]

النموذج الخطي المعمم هو توسيع أو تعميم لنموذج الإحداد الخطي العام وبالشكل الذي يمكن معه اختبار التأثيرات الخطية واللاخطية فضلاً عن ذلك عندما تكون المتغيرات التوضيحية بقيم مستمرة أو غير مستمرة وكذلك عندما يكون توزيع المتغير المعتمد يتبع توزيع العائلة الأسية إضافة إلى التوزيع الطبيعي.

إن المفهوم الأساسي للنموذج الخطي المعمم (Generalized Linear model) يتلخص بكون هذا النموذج يختلف عن النموذج الخطي العام (Generalized Linear model)، والذي يكون نموذج الإحداد الخطي المتعدد حالة خاصة منه بخاصيتين أساسيتين هما:

1. توزيع المتغير المعتمد من الممكن أن يكون غير طبيعي، أي ممكن أن يكون بواسون (Poisson)، (Gamma)، (binomial)، ... الخ.
2. قيم المتغير المعتمد يتم تقديرها من المركبة الخطية من المتغيرات التوضيحية التي ترتبط مع المتغير المعتمد بواسطة دالة الربط (Link function)، (ومن المعروف بأن دالة الربط تستخدم لنمذجة الاستجابات عندما يفترض أن المتغير المعتمد لا يرتبط خطياً مع المتغيرات التوضيحية).
3. المتغير المعتمد من الممكن أن لا يكون متغيراً مستمراً (Continuous) وهذا يعني من الممكن أن يكون (binomial)، (multinomial) أو (ordinal) ... الخ.

إن النموذج الخطي العام بمتغير منفرد يعد حالة خاصة من النموذج الخطي المعمم ففي النموذج الخطي العام قيم المتغير المعتمد يتبع التوزيع الطبيعي، ودالة الربط هي دالة التطابق البسيطة (Simple identity function) وهذا يعني أن المركبة الخطية من قيم المتغيرات التوضيحية لا يتم تحويلها ولتوضيح ذلك :-

في النموذج الخطي العام المتغير المعتمد Y يرتبط خطياً مع المتغيرات التوضيحية:

$$Y_i = b_0 + b_1 X_1 + \dots + b_k X_k + U$$

أما في النموذج الخطي المعمم فإن العلاقة تفترض بالشكل الآتي:

$$Y = g(b_0 + b_1 X_1 + \dots + b_k X_k) + U$$

حيث أن:

$g(b_0 + b_1 X_1 + \dots + b_k X_k)$ هي دالة وأن معكوسها يمثل دالة الربط (Link function (f) وعليه فإن:

$$f(\mu_Y) = b_0 + b_1 X_1 + \dots + b_k X_k$$

وأن

$$E(Y) = \mu_Y$$

وهناك العديد من دوال الربط يمكن اختيارها حسب نوع توزيع المتغير المعتمد فيما إذا كان normal ، Poisson، ... الخ. والآتي بعض من هذه الدوال:-

التوزيع	Link function	$f(Z)$
Normal	Identity	$f(Z) = Z$
Poisson	Log Link	$f(Z) = \text{Log } Z$
Binomial	Log it Link	$f(Z) = \text{Log}(Z/1 - Z)$

إن تقدير معلمات النموذج الخطي المعمم يكون باستخدام طريقة الإمكان الأعظم MLE، والتي تتطلب طرائق تكرارية (iterative) وهناك العديد من الطرائق التكرارية MLE في النموذج الخطي المعمم وتعد (Newton-Raphson)، (fisher - Scoring) من بين الطرائق الأكثر كفاءة وشيوعاً في الاستخدام في النموذج الخطي المعمم.

ويتم اختبار معنوية المعلمات باستخدام اختبار Wald وغيرها.

1.6 نموذج إنحدار بواسون [1],[2],[3] :-

يعرف نموذج إنحدار بواسون على أنه الأسلوب الذي يتم من خلاله نمذجة المتغير المعتمد كونه متغير استجابة (Response Variable)، عندما تكون قيم ذلك المتغير بهيئة قيم معدودة (Count data) أو بهيئة معدلات (Rate data)، ويعد نموذج إنحدار بواسون هو الأداة الملائمة لتحليل الحوادث نادرة الوقوع ذات البيانات التي تكون قيمتها غير سالبة.

إن نموذج إنحدار بواسون يفترض بأن العلاقة ما بين متوسط (Y) والمتغيرات التوضيحية (X's) توصف بالشكل الآتي :-

$$E(Y_i) = e^{x'_i \beta} \quad \dots(1)$$

حيث أن:

β : موجه معاملات ذو الدرجة $((P + 1) \times 1)$.

X_i : تمثل الصف (i) للمصفوفة (X).

X : مصفوفة المتغيرات التوضيحية ذات الدرجة $(n \times (P + 1))$.

