

## Neural Networks for Blood Bank Quantity Prediction of Blood Components for Leukemia Patients

Hani Amasha\*  
Wael Al-Imam\*\*  
Barra Nouman\*\*\*

(Received 25 / 8 / 2019. Accepted 7 / 10 / 2019)

### □ ABSTRACT □

Blood banks are the main provider of the large amounts of blood to Acute Myeloid Leukemia Patients undergoing Stem Cells Treatment transplant. The bank always strives to predict the required amount so through identifying the correlations between blood analysis indicators (independent variables) that are believed to influence the success of Acute Myeloid Leukemia (AML) treatment via Stem Cells Transplantation (SCT) and between the outcome (dependent variables) of treatment which are: Hospitalization Period, Relapse of leukemic cell, and the Overall Survival. Since, as the hospitalization period and lengthens and the relapse rate increases, it means we have to get ready for another round of transplantation, and, hence, prepare new quantities of blood again and again. The overall survival gives a longer prediction period to take into consideration. If it were to be short period, a new quantity of blood is prepared anyway should an emergency occurs

Influencing inputs are quite few but has to be precisely specified from an inferential statistical viewpoint in order to reduce their number. Data from two hospitals (22+20=42 patients) were collected and a sample from external hospital counted 898 patients was also studied, and Kendal Tau-b was adopted using SPSS.

Eighteen indicators (independent variables) were considered, some of which: age, weight, hemoglobin, white blood cells, response to radiation therapy and to chemotherapy prior to transplantation. The study figured out all those indicators that correlates with each of the dependent (or outcome) variables mentioned above.

Dependent (or outcome) variables included hospitalization, relapse, and overall survival. It has been discovered what variables has an influence on the dependent variables; or outcomes. These were used as inputs to the neural network where 85% of the sample were assigned for training, 5% for validation, and 10% for testing. Prediction accuracy reached 88% which is a very good result, with nearly 20 neurons and 12 epochs to reach that result. Gradient has also dropped to very small bottom (0.00125) and Regression was 70%, which has not been available to the physician before. Equipped with this decision aiding tool, the physician can rethink the transplantation protocol. The project has, thus, achieved its goal which is providing a decision support system that enables blood bank managers to reserve such an ever-increasing amount of blood bags for those patients.

**Keywords:** blood bank, Prediction, Stem Cells, Leukemia.

---

\* Professor, Biomedical Engineering Dept., FMEE, Damascus University, Damascus, Syria.

\*\* Associate Professor, Biomedical Engineering Dept., FMEE, Damascus University, Damascus, Syria.

\*\*\* Postgraduate Student (Ph.D.), Biomedical Engineering Dept., FMEE, Damascus University, Damascus, Syria.

## استخدام الشبكات العصبونية لتنبؤ بنك الدم باحتياجات مرضى سرطان الدم

د. هاني عماشه\*

د. وائل الامام\*\*

براءة نعمان\*\*\*

(تاريخ الإيداع 25 / 8 / 2019. قُبِلَ للنشر في 7 / 10 / 2019)

### □ ملخص □

يهتم بنك الدم بتأمين احتياجات مرضى سرطان الدم. ويسعى دوماً للتنبؤ بهذه الاحتياجات من خلال تحديد بعض المتغيرات المستقلة التي تسهم بالتأثير بدرجات متباينة على نتائج علاج سرطان نقي الدم الحاد بالخلايا الجذعية التي تنحصر بمتغيرات تابعة (غير مستقلة) هي فترة الاستشفاء والنكس والبقيا. إذ أنه كلما طالت فترة الاستشفاء وزاد معدل النكس فهذا يدل على أننا يجب أن نستعد لجولة أخرى من الزرع وبالتالي تحضير كميات جديدة من الدم. كذلك الأمر بالنسبة للبقيا الذي ينبئنا بأن هذا المريض سيتجاوز محتته بعد الزرع الأول وبالتالي لن يحتاج لكميات دم جديدة. أما إذا كان البقايا قصيراً فقد نبادر لجولة علاج وتحضير كميات دم احتياطاً لأنه وإن كان المريض محتملاً وفاته بأي لحظة سنحاول محاولة أخيرة وهذا طبعاً يستلزم كميات دم.

تم تحديد المشعرات (المتغيرات المستقلة)، المتوقع ان تسهم بشكل فاعل في نتائج العلاج، بدقة من وجهة نظر إحصائية استدلالية لأجل تقليص عددها وبهذا يقل تكرار العوامل المترابطة. بلغ حجم العينة المحلية 42 مريضاً من مشفيين محليين أما العينة الخارجية فحصلنا عليها من مشفى آخر وعددها 898 واستعملت طريقة معامل ترابط كندال- تاو-ب التراتبي لكشف الارتباط الاحصائي بين المتغيرات المستقلة بحد ذاتها، ثم كشف الارتباط بين المتغيرات المستقلة وغير المستقلة.

وقد تم تحديد ثمانية عشر مشعراً منها: العمر، الوزن، الخضاب، الكريات البيض، والاستجابة للعلاج الإشعاعي والكيمائي، الخضاب، الكريات البيض، الصفائح الدموية، نوع الزرع ودرجة قربي المتبرع وغيرها. وقد بينت هذه الدراسة وفرزت كل من تلك المتغيرات المستقلة التي لها أثر واضح (ترابط) على كل تابع خرج غير مستقل مما سبق. بعدها أدخلت هذه المتغيرات المترابطة إلى الشبكة العصبونية. حيث خصّصت 85% من عناصر العينة للتدريب و 5% للتحقق و 10% للتجريب. بلغت دقة تنبؤ الشبكة 88% وهي نسبة جيدة جداً وتحقق هذا بعشرة عصبونات وعدد تكرارات لا تزيد عن 12. كما تناقص التدرج الانحداري لقيم جداً صغيرة (0.00125)، كذلك الحال مع معامل الترابط (R= Regression) الذي بلغ حوالي 0.7، وهذا ينم على أنّ متغيرات الدخل المترابطة هذه قادرة على تفسير النتائج بنسبة سبعين بالمائة وهذا بدوره أمر لم يكن متوافقاً قبلاً للطبيب أثناء إعداد بروتوكولات العلاج المعتادة. وهكذا فقط حقق النظام الغاية المرجوة منه وهي توفير آلية لدعم القرار مدير بنك الدم لتأمين كمية الدم المطلوبة لهذا العدد المتزايد من المرضى.

**الكلمات المفتاحية:** سرطان الدم، التنبؤ، خلايا جذعية، بنك الدم.

\* أستاذ، قسم الهندسة الطبية، كلية الهك، جامعة دمشق، دمشق، دمسق، سورية.

\*\* أستاذ مساعد، قسم الهندسة الطبية، كلية الهك، جامعة دمشق، دمشق، سورية.

\*\*\* طالبة دكتوراه، قسم الهندسة الطبية، كلية الهك، جامعة دمشق، سورية.

**مقدمة:**

من المعروف أن استخدام الخلايا الجذعية أصبح من الطرق الأساسية في علاج سرطان الدم [1، 2، 3، 4]. وهناك العديد من العوامل التي تساعد أو تعيق عملية الشفاء مثل الوزن [5، 6] والعمر [7، 8] وغيرها من العوامل التي درست. أجريت دراسة إحصائية راجعة على هذه البيانات لتحديد المفيد منها [9، 10].

انه من المهام العديدة لبنك الدم أن يجهز ويزود مرضى سرطان الدم بمكونات الدم الضرورية أثناء استخدام الخلايا الجذعية في علاجه. ونظرا لكثرة المتغيرات التي قد تتحكم بنتائج العلاج وجب التحقق من تأثير كل من هذه المتغيرات على نتائج العلاج.

**1- الدراسة الإحصائية للعينتان المحلية والخارجية:**

أجريت دراسة إحصائية راجعة باستخدام برمجية (SPSS) واستعملت طريقة معامل ترابط **كندال-تاو-ب** ( Kendall Tau\_b) الترتيبي لكشف الترابط أو الاقتران الاحصائي ضمن المتغيرات نفسها ومع مخارج الدراسة [11]. رافقت مرحلة تحصيل وتهيئة البيانات العديد من الصعوبات منها على سبيل الذكر لا الحصر: عدم وجود كامل البيانات بشكل مفهرس حاسوبياً في المشفيين المحليين وعدم توافر كادر طبي متفرغ ولديه متسع من الوقت في المشفيين المحليين ذي خبرة للبحث في السجلات وفهم السجلات لكي يستخلص لنا البيانات من السجلات والملفات الورقية والحاسوبية لأنه لا يجوز وفقاً للأخلاقيات الطبية أن يطلع على البيانات شخص غير مخول لأن ذلك يخرق قواعد الأخلاقيات الطبية. هذا أخرجنا كثيراً في عينتنا المحلية. بعد الحصول على بيانات الأسماء بشكل مرمز، تبين أنها تعود لعدة كيانات إمرضية، فاضطررنا لاستبقائها جميعاً. وعدم تطابق توافر الحقول المسجلة في المشفيين بشكل كامل حدا بنا لحذف المعلومات غير المشتركة.

أما في بيانات العينة الخارجية فقد وجدنا أن العديد من البيانات ناقصة وبعضها منقوصة الحقول بشكل سيئ. أحيانا بعض التواريخ خاطئة مثل ورود تاريخ النكس بعد تاريخ الوفاة. مما اضطرنا لإعادة التواصل مع المشفى لتصحيح هذه الأخطاء. وفي أحيان كثيرة اضطررنا لحذف بعض الأسطر بسبب التأخير في الإجابة من قبلهم أو في حال كانت النواقص كثيرة تعيق الاستفادة من هذا السجل.

**1-1: وصف عينة متغيرات الدخل:**

تم استقصاء 18 متغيراً يعتقد أن لها تأثيراً كبيراً في نجاح العلاج. هذه المتغيرات هي:

- 1- الجنس (Sex) وترميزه (F/M).
- 2- العمر (Age) بالسنوات (Years).
- 3- الوزن (W) بالكيلوغرام (Kg).
- 4- الطول (H) بالسنتيمتر (Cm).
- 5- زمرة الدم ( $A^{\pm}, B^{\pm}, O^{\pm}, \dots$ ).
- 6- مدخن (Smk) ام لا ( $Y=1/N=0$ ).
- 7- يتناول الكحول (Alcol) ام لا ( $Y=1/N=0$ ).
- 8- الحالة البدنية للمريض (Physical Status). تتراوح بين قدرته على التجوال بحرية خارج المنزل إلى قليل النشاط.
- 9- المرحلة السريرية عند بدء المرض (SD).
- 10- التشخيص المرضي (DS) (ينقسم المرضى على أربعة كيانات إمرضية).

- 11- التتميط المناعي (FC) (في حال وجود / عدم وجود أثر للمرض)
  - 12- الخضاب (Hb) وتغيراته عادة بطيئة قياسا بغيره من العوامل.
  - 13- كريات الدم البيضاء (WBC). تم قياسها تماما قبل الزرع. وتتبع تغيراتها الكيان الإمبراضي، وعادة تكون تلك التغيرات حادة وشديدة سواء ارتفاعا أو انخفاضاً.
  - 14- الصفائح (Plt): قيست قبل الزرع أيضا.
  - 15- عدد الخطوط العلاجية (Ln): وهي عدد المرات السابقة التي خضع فيها المريض للعلاج قبل أن يأتي للدراسة.
  - 16- هل خضع المريض للعلاج الإشعاعي أم لا؟
  - 17- الاستجابة للعلاج الإشعاعي (PTr). وتتراوح بين (عدم استجابة) إلى (استجابة تامة)
  - 18- الاستجابة للعلاج الكيماوي قبل الزرع (CTr). وتتراوح بين (عدم استجابة) إلى (استجابة تامة).
- 1-2: المخرجات:**

قدمت الدراسة السابقة معرفة ارتباط زمن الاستشفاء والنكس والبقيا ببعض متغيرات الدخل. والمقصود بفترة الاستشفاء بالأيام (Days) هي الفترة بين الدخول بالتنشيط وارتفاع الكريات البيض فوق حد معين (حيث أن لحظة الزرع هو اليوم صفر). وتقاس فترة الاستشفاء بالأيام (days). وقد لوحظ أن زمن الاستشفاء لا يتوزع توزيعاً طبيعياً (Normal Distribution). أما الحالات التي تأخر الاستشفاء فوق الثلاثين يوماً فهي حالات لم يحصل فيها استشفاء بل هي حالات نكس. إذا فإن متغير النكس (Relapse) هو أن يمر 40 يوم دون ارتفاع الكريات فوق الحد المطلوب. أو أن يرتفع عدد الكريات لفترة من الزمن ولكن يعود المرض بعد حين. بينما تم ترميز البقيا (OS) أو البقاء على قيد الحياة بالسنوات.