ويبنى نموذج إنحدار بواسون على ثلاثة افتراضات أساسية تدعم هذا المفهوم وهي :-

الافتراض الأول:

إن الدالة الاحتمالية الشرطية للمتغير المعتمد Y علماً بأن معلمة التوزيع μ تتبع توزيع بواسون بمعلمة قدرها μ ، وأن الصيغة العامة لتوزيع تلك الدالة هي ذات صيغة التوزيع الواردة سابقاً والتي كانت:

$$f(Y|\mu) = \frac{e^{-\mu} \mu^y}{y!}, \quad Y = 0, 1, 2, \dots$$

وأن هذا الافتراض واجب التحقق عندما $(\mu > 0)$.

الافتراض الثاني:

يمكن التعبير عن معلمة التوزيع للمتغير المعتمد (Y) بأنها مساوية إلى :-

$$\mu_i = e^{x'_i \beta}$$

الافتراض الثالث :

إن الأزواج المرتبة للمشاهدات (Y_i, X_i) لها توزيع مستقل، إجمالاً يمكن من خلال دمج الافتراضين الأول والثاني للحصول على الدالة الاحتمالية الشرطية الآتية^[3]:

$$f(Y_i|X_i) = \frac{e^{-e^{x'_i \beta}} e^{Y_i x'_i \beta}}{Y_i!}$$

إن نموذج إنحدار بواسون يحقق دالة الوسط الحسابي الشرطية الأسية (الخطية – اللوغارتمية) كون النموذج سيتحول بأخذ اللوغاريتم الطبيعي إلى إنموذج خطي:

$$E(Y_i|X_i) = \mu_i = e^{x'_i \beta}$$

ودالة التباين الشرطية الأسية (الخطية – اللوغارتمية)

$$Var(Y_i|X_i) = \mu_i = e^{x'_i \beta}$$

وكما ذكرنا آنفاً بأن توزيع بواسون فيه خاصية تساوي الوسط الحسابي والتباين، وهنا أيضاً تنطبق هذه الخاصية في نموذج إنحدار بواسون. إن نموذج بواسون في الأساس هو نموذج إنحدار غير خطي (non-Linear) حيث أن تأثير المتغيرات التوضيحية ترتبط بمتغير الاستجابة بصورة لا خطية ولكن من الممكن تحويلها إلى صيغة خطية من خلال رابط (Connector) أو ما يسمى بدالة الربط f(Z) (Link function) المذكورة آنفاً، حيث أن:

$$f(z) = \text{Log } z$$

$$Z = E(Y)$$

وذلك في حالة التوزيع بواسون وبالتالي فإن الصيغة الخطية لـ μ_i تصبح:-

$$\text{Log } \mu_i = X'_i \beta$$

هي الصيغة الأكثر شيوعاً والتي تسمى (Log-Linear model) والتي تعد حالة خاصة من أنموذج الإنحدار المعمم (Generalized Linear model) الذي تم عرض تفاصيله آنفاً.

1.7 تقدير معلمات نموذج إندار بواسون [1],[2]:

يمكن استخدام طريقة الإمكان الأعظم MLE لتقدير المعلمات ، حيث يتم الاستناد إلى الافتراضات الثلاثة المذكورة آنفاً وكما مبين في أدناه :-

$$Pr(Y = Y_i) = \frac{e^{-\mu_i} \mu_i^{Y_i}}{Y_i!}, i = 1, 2, \dots, n$$

ومن خلال تعظيم المشاهدات لتوزيع المتغير المعتمد (Y_i) ، تكون دالة الإمكان الأعظم بالشكل الآتي:-

$$L(Y_1, Y_2, \dots, Y_n; \mu_i) = \frac{E^{-\sum_{i=1}^n \mu_i} \mu_i^{\sum_{i=1}^n Y_i}}{\prod_{i=1}^n Y_i!}$$

وبأخذ اللوغاريتم الطبيعي لدالة الإمكان للمشاهدات أعلاه نحصل على :-

$$\log L(Y_i | X_i, \underline{\beta}) = - \sum_{i=1}^n \mu_i + \sum_{i=1}^n Y_i (\log \{\mu_i\}) - \log \left\{ \prod_{i=1}^n Y_i! \right\} \quad \dots (2)$$

وبالاعتماد على الافتراض الثاني من الفروض الأساسية لنموذج إندار بواسون $\mu_i = e^{\hat{X}_i \underline{\beta}}$ يتم التعويض في المعادلة (2) أعلاه وكما يلي:-

$$\log L(Y_i | X_i, \underline{\beta}) = - \sum_{i=1}^n e^{\hat{X}_i \underline{\beta}} + \sum_{i=1}^n Y_i (\log \{e^{\hat{X}_i \underline{\beta}}\}) - \log \left\{ \prod_{i=1}^n Y_i! \right\} \quad \dots (3)$$

وباشتقاق المعادلة (3) بالنسبة لموجه المعلمات $\underline{\beta}$ نحصل على:-

$$\frac{\partial \log L}{\partial \underline{\beta}} = - \sum_{i=1}^n (Y_i - e^{\hat{X}_i \underline{\beta}}) X_i$$

وبمساواة مشتقة دالة الإمكان بالنسبة للمعلمات بالصفر.