### 1-3: معامل الارتباط (Correlation) ومعامل الأهمية (Significance):

انقسمت الدراسة الاحصائية إلى قسمين : الجزء الأول يعنى بدراسة الترابط المتعارض بين المتغيرات المستقلة نفسها حتى يمكن تقليص تلك المتغيرات التي قد لا تضيف معلومات إضافية مفيدة بوجودها. حيث لم نجد ارتباطات بين غالبية هذه المتغيرات. بينما وجد ارتباط بين التتميط المناعي والتشخيص المرضي مثلاً ولكون التتميط لا يجرى لكل الكيانات الامراضية فلم يتم حذف أحدهما بل استبقي عليه لبيان الأثر على متغيرات الخرج. كذلك الامر فيما يتعلق بعدد الكريات البيض والكيانات الامراضية وللسبب نفسه لم يتم إقصاء أي منها. إن القيم السالبة في معامل الارتباط تعني أن الارتباط عكسي بين متغير الدخل والخرج، أي كلما زاد الدخل نقص الخرج، في حين القيم الموجبة تفيد التناسب الطردي بين متغيرات الدخل والخرج.

أما الجزء الآخر من الدراسة فيخبرنا عن ارتباط كل مخرج من مخرجات الدراسة بعدد من متغيرات الدخل حيث كانت قاعدة القرار هي:  $P\text{-value} > 0.05$

### 1-4: تحليل معاملات الارتباط:

إن تحليل النتائج الاحصائية تظهر الترابط بين كل المتغيرات الدخل سواء مع بعضها أو ارتباطها بمتغيرات التابعة. حيث يبين التدقيق في قيم الجدول (1) أدناه كل هذه الارتباطات. ويمكن أن نستنتج ما يلي:

- زمن الاستشفاء: وهو الزمن المنقضي بعد عملية الزرع حتى ترتفع الكريات البيض للحد المطلوب (ويقاس بالأيام). وهو متغير خرج ثابت القيمة عادة ما لم يحدث نكس ثاني وفترة استشفاء جديدة. وحتى في مثل حالة كهذه يجب الحفاظ على فترة الاستشفاء الأولى وتدوين الفترة الثانية والثالثة... إلخ. يبين العمود الأول من الجدول قيم الترابط

ما بين زمن الاستشفاء ومتغيرات الدخل كافة. وبمعايينة القيم المظللة التي يقل فيها عامل الأهمية عن (0.05) نلاحظ أن متغير تعاطي الكحول (0.023) يؤثر في زمن الاستشفاء بشكل طردي أي أن تناول الكحول يزيد من فترة الاستشفاء وهذا يعد تأثيراً سلبياً من وجهة النظر الطبية. بينما يؤثر متغير الحالة البدنية بشكل عكسي على زمن الاستشفاء، وهذا منطقي إذ أنه كلما كانت الحالة البدنية أسوأ أي أنه كان قليل النشاط (PS=3) كلما استغرق وقتاً أطول للبقاء في فترة الاستشفاء. أما ارتباط زمن الاستشفاء بالتشخيص فهو لأمر معروف كونه يختلف باختلاف الكيان الإمراضي وهذا يجب دراسته منفرداً في المستقبل. كذلك الأمر ارتباط التتميط المناعي بشكل عكسي بزمن الاستشفاء لكون التتميط المناعي يرتبط أصلاً بالتشخيص. وترتبط الكريات البيض بشكل عكسي مع زمن الاستشفاء فكلما زاد عدد الكريات نقصت فترة الاستشفاء. أما عدد الخطوط العلاجية فيرتبط طرداً مع زمن الاستشفاء، أي أنه كلما احتاج المريض إلى خطوط علاجية أكثر فإنه يستلزم فترة استشفاء أطول. أما متغيري الاستجابة للعلاج الشعاعي والاستجابة للعلاج الكيماوي فيتناصبان طرداً مع فترة الاستشفاء، لأن ترميزهما أصلاً يفيد بالاستجابة الجزئية من أجل الترميز العالي (PTr =2, CTr =3).

• النكس: هو أن تتخفف الكريات البيض مجدداً تحت العدد المطلوب وبالتالي يعود المرض من جديد، وهذا قد يحدث في نفس يوم الزرع أو بعده بأيام أو بسنوات. وعندها يجب إعادة الزرع من جديد، وإعطاء ترميز جديد وفي الواقع يجب إجراء دراسة أخرى مستقلة لأولئك الذين ينكسون أكثر من مرة وبيان الأسباب التي تسهم في عملية النكس المتكرر. يعطينا العمود الثاني من الجدول معلومات عن المتغيرات ذات الأثر في موضوع النكس وقابليته. حيث يرتبط بالجنس وهذا الأمر المحير صدفة وليس هنالك دلالة طبية. أما ارتباطه بالمرحلة السريرية طرداً شئاً طبيعياً لأن المراحل السريرية المتقدمة تعني حالة متدهورة للمريض وبالتالي يزداد احتمال النكس لديه. يرتبط النكس عكسياً مع الصفحات، وهذا يمكن فهمه وتفسيره طبيياً من خلال ربط تدني الصفحات بالمرحلة المتقدمة أو حسب الكيان الإمراضي نفسه.

• البقيا: وهو الزمن المنقضي بعد عملية الزرع وحتى وفاة المريض أو إلى يومناً هذا إذا كان لا يزال على قيد الحياة. وهنا يجب أن يتم تبني آلية محددة بين إدارة المشفى والمريض أو أهله لمتابعة المريض كل فترة محددة وبهذا تبقى القيم محدثة وهو أيضاً مدعاة لتكرار الدراسة الإحصائية والعصبونية كل فترة لندخل هذا التبدل في المعطيات بالحسبان. وبالنظر للعمود الثالث من الجدول وبالتحديد تلك المربعات الداكنة التي تقيد الارتباط بين تلك المتغيرات مع البقيا ومدى تأثيرها عليه. فالحالة البدنية للمريض ترتبط طرداً مع البقيا وهذا الأمر لا يحتاج لتبيان. أما تأثير التشخيص العكسي على البقيا فمرده إلى أن الكيان الإمراضي الرابع (HD) أقل نسبة في البقيا. أما دور التتميط المناعي في التأثير على البقيا فينسجم مع فكرة أن التتميط المناعي أصلاً مرتبط بالكيان الإمراضي. كذلك فإن تأثير الكريات البيض على البقيا هو أمر غني عن الشرح والتفسير. أما متغيري الاستجابة للعلاج الشعاعي والاستجابة للعلاج الكيماوي فيتناصبان عكسياً مع البقيا، لأن ترميزهما أصلاً يفيد بالاستجابة التامة من أجل الترميز المنخفض (PTr =1, CTr =1). كذلك فإنه عند توافر بيانات لمرضى أكثر يفضل تقسيم الدراسة لأربعة أجزاء لكل كيان إمراضي.