يمكن الحصول على تقدير لمعلمات نموذج إندار بواسون:

$$\frac{\partial \log L}{\partial \underline{\beta}} = 0$$

$$\sum_{i=1}^n (Y_i - e^{\hat{X}_i \hat{\underline{\beta}}}) X_i = 0 \quad \dots (4)$$

إن معادلة (4) هي غير خطية بالنسبة لموجه المقدرات $\hat{\underline{\beta}}$ ولحل هذه المعادلة نستخدم إحدى الطرق التكرارية المذكورة آنفاً، وقد تم استخدام الطريقة (fisher – Scoring) في هذا البحث، وهناك العديد من الطرائق التكرارية لتقدير ML في النموذج الخطي المعمم generalized linear model حيث تعد طريقة Newton-Raphson وطريقة Fisher-scoring من بين أهم وأكفأ الطرائق التكرارية والاكثر شيوعاً واستخداماً. أن الطريقة التي اعتمدت هي طريقة Fisher-scoring أو ماتسمى [Iterative re-weighted least square] وتتصف بكونها تجهز أو تعطي خوارزمية موحدة (unified) لكافة النماذج الخطية المعممة ، فضلاً عن أنها تعطي تقدير لمصفوفة التباين والتباين المشترك للمعلمات المقدرة .

1.8 إخبار المعلمات [2]:

لقد تم في هذا البحث استخدام إحصاءة (Wald) لاختبار مغنوية معلمة الإندار، حيث تستند هذه الإحصاءة على خاصية التقريب للتوزيع الطبيعي لتقدير الإمكان الأعظم، وتحسب إحصاءة Wald وفق الصيغة التالية :-

$$W = \frac{\hat{\beta}^2}{Var(\hat{\beta})} \quad \dots (5)$$

1.9 نموذج إندار Step Wise

في حالة وجود عدد كبير من المتغيرات التوضيحية لا يمكن افتراض تواجدها جميعاً في النموذج وذلك لاحتمال حدوث مشكلة التعدد الخطي ، فضلاً عن تعقيد الحسابات وغيرها، وهناك العديد من الطرائق التي يمكن استخدامها التي تتيح عدم التعامل مع هذه المتغيرات دفعة واحدة ومن بين هذه الطرق التي تعد من أفضل الطرائق التي يمكن اتباعها لبناء النموذج هي طريقة (Step Wise) التي تنطوي على إدخال المتغيرات التوضيحية واحداً تلو الآخر في ضوء أعلى ارتباط بالمتغير المعتمد، فضلاً عن اختبار معنوية المتغير الجديد الداخل بوجود المتغيرات السابقة وبذلك اختبار معنوية المتغيرات السابقة بوجود المتغير الجديد الأمر الذي يضمن احتواء النموذج على المتغيرات المعنوية فقط، مما يزيد كفاءة النموذج . وقد تم اعتماد هذه الطريقة في بناء النموذج في هذا البحث.

1.10 جدول النوفيق Goodness of fit [3]

بعد توفيق أنموذج إندار بواسون للبيانات لابد من معرفة مدى ملائمة وجودة النموذج الموضوع، إن توزيع بواسون يعد من التوزيعات الخاصة جداً حيث أن المتوسط μ والتباين σ^2 هما متساويان فإذا ما استخدمنا النسبة $\frac{variance}{mean}$ كمعلمة تشتت (dispersion Parameter) فإن قيمة هذه المعلمة إذا كانت تساوي (1) فإن ذلك يشير إلى نموذج إندار بواسون القياسي (Standard Poisson model) وإن النموذج ملائم تماماً، في حين إذا كانت قيمة معلمة التشتت أكبر من (1) فإن ذلك يشير إلى أن الأنموذج فوق التشتت (over dispersion model) والنموذج الموضوع غير ملائم، وإذا كانت قيمة المعلمة أقل من (1) فإن الأنموذج تحت التشتت (under dispersion model) والنموذج أيضاً غير ملائم، وعند حصول حالة عدم الملاءمة لابد من إعادة توفيق النموذج فوق التشتت، أو النموذج تحت التشتت، إن معلمة التشتت تؤدي دوراً مهماً وأساسياً في إعادة التوفيق، ويمكن التعبير بالشكل الآتي:

$$Var(Y) = \theta \mu$$

بدلاً من:

$$Var(Y) = \mu$$

حيث أن θ هي معلمة التشتت.

ويتم إعادة التوفيق سواء الأنموذج فوق التشتت أو تحت التشتت بالشكل الآتي :-
يتم توفيق النموذج بالشكل الاعتيادي حيث يجري تقدير المعلمات وحساب مصفوفة التباين والتباين المشترك للمعلمات المقدرة، ثم يتم ضرب معلمة التشتت (θ) في هذه المصفوفة.
بعد إعادة التوفيق هناك إنعكاسات على التحليل وحساب النتائج وذلك بالنسبة للأنموذج فوق التشخيص وكذلك تحت التشخيص نعرضها كما يأتي:

أولاً : بالنسبة للنموذج فوق التشتت :

إن إعادة التوفيق سوف تنعكس على مصفوفة التباين والتباين المشترك إذ تكون التباينات للمعلمات المقدرة بقيم أكبر وهذا يقود بالتالي إلى أن تكون الأخطاء المعيارية بقيم أكبر مما يؤدي إلى أن تكون قيم إحصاء الاختبار Wald بقيم أصغر مما يجعل بعض المتغيرات التوضيحية غير معنوية في حين كانت معنوية قبل إعادة التوفيق، وكذلك فترات الثقة تكون أطول مما هي عليه قبل إعادة التوفيق، أما قيم المقدرات للمعلمات (Point estimation) لا تتأثر وتبقى نفسها.

ثانياً : بالنسبة للأنموذج تحت التشتت:

إن الإنعكاسات بعد إعادة التوفيق تكون عكس النموذج فوق التشتت إذ أن مصفوفة التباين والتباين المشترك تكون بقيم أقل مما يعني أن الأخطاء المعيارية سوف تكون أقل وهذا يقود إلى أن بعض المتغيرات التوضيحية التي لم تكن معنوية تصبح معنوية، وفترات الثقة تكون أقصر.

إن قيمة المعلمة (θ) يتم وضعها إما على أساس معلمة القياس (Scaled deviance) أو (Scaled Pearson chi-square) والجدير بالذكر أن قيمة أحدهما إذا ما اعتمد كقيمة لـ θ سوف تصبح مساوية إلى (1) والأخرى قريبة من الواحد وذلك بعد إعادة التوفيق.

ولابد من الإشارة إلى أن هنالك نوعين رئيسيين من البواقي يتم التعامل معهما الأول Pearson residuals والذي يقوم على أساس الفرق بين قيمة Y والقيمة المقدرة لها $\hat{Y}$. والثاني (deviance) والذي يقوم على أساس مساهمة الاستجابات المشاهدة لأحصاء Log-Likelihood.

1.11 الخطورة النسبية :

في حالة وجود جدول توافق (2x2) كما في الجدول الآتي:-

Case \ Group	+	-
Group (1)	A	b
Group (2)	C	d

والذي يبين وجود حالة معينة (+) وعدم وجودها في مجموعتين (1)، (2) فيمكن حساب الخطورة النسبية بالصيغة الآتية:

$$\hat{\Psi} = \frac{a \cdot d}{c \cdot b} \quad \dots (6)$$

وتعرف الخطورة النسبية على إنها احتمال تضاعف حدوث الحالة (+) في المجموعة (1) بمقدار Ψ عن حدوثها في المجموعة (2).
أما حدود الثقة لها فتحسب كالآتي:

$$\log_e(\hat{\Psi}) \pm 1.96 \sqrt{\text{Var}(\log_e \hat{\Psi})} \quad \dots (7)$$

$$\text{Var}(\log_e \hat{\Psi}) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \quad \dots (8)$$

2- الجانب التطبيقي

2.1 جمع البيانات :

إن هدف هذا البحث وكما ذكرنا هو بناء نموذج إنحدار بواسون لتقدير معدل تكرار الإصابة بإحتشاء العضلة القلبية للأشخاص 62 فأكثر في ضوء عدد من المتغيرات التوضيحية المفترضة المتمثلة بالعمر، الضغط، الكليسترول، السكر، الوزن، الجنس، الوراثة، التدخين، وعلى هذا الأساس فإن متغيرات هذا البحث تعرض بالشكل الآتي:
 Y المتغير المعتمد الذي يمثل عدد مرات الإصابة بإحتشاء العضلة القلبية للأشخاص 62 فأكثر خلال العام.
المتغيرات التوضيحية تتمثل بما يأتي :-

- X_1 العمر
- X_2 الضغط
- X_3 الكليسترول
- X_4 السكر
- X_5 الوزن
- X_6 الجنس
- X_7 الوراثة
- X_8 التدخين

وقد تم جمع البيانات المتعلقة بهذه المتغيرات من مستشفى اليرموك لـ (151) مصاباً وبشكل عشوائي، والتي تشكل بالتالي عينة البحث.

2.2 التحليل الإحصائي :

لقد تم تطبيق كافة المراحل الإحصائية الخاصة ببناء النموذج من اختيار للمتغيرات التوضيحية وتقدير المعلمات واختبارها وبيان ملاءمة النموذج وجودة التوفيق التي تم ذكرها آنفاً، وذلك باستخدام البرنامج الجاهز Statistica Version 10 إذ تم الحصول على النتائج وكما مبين في أدناه، وتجدر الإشارة إلى أن المتغير المعتمد والمتغيرات التوضيحية المذكورة أعلاه كانت بمدخلات حاسبة هي Var(1) متغير معتمد و Var(2) إلى Var(9) هي متغيرات توضيحية وبالتسلسل المذكور أعلاه.

2.2.1 طريقة Step Wise

إن النتائج الخاصة بتطبيق طريقة (Step Wise) لاختيار أفضل المتغيرات التوضيحية موضحة في جدول (1) والتي تتضمن مرحلتين، ومن ملاحظة جدول المرحلة الثانية نلاحظ أن المتغير Var(2) الذي يمثل العمر هو المتغير الوحيد المعنوي الداخل إلى النموذج وذلك بمقارنة $Wald P < 0.05$ الخاصة بهذا المتغير وبذلك نرفض H_0 والفروق معنوية فقط بالنسبة لـ Var(2).