• إن النتائج التي حصلنا عليها حتى اللحظة قابلة للتطور والتغير عند ازدياد عناصر العينة، ولذا يجب إعادة الدراسة الإحصائية بين الفينة والأخرى. كما معظم النتائج تنسجم مع التبريرات الطبية السائدة، وبعضها الآخر بحاجة للتحقق مع الدراسات السابقة، ومن الاكيد أن يكون التباين في بعض هذه الحالات مرده صغر العينة لدينا ولهذا لم يظهر في الدراسة ارتباط بين الوزن ومعدلات النكس كما هو في إحدى الدراسات المرجعية.

جاءت الدراسة الإحصائية لتكشف ترابطات المتغيرات المستقلة (مشعرات الدخل) بعضها ببعض وبالتالي تقليص عددها للمراحل اللاحقة، ثم دراسة الترابط بين المتغيرات المستقلة المتبقية بعد التقليص ومتغيرات الخرج. وهذا تمهيداً لإدخال هذه المتغيرات جميعها (دخل وخرج) للشبكة العصبونية. تقوم الشبكة العصبونية عندها بالتنبؤ بالمخارج وتخدم الطبيب المعالج عندها كأداة اتخاذ قرار مساعدة للتصرف حيال خطة علاج المريض الجديد.

تم في البداية إجراء الدراسة الإحصائية للبيانات الخارجية لكونها وردت لاحقاً ولم تدمج الدراستين كما أسلفنا لكون متغيرات الدخل المستقلة في العينتين متباينة.

#### 5-1: الدراسة الإحصائية للعينة الخارجية:

كما لاحظنا في السابق تم استقصاء 16 متغيراً بعضها طبي والآخر ذي دلالة اجتماعية. وتمت محاولة دراسة تأثير متغيرات الدخل على كل من زمن الاستشفاء والبقيا كما في الدراسة المحلية ولكننا فوجئنا بالنقص الحاصل في البيانات ومردده عدم متابعة المريض لمعرفة زمن وفاته فيما لو توفي. كذلك ورد أكثر من ثلاثون تاريخ وفاة حاصل قبل تاريخ النكس وهذه أخطاء يجب تصحيحها والتحقق منها قبل المتابعة بدراسة البقيا وزمن الاستشفاء ودراسة الشبكة العصبونية لهما.

- المتغيرات الطبية : نوع الزرع، زمرة الدم، حالة المرض عند الزرع، حالة المرض الحالي، حي أم لا، درجة قربي المتبرع، الكرياتينين قبل الزرع، الوزن، العمر، الطول، الكريات البيض قبل الزرع، الكريات البيض بعد الزرع.
  - المتغيرات الاجتماعية: جنس المريض، جنس المتبرع، مدخن، المهنة، المستوى التعليمي، متزوج أم لا.
- وفيما يلي ندرج فقط نتائج الدراسة الإحصائية لمتغير النكس مع مشعرات الدخل المرتبطة معه. مشعرات الدخل المترابطة مع النكس (Relapse) هذه تساهم هذه المتغيرات باتخاذ القرار بحال وجود أو عدم وجود النكس.

جدول 1: قيم ترابط كل من زمن الاستشفاء والنكس والبقيا مع المتغيرات المستقلة. حيث الرقم الأول في كل مربع هو معامل الارتباط  $\tau_b$  والرقم الثاني اسفله معامل الأهمية P-value

متغير الدخل	زمن الاستشفاء	النكس	البقيا
الجنس	0.161	0.377	0.040
	0.226	0.013	0.772
العمر	0.127	0.070	0.097
	0.256	0.581	0.499
الوزن	0.121	0.066	0.043
	0.275	0.600	0.707
الطول	0.168	0.209	0.003
	0.134	0.101	0.982
مدخن	0.219	0.183	0.179
	0.100	0.225	0.195
كحول	0.302	0.216	0.201
	0.023	0.154	0.147
الحالة البدنية	-0.393	-0.133	0.676
	0.002	0.358	0.000
تشخيص مرضي	0.331	0.035	-0.269
	0.007	0.804	0.036
المرحلة السريرية	0.113	0.267	-0.013
	0.335	0.055	0.917
التتميط المناعي	-0.425	0.226	0.535
	0.001	0.134	0.000
الخضاب	-0.077	0.197	0.056
	0.491	0.118	0.626
الكريات البيض	-0.525	0.125	0.555
	0.000	0.319	0.000
الصفائح	0.043	-0.373	0.116
	0.695	0.003	0.314
الخطوط العلاجية	0.312	-0.225	-0.184
	0.010	0.103	0.144
خضع للإشعاعي	0.180	0.167	-0.081
	0.176	0.271	0.560
الاستجابة للإشعاعي	0.355	-0.065	-0.288
	0.006	0.658	0.032
الاستجابة للكيمائي	0.324	0.004	-0.335
	0.009	0.978	0.010

الجدول (2) قيم ترابط النكس مع متغيرات الدخل المستقلة

عدد المرضى المدون لهم هذه القيمة	معامل الأهمية P-value	معامل الارتباط $\tau_b$	المتغير (مع الترميز إن وجد)
898	0.000	- 0.135	نوع الزرع (0=ذاتي، 1=غيري)
770	0.032	0.077	جنس المتبرع (0=أنثى، 1=ذكر)
898	0.000	- 0.107	العمر عند الزرع (بالسنين)
507	0.007	0.112	المهنة (0=فلاح، عامل، حرفي، 1=موظف، مدير، 2=ربة منزل، 3=غير معروف أو صغير جدا على المهنة)
898	0.001	0.112	الوضع الأسري (0=لم يتزوج مطلقاً، 1=متزوج أو أرمل أو مطلق)
898	0.000	- 0.510	الوضع الحياتي (0=متوفى، 1=على قيد الحياة)
544	0.019	- 0.088	قيمة الكرياتينين قبل الزرع