جدول (1)
نتائج تطبيق Step wise

	Effect	Degr. of Freedom	Wald Stat.	Wald p	Score Stat.	Score p	Var.Status
Step 1	VAR2	1			31.33463	2.17E-08	Entered
	VAR3	1			16.18562	5.74E-05	Out
	VAR4	1			3.211412	0.073126	Out
	VAR5	1			14.85927	0.000116	Out
	VAR6	1			1.319078	0.250758	Out
	VAR7	1			7.119955	0.007623	Out
	VAR8	1			0.067312	0.795292	Out
	VAR9	1			2.249472	0.13366	Out
	1*2	1			0.053429	0.817201	Out
	1*3	1			2.107063	0.146621	Out
	2*3	1			0.948694	0.330052	Out
	1*2*3	1			0.002507	0.96007	Out
Step 2	VAR2	1	31.1063	2.44E-08			In
	VAR3	1			2.190142	0.138896	Out
	VAR4	1			0.687119	0.407146	Out
	VAR5	1			2.953759	0.085678	Out
	VAR6	1			0.225607	0.634801	Out
	VAR7	1			0.66774	0.413841	Out
	VAR8	1			0.015956	0.899481	Out
	VAR9	1			0.130251	0.718172	Out
	1*2	1			0.652317	0.419286	Out
	1*3	1			0.008981	0.924499	Out
	2*3	1			0.104338	0.746685	Out
	1*2*3	1			0.100414	0.751333	Out

2.2.2. اختبار Wald

إن نتائج تطبيق اختبار Wald موضحة في جدول رقم (2) حيث يبين هذا الجدول أيضاً أن Var(2) متغير العمر هو المتغير الوحيد المعنوي إذ أن الاحتمال الخاص به $P = 0.002809$ هو أصغر من 0.05، هذا فضلاً عن معنوية الحد الثابت .

جدول (2)
نتائج اختبار Wald

Effect	Level of effect	Column	Estimate	Standard Error	Stat.	P
Interc		1*	-3.12067	1.201733*	6.743433*	0.009409*
VAR2		2*	0.02889*	0.009670*	8.927359*	0.002809*
VAR3		3	0.07487	0.07365	1.033408	0.309360
VAR4		4	0.08033	0.002123	0.024118	0.876586
VAR5		5	0.00195	0.001713	1.294547	0.255212
VAR6		6	-0.00016	0.005057	0.000943	0.975505
VAR7	0	7	-0.05438	0.084222	0.416918	0.518479
VAR8	0	8	-0.03552	0.074352	0.228253	0.632822
VAR9	0	9	0.03405	0.073814	0.212823	0.644564
VAR7*VAR8	1	10	-0.05763	0.073445	0.615698	0.432650
VAR7*VAR9	1	11	0.00117	0.074024	0.000251	0.987352
VAR8*VAR9	1	12	-0.02124	0.07276	0.085222	0.770342
VAR7*VAR8*VAR9	1	13	-0.01261	0.072984	0.029840	0.862854
Scale			1	0.000000		

3.2.2. مصفوفة التباين والتباين المشترك
وقد تم أيضاً حساب فترة الثقة للمعاملات المقدرة وكما هي موضحة في جدول رقم (3)، وكذلك مصفوفة التباين والتباين المشترك وهي كما موضحة في جدول (4).

جدول (3)
فترة الثقة للمعاملات المقدرة

Effect	Level of effect	Column	Lower CL 95%	Upper CL 95%
Interc		1	-5.47603	-0.765321
VAR2		2	0.00994	0.047846
VAR3		3	-0.06948	0.219222
VAR4		4	-0.00383	0.004491
VAR5		5	-0.00141	0.005306
VAR6		6	-0.01007	0.009757
VAR7	0	7	-0.21945	0.110691
VAR8	0	8	-0.18125	0.110205
VAR9	0	9	-0.11062	0.178725
VAR7*VAR8	1	10	-0.20158	0.086320
VAR7*VAR9	1	11	-0.14391	0.146258
VAR8*VAR9	1	12	-0.16385	0.121366
VAR7*VAR8*VAR9	1	13	-0.15565	0.130439