استخدمت طرق الذكاء الصناعي المختلفة [12، 13، 14] ومنها الشبكات العصبونية في كثير من الأبحاث المتعلقة بعمليات زرع الخلايا الجذعية سواء من مصدر ذاتي أو غيري. حيث استخدمت برمجية الماتلاب لتشكيل هذه الشبكة. أدخلت هذه المتغيرات المترابطة إلى الشبكة العصبونية. حيث خصّصت 85% من عناصر العينة للتدريب و5% للتحقق و10% للتجريب. بلغت دقة تنبؤ الشبكة 88% وهي نسبة جيدة جداً وتحقق هذا بعشرة عصبونات وعدد تكرارات لا تزيد عن 12. كما تناقص التدرج الانحداري لقيم جداً صغيرة (0.00125)، كذلك الحال مع معامل الترابط (R= Regression) الذي بلغ حوالي 0.7، وهذا ينمّ على أنّ متغيرات الدخل المترابطة هذه قادرة على تفسير النتائج بنسبة سبعين بالمائة وهذا بدوره أمر لم يكن متوافراً قبلاً للطبيب أثناء إعداد بروتوكولات العلاج المعتادة. وهكذا فقط حقق النظام الغاية المرجوة منه وهي توفير آلية لدعم القرار بناءً على نتائج الشبكة. وعند بناء الشبكة كان هناك عدة تجارب للوصول إلى أفضل قيم عند بناء كل شبكة وسنذكر في هذه الدراسة أفضل نتيجة توصلنا إليها لكل مجموعة حيث وجدنا أنّ نتيجة الشبكة تتأثر بشكل كبير بعدد العصبونات في الطبقة الخفية التي تولد الشبكة. هناك عدة مؤشرات تدل على نتائج بناء الشبكة العصبونية قمنا بمراقبتها ودراستها للوصول إلى أفضل شبكة واعتمادها بشكل أساسي لإجراء اختباراتنا فيما بعد.

## النتائج والمناقشة:

### 2: نتائج الشبكات العصبونية ونقاشها:

سنبين فيما يلي نتائج الشبكات العصبونية لكلتا العينتين المحلية والخارجية.

#### 1-2: نتائج الشبكة العصبونية للبيانات المحلية:

بعد الدراسة الإحصائية وتحديد المتغيرات التي تؤثر بزمّن الاستشفاء والنكس والبقيا، تم التوصل إلى تحديد الشبكات العصبونية التي سنقوم ببنائها وهي أربع شبكات:

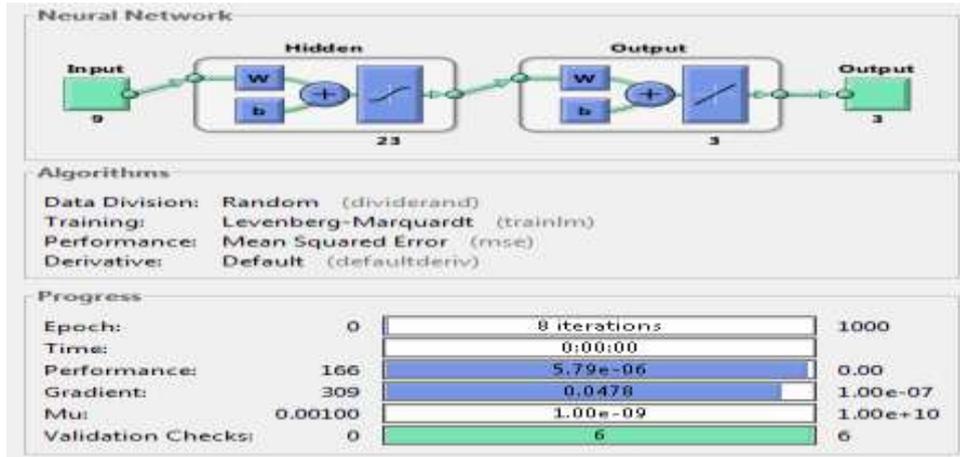
1. الشبكة العصبونية الأولى: تتعلق بزمّن الاستشفاء كخرج للشبكة وتم أخذ جميع المتغيرات المرتبطة به (الاستجابة للأشعة، الاستجابة للكيماوي، عدد الخطوط العلاجية، الكريات البيض، التتميط المناعي، التشخيص المرضي، الحالة البدنية، تناول الكحول) كمدخل للشبكة.

2. الشبكة العصبونية الثانية: تتعلق بالنكس كخرج للشبكة وتم أخذ جميع المتغيرات المرتبطة به (الجنس، المرحلة السريرية عند بدء المرض، عدد الصفحات) كمدخل للشبكة.

3. الشبكة العصبونية الثالثة: تتعلق بالبقيا كخرج للشبكة وتم أخذ جميع المتغيرات المرتبطة به (الاستجابة للأشعة، الاستجابة للكيماوي، الكريات البيض، التتميط المناعي، التشخيص المرضي، الحالة البدنية) كمدخل للشبكة.

4. الشبكة العصبونية الرابعة: تتعلق بزمّن الاستشفاء والنكس والبقيا كمخارج للشبكة وتم أخذ جميع المتغيرات المرتبطة بها (الاستجابة للأشعة، الاستجابة للكيماوي، عدد الخطوط العلاجية، الكريات البيض، التتميط المناعي، التشخيص المرضي، الحالة البدنية، المرحلة السريرية، عدد الصفحات، تناول الكحول) كمدخل للشبكة.

بنيت الشبكات الأولى والثانية والثالثة واستنتجت منها نتائج شبيهة لما وجد في الشبكة الرابعة التي تجمع الثلاث شبكات. وسنعرض فيما يلي الشبكة الرابعة الجامعة درءا للتكرار وهي الشبكة العصبونية لمتغيرات زمن الاستشفاء والنكس والبقيا المبينة في الشكل (1).



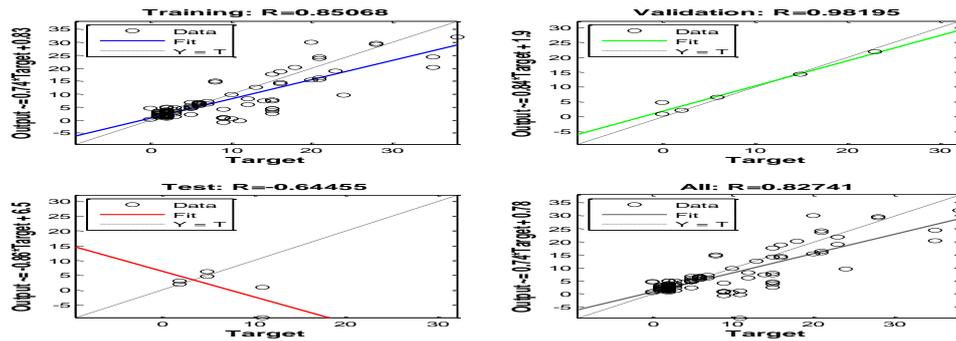
الشكل (1) الشبكة العصبونية لمتغيرات زمن الاستشفاء والنكس والبقيا في العينة المحلية.