جدول (4)
مصفوفة التباين والتباين المشترك

Effect	Interc	VAR2	VAR3	VAR4	VAR5	VAR6	VAR71	VAR81	VAR91	1*21	1*31	2*31	1*2*31
Interc	1.444161	-0.001577	-0.066574	-0.000252936	0.00045758	-0.00245	-0.033213	0.001663	0.008385	0.001567	0.002296	-0.002717	-0.007703
VAR2	-0.001577	9.35E-05	-0.000285	-1.04572E-06	-3.31985E-06	3.23E-06	0.00015	5.46E-05	-0.000101	-8.69E-05	0.000101	3.58E-05	1.63E-05
VAR3	-0.066574	-0.000285	0.005424	-9.6994E-06	-2.75196E-05	2.79E-05	0.000448	-0.00055	-0.000206	0.000247	-0.000513	-3.19E-05	0.000176
VAR4	-0.000253	-1.05E-06	-9.7E-06	4.50859E-06	-7.21326E-07	-2.35E-06	1.05E-05	-6.22E-06	1.08E-05	1.5E-05	1.22E-05	-2.25E-06	7.47E-06
VAR5	0.000458	-3.32E-06	-2.75E-05	-7.21326E-07	2.93357E-06	-1.4E-06	1.63E-05	-6.7E-07	9.06E-06	5.43E-06	-1.97E-07	5.5E-06	3.48E-06
VAR6	-0.00245	3.23E-06	2.79E-05	-2.35486E-06	-1.39881E-06	2.56E-05	0.000133	7.09E-05	-2.45E-05	-2.07E-05	-4.43E-05	-3.49E-06	1.33E-05
VAR7	-0.033213	0.00015	0.000448	1.05087E-05	1.63473E-05	0.000133	0.007093	0.001518	-0.001156	0.001764	-0.001241	-0.000107	-7.92E-05
VAR8	0.001663	5.46E-05	-0.00055	-6.22376E-06	-6.70231E-07	7.09E-05	0.001518	0.005528	-0.000468	0.002764	-0.000315	-0.001225	-0.000864
VAR9	0.008385	-0.000101	-0.000206	1.08433E-05	9.05846E-06	-2.45E-05	-0.001156	-0.000468	0.005448	-0.000128	0.002822	0.001802	0.001121
VAR7*VAR8	0.001567	-8.69E-05	0.000247	1.49931E-05	5.43232E-06	-2.07E-05	0.001764	0.002764	-0.000128	0.005394	-0.000436	-0.000911	-0.001207
VAR7*VAR9	0.002296	0.000101	-0.000513	1.2207E-05	-1.96846E-07	-4.43E-05	-0.001241	-0.000315	0.002822	-0.000436	0.00548	0.001195	0.001856
VAR8*VAR9	-0.002717	3.58E-05	-3.19E-05	-2.25263E-06	5.50039E-06	-3.49E-06	-0.000107	-0.001225	0.001802	-0.000911	0.001195	0.005294	0.002891
VAR7*VAR8*VAR9	-0.007703	1.63E-05	0.000176	7.47168E-06	3.48226E-06	1.33E-05	-7.92E-05	-0.000864	0.001121	-0.001207	0.001856	0.002891	0.005327

4. 2. 2. جدول التوفيق goodness of fit
لقد تم حساب معلمة القياس بنوعيه (Scaled deviance) و (Scaled person chi-square) وكما هي موضحة في جدول (5) .

جدول (5)
إحصاءات جودة التوفيق

Criterion	d.f	Value	Scaled value
Deviance	137	28.163	0.205568
person chi-square	137	28.179	0.205688

ومن ملاحظة جدول (5) الحقل الأخير (Scaled value) نجد أن النموذج غير ملائم، وذلك لأن المعلمتين (Scaled deviance) و (Scaled person chi-square) كلاهما أقل من (1) وبذلك فإن النموذج تحت التششت (Under dispersion)، وعليه لا بد من إعادة توفيق النموذج تحت التششت.

5. 2. 2. توفيق النموذج تحت التششت
لقد تم اعتماد المعلمة (Scaled person chi-square) كمعلمة تششت χ^2 و تم ضرب مصفوفة التباين والتباين الموضحة في جدول رقم (4) بهذه المعلمة حيث تم الحصول على النتائج الآتية:-

أولاً : مصفوفة التباين والتباين المشترك جديدة
وكما هي مبينة في جدول (6) ، ويلاحظ على قيم قطر هذه المصفوفة مقارنة مع قيم قطر مصفوفة التباين والتباين المشترك جدول (4)، بأنها أصغر بكثير فعلى سبيل المثال المتغير (2) Var في (4) كان التباين الخاص به 0.000094 في حين أصبح 0.000019 وهكذا بالنسبة للتباين الخاص بكافة المتغيرات

جدول (6)