حيث نلاحظ نتائج الشبكة في الأشكال (2، 3، 4) التالية:

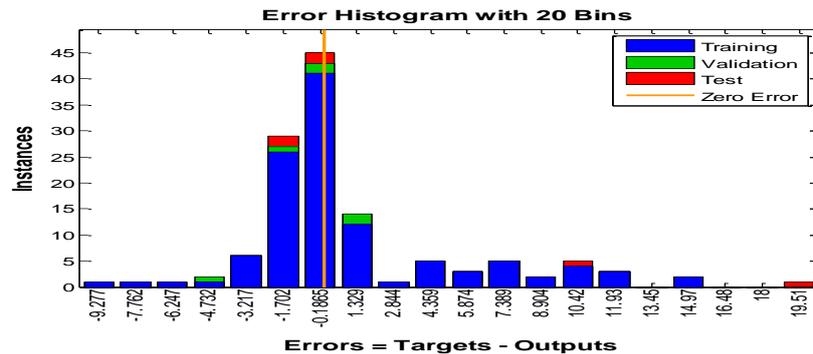
1. إن تابع الهدف قد تناقص من 166 إلى  $5.79 \times 10^{-6}$ .

2. إن تابع الميل قد تناقص من 306 إلى 0.0478.

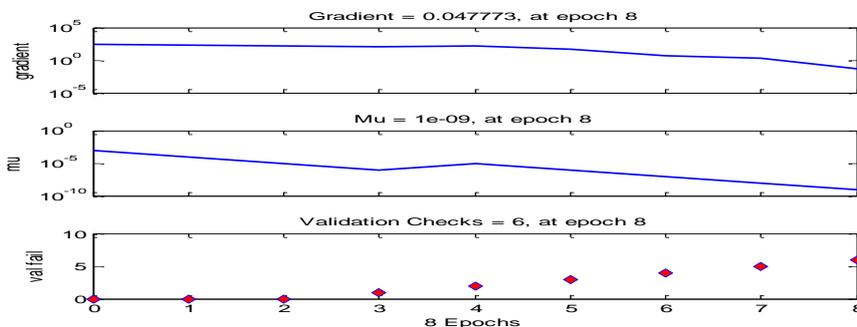
3. إن قيمة  $R=0.827$  وهي قيمة جيدة وهي أكبر من البنية الأولى والثانية.
4. إن المدرج التكراري للأخطاء يوضح توزيعاً قريباً من الطبيعي، ولذلك اعتمدنا هذه البنية.
- ويفسر معامل الانكفاء الخطي في الشبكات العصبونية كما يلي: قيمة  $R=85\%$  هي النسبة التي تتبنا بالكم الذي تستطيع مشعرات الدخل تفسير مشعر الخرج النكس وهذه القيمة تساعد كثيراً على اتخاذ القرار الطبي، لأنها تعطي معلومة طبية إضافية للطبيب لم تكن أصلاً لديه بالتالي تعينه على اتخاذ تدابير جديدة ما كان ليأخذها بعين الاعتبار لولا أن هذه المعلومات قد أضحت متاحة له.



الشكل (2) معامل الانكفاء الخطي R لمتغيرات الدخل مع متغير الخرج زمن الاستشفاء والنكس والبقيا في العينة المحلية



الشكل (3) المنحنى التكراري لخطأ الشبكة العصبونية لمتغيرات زمن الاستشفاء والنكس والبقيا في العينة المحلية



الشكل (4) المخطط الإنحداري لخطأ الشبكة العصبونية لمتغيرات زمن الاستشفاء والنكس والبقيا في العينة المحلية.

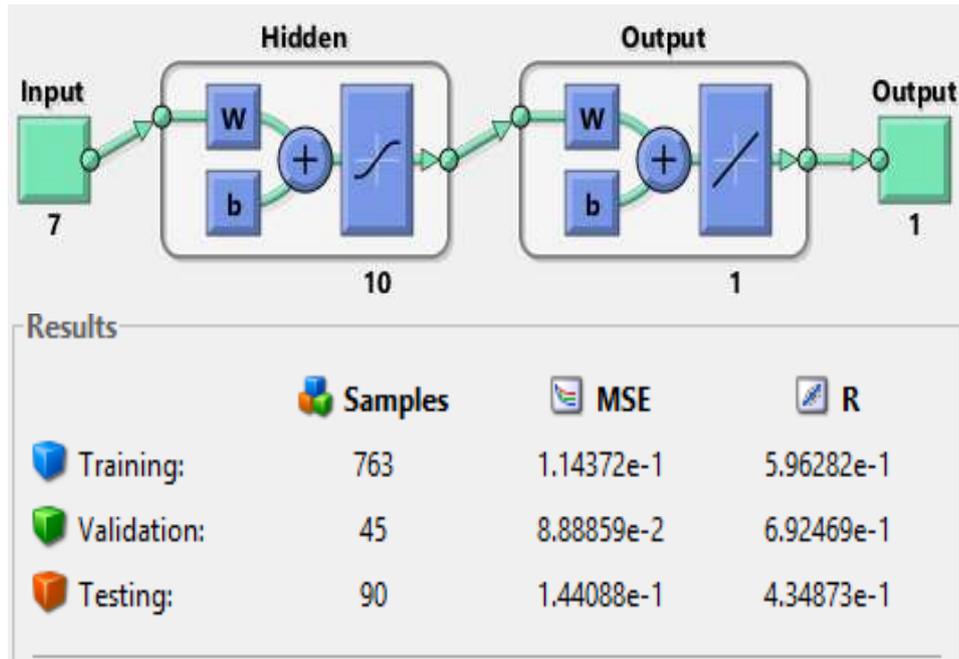
بلغت دقة تنبؤ الشبكة 88% وهي نسبة جيدة جداً وتحقق هذا بحوالي العشرين عصبوناً تقريباً وعدد تكرارات لا تزيد عن 12. كما تناقص التدرج الانحداري لقيم جداً صغيرة (0.00125)، كذلك الحال مع القيمة العالية لمعامل الترابط ( $R=$  Regression) الذي بلغ حوالي 0.7، وهذا ينم على أنّ متغيرات الدخل المترابطة هذه قادرة على تفسير النتائج بنسبة سبعين بالمائة وهذا بدوره أمر لم يكن متوافقاً قبلاً للطبيب أثناء إعداد بروتوكولات العلاج المعتادة وطلب كميات الدم اللازمة للمريض من بنك الدم سواء كان سيشفى أم سينكس. وفي حالة النكس سيحتاج لضعف كمية الدم أما إذا تنبأت الشبكة بأنه لن ينجو فسيقدم له كمية الدم الأولية وينتظر على أية حال ما ستؤول إليه حالته. وهكذا فقط حققت الشبكة الغاية المرجوة منها.

## 2-2 : نتائج الشبكة العصبونية للعينة الخارجية:

قمنا بتجريب الشبكة العصبونية بعدد من العصبونات 10 ثم 13 ثم 16 ثم أكثر من ذلك بكثير ولم نلاحظ أي تحسن على نتائجنا، ولذلك اكتفينا بتبني أصغر عدد ممكن من العصبونات لتحقيق السرعة في النتائج رغم أن السرعة ليست محدداً خطيراً لكون النظام ليس للتشخيص الفوري (instantaneous or online). ولدى حساب متوسطات الأخطاء عند 10، 13، و16 عصبون وجدنا الآتي:

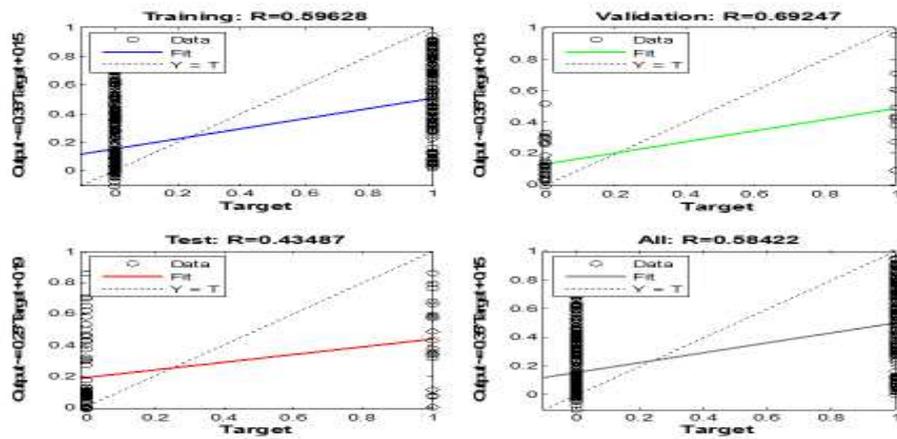
- مجموعة التدريب = 12 %
- مجموعة التحقق = 11.4 %
- مجموعة الاختبار = 11.8 %

تم تبني نسبة الخطأ الواردة الكبرى فيما بينهم وهي 12%. كذلك فإن هذه النسب تتزاح قليلاً حسب العينات المأخوذة في كل مرة كعينات تدريب أو تحقق أو اختبار. وهذا الذي دفعنا باتجاه أخذ أقل دقة بين الثلاثة، أي أسوأ دقة. وبناءً عليه اعتبرنا أن النظام يحقق التنبؤ بدقة تبلغ 88%. وتعد هذه النسبة عالية جداً، إذ أنها وبدقة عالية تنتبأ بما سيؤول إليه حال المريض الجديد ذي المشعرات (أو متغيرات الدخل التي سندونها للمريض) بعد الزرع. وفيما يلي نستعرض نتائج الشبكة لعدد 10 عصبونات ومجموعة التدريب 85% ونسبة العينات في مجموعة التحقق 5% ونسبة العينات في مجموعة الاختبار 10% الوارد أعلاه.



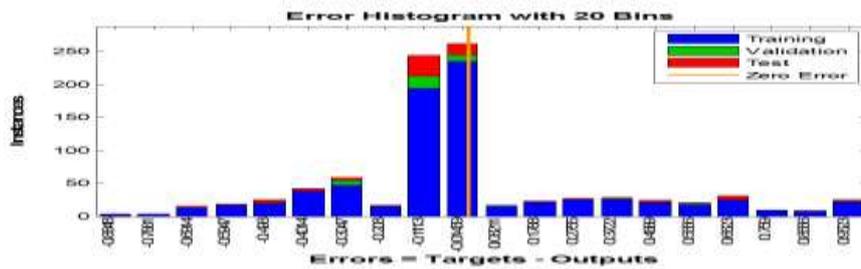
الشكل (5) الشبكة العصبونية للعينة الخارجية بعدد عصبونات 10

نتائج الشبكة العصبونية



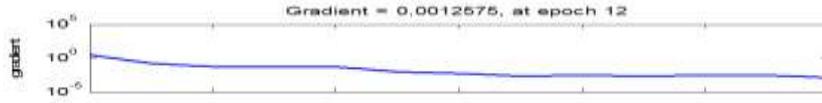
الشكل (6) قيم R للشبكة العصبونية المستخدمة للبيانات الخارجية

المنحني التكراري للخطأ



الشكل (7) المنحني التكراري للخطأ للشبكة العصبونية المستخدمة للبيانات الخارجية

## المخطط الانحداري للخطأ



الشكل (8) المخطط الانحداري للخطأ للشبكة العصبونية المستخدمة للبيانات الخارجية.

حاولنا المقارنة بين نتائج العينة الخارجية والعينة المحلية بخصوص النكس ولكن المقارنة لم تكن ممكنة أو عادلة بسبب عدم تطابق مشعرات الدخول بالعينتين من حيث العدد أو النوع رغم أنهم اشتركوا فيما بينهم بعدد قليل من متغيرات الدخول. كذلك فإن عدد عناصر العينتين متباين بشكل كبير ما بين 42 عنصر للعينة المحلية و 898 مدخلاً في العينة الخارجية وقف عائقاً أمام المقارنة العادلة بين العينة الخارجية والداخلية.

نلاحظ أنه تم التوقف عند نهاية صغرى محلية وليس عند نهاية صغرى مطلقة لقيمة الميل الانحداري (gradient). وهذا ليس بالأمر السيئ أصلاً لكون القيمة النهائية التي بلغت عند هذه النهاية الصغرى تبلغ 0.049 في عينتنا المحلية، بينما تبلغ 0.00125 في العينة الخارجية وهذا الرقم أكثر من جيد وخاصة أنه ظهر عند التكرار 12 بعدد 10 عصبونات فقط.