مصفوفة التباين والتباين المشترك لنموذج تحت التشتت

Effect	Interc	VAR2	VAR3	VAR4	VAR5	VAR6	VAR71	VAR81	VAR91	1*21	1*31	2*31	1*2*31
Interc	0.297047	-0.00032	-0.01369	-5.2E-05	9.41E-05	-0.0005	-0.00683	0.000342	0.001725	0.000322	0.000472	-0.00056	-0.00158
VAR2	-0.00032	1.92E-05	-5.9E-05	-2.2E-07	-6.8E-07	6.64E-07	3.09E-05	1.12E-05	-2.1E-05	-1.8E-05	2.08E-05	7.37E-06	3.36E-06
VAR3	-0.01369	-5.9E-05	0.001116	-2E-06	-5.7E-06	5.73E-06	9.21E-05	-0.00011	-4.2E-05	5.08E-05	-0.00011	-6.6E-06	3.61E-05
VAR4	-5.2E-05	-2.2E-07	-2E-06	9.27E-07	-1.5E-07	-4.8E-07	2.16E-06	-1.3E-06	2.23E-06	3.08E-06	2.51E-06	-4.6E-07	1.54E-06
VAR5	9.41E-05	-6.8E-07	-5.7E-06	-1.5E-07	6.03E-07	-2.9E-07	3.36E-06	-1.4E-07	1.86E-06	1.12E-06	-4E-08	1.13E-06	7.16E-07
VAR6	-0.0005	6.64E-07	5.73E-06	-4.8E-07	-2.9E-07	5.26E-06	2.74E-05	1.46E-05	-5E-06	-4.3E-06	-9.1E-06	-7.2E-07	2.73E-06
VAR7	-0.00683	3.09E-05	9.21E-05	2.16E-06	3.36E-06	2.74E-05	0.001459	0.000312	-0.00024	0.000363	-0.00026	-2.2E-05	-1.6E-05
VAR8	0.000342	1.12E-05	-0.00011	-1.3E-06	-1.4E-07	1.46E-05	0.000312	0.001137	-9.6E-05	0.000568	-6.5E-05	-0.00025	-0.00018
VAR9	0.001725	-2.1E-05	-4.2E-05	2.23E-06	1.86E-06	-5E-06	-0.00024	-9.6E-05	0.001121	-2.6E-05	0.000581	0.000371	0.000231
VAR7*VAR8	0.000322	-1.8E-05	5.08E-05	3.08E-06	1.12E-06	-4.3E-06	0.000363	0.000568	-2.6E-05	0.00111	-9E-05	-0.00019	-0.00025
VAR7*VAR9	0.000472	2.08E-05	-0.00011	2.51E-06	-4E-08	-9.1E-06	-0.00026	-6.5E-05	0.000581	-9E-05	0.001127	0.000246	0.000382
VAR8*VAR9	-0.00056	7.37E-06	-6.6E-06	-4.6E-07	1.13E-06	-7.2E-07	-2.2E-05	-0.00025	0.000371	-0.00019	0.000246	0.001089	0.000595
VAR7*VAR8*VAR9	-0.00158	3.36E-06	3.61E-05	1.54E-06	7.16E-07	2.73E-06	-1.6E-05	-0.00018	0.000231	-0.00025	0.000382	0.000595	0.001096

ثانياً: فترات الثقة :

قد تم الحصول على فترات ثقة جديدة وكما هي موضحة في جدول رقم (7) إذ نلاحظ أن فترات الثقة هي أقصر من فترات الثقة الموضحة في جدول رقم (3) فمثلاً بالنسبة $Var(2)$ نجد أن فترة الثقة الخاصة به بعد ضرب المعلمة θ هي $U=0.03749$ ، $L=0.02030$ في حين كانت $U=0.047846$ ، $L=0.00994$ ، وهكذا بالنسبة لبقية المتغيرات.

جدول (7)
فترة الثقة المقدرة الجديدة

Effect	Level of effect	Column	Lower CL 95%	Upper CL 95%
Interc		1	-4.18889	-2.765321
VAR2		2	0.02030	0.03749
VAR3		3	0.00940	0.14034
VAR4		4	-0.00156	0.00222
VAR5		5	0.00043	0.00347
VAR6		6	-0.00465	0.00434
VAR7	0	7	-0.12925	0.02048
VAR8	0	8	-0.10161	0.03057
VAR9	0	9	-0.03156	0.09967
VAR7*VAR8	1	10	-0.12291	0.00766
VAR7*VAR9	1	11	-0.06463	0.06697
VAR8*VAR9	1	12	-0.08592	0.04344
VAR7*VAR8*VAR9	1	13	-0.07748	0.05227

ثالثاً : الأخطاء المعيارية واختبار Wald

لقد تم الحصول على قيم جديدة للأخطاء المعيارية، وكذلك قيم جديدة لإحصاءة الاختبار $Wald$ ، في حين أن مقدرات المعلمات بقيت بنفس القيم، وكما موضح هو في جدول رقم (8)، ويلاحظ أن قيم الأخطاء المعيارية في هذا الجدول مقارنة بقيمها في جدول رقم (2)، تصنف بكونها أصغر بكثير، الأمر الذي انسحب على أن إحصاءة اختبار $Wald$ التي كانت غير معنوية لبعض المتغيرات قبل إعادة التوفيق أصبحت معنوية بعد إعادة التوفيق، إذ نلاحظ في جدول رقم (8) أن المتغيرين $Var(3)$ و $Var(5)$ أصبح كلاهما معنوي فضلاً عن معنوية $Var(2)$ وذلك عند مقارنة الاحتمال (P) الخاص بكل منها مع 0.05 في حين كان $Var(3)$ و $Var(5)$ غير معنويتين قبل إعادة توفيق. وعليه فإن المتغيرات المعنوية سوف تكون :-

$Var(2)$ ويمثل العمر
 $Var(3)$ ويمثل الضغط
 $Var(5)$ ويمثل السكر

جدول (8)
مقدرات المعلمات

Effect	Level of effect	Column	Estimate	Standard Error	Stat.	P
Interc		1*	-3.12067	0.545020*	32.78475*	0.000000*
VAR2		2*	0.02889*	0.004386*	43.40241*	0.000000*
VAR3		3*	0.07487	0.033402*	5.02415*	0.024996*
VAR4		4	0.08033	0.000963	0.11725	0.732033
VAR5		5*	0.00195	0.000777*	6.29374*	0.012117*
VAR6		6	-0.00016	0.002294	0.00458	0.946023
VAR7	0	7	-0.05438	0.038197	2.02695	0.154531
VAR8	0	8	-0.03552	0.033721	1.10970	0.292146
VAR9	0	9	0.03405	0.033477	1.03469	0.309060
VAR7*VAR8	1	10	-0.05763	0.033309	2.99336	0.083607
VAR7*VAR9	1	11	0.00117	0.033572	0.00122	0.972116
VAR8*VAR9	1	12	-0.02124	0.032999	0.41432	0.519783
VAR7*VAR8*VAR9	1	13	-0.01261	0.033100	0.14507	0.703288
Scale			0.45353	0.000000		

رابعاً : جدول التوفيق

قد تم حساب الحصول على معلمتي القياس الجديدة وهما (Scaled و Scaled deviance) (person chi-square) وكما موضحة في جدول رقم (9).