لم يظهر العمر في النتائج المحلية أو الخارجية بسبب أن تأثير موضوع العمر عادة يظهر للأعمار المسنة (فوق الستين) [7]. ولما كان عدد الأفراد فوق الستون قليلاً في العينتين، لم يظهر التأثير واضحاً مما بدا على أنه تعارض من الدراسات المرجعية. ولو تذكرنا أن معظم دراسات المرجعية حول العمر تناولت المرضى فوق الستين، مما قلل من فرص ظهور الترابط كما هو متوقع.

لقد تمكن البحث من تحقيق المطلوب منه، حيث تم بناء نظام خبير للتنبؤ بدقة عالية (88%) بمعدل نجاح سرطان الدم النخاعي الحاد بالخلايا الجذعية. هذا التنبؤ سيشكل ركيزة نظام دعم القرار فيما يخص التعديلات الواجب تعديلها على البروتوكول السائد المتبع عادة لكل المرضى قبل زرع الخلايا الجذعية. وبالتالي توفير آلية لدعم القرار مدير بنك الدم لتأمين كمية الدم المطلوبة لهذا العدد المتزايد من المرضى.

إن الارتباط بين النكس وبعض المدخلات الاجتماعية أمر غير مستغرب، لأن هذه المدخلات الاجتماعية (حسب اعتقادنا) تخفي في طبيعتها عادات غذائية وسلوكية وجسدية ونفسية متباينة ومن هنا قد يكون مرد تأثيرها على موضوع النكس.

## الاستنتاجات والتوصيات:

إذا فقد توصلنا من خلال الشبكات العصبونية للتنبؤ بنسبة عالية بما سيحصل لمرضى سرطان الدم من حيث تطور حالته نحو الشفاء والنكس إثر عملية زرع الخلايا الجذعية، وهكذا أصبح لدينا مؤشراً حيويًا على كمية الدم التي سيحتاج إليها وهذا يساعد بنك الدم على التنبؤ والتخطيط لكميات كما يلي: قيمة  $R=85\%$  هي النسبة التي تنبئنا بالكم الذي تستطيع مشعرات الدخول تفسير مشعر الخرج النكس وهذه القيمة تساعد كثيراً على اتخاذ القرار الطبي، لأنها تعطي معلومة طبية إضافية للطبيب لم تكن أصلاً لديه وبالتالي تعينه على اتخاذ تدابير جديدة ما كان ليأخذها بعين الاعتبار لولا أن هذه المعلومات قد أوضحت متاحة له. وبالتالي ستمكنه من طلب كميات الدم اللازمة للمريض من بنك

الدم سواء كان سيشفى أم سينكس. وفي حالة النكس سيحتاج لضعف كمية الدم أما إذا تبيأت الشبكة بأنه لن ينجو فسيقدم له كمية الدم الأولية وينتظر على أية حال ما ستؤول إليه حالته. وإثر النتائج الجيدة التي حصلنا عليها، يجب أن يتم التدقيق أكثر على موضوع إجراء تجارب سريرية مبرمجة منذ الآن وصاعداً على صعيد المتغيرات الطبية والمتغيرات الاجتماعية وتخصيص فريق طبي اجتماعي لتفسير بعض النتائج بناء على واقع المرأة مثلاً، أو نوع الوظيفة والمؤثرات النفسية والجسدية المرافقة. ويجب توحيد تسجيل كافة المتغيرات الطبية والاجتماعية في كل المشافي العاملة بهذا المجال في القطر. حتى في حال التخصص في العلاج ضمن المشافي المحلية الخاصة والعامة، يجب التنسيق فيما بينها على التوثيق والبروتوكولات المتبعة. وينصح بإعادة الدراسة الاحصائية كلما ازداد عدد العينة، ومناقشة النتائج في كل مرحلة.

### المراجع:

- [1] Murnaghan I. (2013), "Why Perform a Stem Cell Transplant? Bone Marrow Transplantation and Peripheral Blood Stem Cell Transplantation". In National Cancer Institute.
- [2] Apperley J., Carreras E., Gluckman E., Masszi T., (2012), "The EBMT-ESH Handbook on Haematopoietic Stem Cell Transplantation", Revised edition
- [3] Dick J. E., et al. (2014) "Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia". *Nature*. 20; 506(7488):328-33.
- [4] Karanes C, et al. (2008). "Twenty years of unrelated donor hematopoietic cell transplantation for adult recipients facilitated by the National Marrow Donor Program". *Biology of Blood and Marrow Transplantation* 14 (9 Suppl).
- [5] Giebel, S. et al., (2010) "Association of Human Development Index with rates and outcomes of hematopoietic stem cell transplantation for patients with acute leukemia.", *Blood Journal, American Society of Hematology*, 216(1); 122-28.
- [6] Aplenc R. et al, (2014), "Effect of body mass in children with hematologic malignancies undergoing allogeneic bone marrow transplantation.", *Blood.*; 123(22):3504-11.
- [7] Lim ZiYi, (2010), "Allogeneic Hematopoietic Stem-Cell Transplantation for Patients 50 Years or Older With Myelodysplastic Syndromes or Secondary Acute Myeloid Leukemia", *Journal of Clinical Oncology*, Vol. 28 No. 3 .
- [8] Bryan Liz, (2012) "Reduced intensity regimen prior to marrow transplant better for older leukemia patients", Ohio State University Medical Center, American Society of Hematology (ASH) Annual Meeting in Atlanta, GA., 443-506:1931.
- [9] Petrie A., Sabin C., (2009), *Medical Statistics at a Glance*, 3rd Edition, Wiley-Blackwell.
- [10] Peacock J., and Peacock P., (2011), *Oxford Handbook of Medical Statistics*, Oxford University Press.
- [11] Noman, B., Al Imam W., Amasha H., (2016), "Correlation between Blood Analysis Indicators (Independent Variables) and Results of Acute Myeloid Leukemia Treatment using Stem Cells", *Damascus University Journal for Engineering Sciences*.
- [12] Sivanandam S. N, Sumathhi S, and Deepa S. N., (2006) , " *Introduction to neural networks using Matlab 6.0.* " , Tata McGraw-Hill Education.
- [13] Giovanni Caocci, (2010) "Comparison between an artificial neural network and logistic regression in predicting acute graft-vs-host disease after unrelated hematopoietic stem cell transplantation in thalassemia patients", Elsevier, *Journal for Hematology, Stem cell Biology and Transplantation*, Volume 38, Issue 5, Pages 426–433.

[14] Grimwade D., (2009), Prospective Minimal Residual Disease Monitoring to Predict Relapse of Acute Promyelocytic Leukemia and to Direct Pre-Emptive Arsenic Trioxide Therapy., American J. of Clinical Oncology., vol 27, no. 22.