جدول (9)

معلمتي القياس الجديدة

Criterion	d.f	Value	Scaled value
Deviance	137	136.920	0.999418
person chi-square	137	137	1

ومن ملاحظة الجدول نجد أن معلمة التشنت (Scaled person chi-square) تساوي (1) تماماً و (Scaled value) تقترب من الواحد وبذلك فإن النموذج الذي تم بناؤه ملائم تماماً للبيانات.

وفي ضوء ما تقدم فإن نموذج إنحدار بواسون المقدر الذي تم بناؤه هو كالآتي:

$$\hat{\mu}_i = e^{-3.12067 + 0.02889 X_1 + 0.07487 X_2 + 0.00195 X_4}$$

والذي يمكن بواسطته تقدير معدل تكرار الإصابة باحتشاء العضلة القلبية خلال العام بعد معرفة العمر، الضغط والسكر وذلك للأشخاص 62 عاماً فأكثر .

6.2.2. الخطورة النسبية Relative Risk

لقد تم تصنيف البيانات في جدول توافق 2×2 وكما هو موضح في جدول (10):

جدول (10)

تصنيف البيانات

تكرار الإصابة العمر	$\geq \bar{X}$	$\leq \bar{X}$	
65 فأكثر	101	25	126
64 فأقل	3	22	25
	104	47	151

حيث أن $\bar{X}$ تمثل متوسط تكرار الإصابة باحتشاء العضلة القلبية عند المرضى خلال العام .

وفي ضوء معطيات هذا الجدول تم احتساب الخطورة النسبية لتكرار الإصابة باحتشاء العضلة القلبية بتطبيق معادلة (6) حيث كانت قيمتها تساوي $\Psi = 29.6$ وهذا يعني احتمال تكرار الإحتشاء في العضلة القلبية يزداد بمقدار 30 مرة تقريباً في المجموعة التي أعمارها 65 فأكثر من المجموعة التي أعمارها 64 فأقل.

كما تم احتساب حدود الثقة لـ Ψ بتطبيق معادلة (7)، (8):

$$U = 106.8044$$

$$L = 8.2038$$

3. الاستنتاجات والتوصيات

3.1. الاستنتاجات:

من خلال مجريات هذا البحث تم التوصل إلى ما يأتي:

1. تم بناء نموذج إنحدار بواسون والمتمثل بالصيغة التقديرية الآتية:

$$\hat{\mu}_i = e^{-3.12067 + 0.02889 X_1 + 0.07487 X_2 + 0.00195 X_4}$$

والذي يمكن استخدامه لتقدير معدل تكرار الإصابة باحتشاء العضلة القلبية للأشخاص بعمر 62 عام فأكثر. وذلك في ضوء قيم معلومة لكل من العمر X_1 والضغط X_2 والسكر X_4 .

2. إن المتغيرات التي لها تأثير معنوي على معدل تكرار الإصابة باحتشاء العضلة القلبية للأشخاص 62 عاماً فأكثر تمثلت بالعمر والضغط والسكر في حين أن الكوليسترول والوزن والجنس والوراثة والتدخين لم يكن لها تأثير معنوي في النموذج.

3. إن احتمال تكرار الإحتشاء في العضلة القلبية يزداد بمقدار 30 مرة تقريباً في مجموعة الأشخاص بأعمار 65 فأكثر عن مجموعة الأشخاص بعمر 64 فما دون حيث بلغت الخطورة النسبية $\Psi = 30$.

2.3. التوصيات:

نوصي باستخدام نموذج إحداد بواسون المقدر الذي تم بناؤه في هذا البحث لتقدير معدل تكرار الإصابة في احتشاء العضلة القلبية للأشخاص بعمر 62 فأكثر بعد معرفة العمر، الضغط والسكر، إذ أن تقدير هذا المعدل يعد أحد الجوانب التحذيرية المهمة للمصابين والتي تستدعي منهم المراقبة والسيطرة على عاملي الضغط والسكر لديهم .

المصادر:

1. Cameron A. C & Trivedi P. K (1998), "Regression Analysis of Count Data", Cambridge University Press New York USA.
2. Damodar N. Gujarati (2005), "Basic Econometric", Fourth ed.
3. William H. Greene (1997), "Econometric Analysis", third ed New York.
4. Winkelmann R (2008), "Economic Analysis of Count Data", 5 ed, Springer, Verlag Bertin.
5. Wolfgang Hardle 2004."Non parametric and semi parametric modle"

.....
.....
.....