JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ
العدد السابع والاربعون
(٣٣٩) (٣٥٥)

تحليل السلاسل الزمنية في الطب الحيوي والصحة العامة في محافظة صلاح الدين

م.م. عائدة عبد الحسين محمد
المديرية العامة لتربية صلاح الدين/ الكلية التربوية المفتوحة، فرع بلد، العراق
ayeda.mohammed@st.tu.edu.iq
د. انتصار أحمد محمد
المديرية العامة لتربية صلاح الدين/ قسم تربية بلد، إعدادية بلد للبنات، العراق
eahmed.math@gmail.com | ORCID

المستخلص:

تهدف هذه الدراسة إلى تحليل الاتجاهات الزمنية لمجموعة من المؤشرات الصحية والديمغرافية في محافظة صلاح الدين خلال المدة ٢٠١٠–٢٠٢٢، وتحديد طبيعة ارتباطها بالنمو السكاني ومعدل الوفيات، وبناء نموذج يفسر التغير في الوفيات اعتمادًا على مؤشرات المرض المتاحة. اعتمدت الدراسة على ١٣ مشاهدة سنوية لالتهاب الكبد الفيروسي والملاريا والتسمم الغذائي ومرض السكري وأمراض القلب والأوعية الدموية، إلى جانب النمو السكاني والوفيات. استخدم التحليل الإحصاء الوصفي، ومعامل ارتباط بيرسون، والانحدار الخطي المتعدد، ومقاييس دقة الملاءمة التنبؤية. أظهرت النتائج أن أعلى متوسط سنوي سُجل لأمراض القلب والأوعية الدموية (٦٢٦٥.٦٩٢)، تلاها التهاب الكبد الفيروسي (٣٣٩٧.٨٤٦). وبلغ ارتباط النمو السكاني بمرض السكري =٠.٥٥r وبأمراض القلب والأوعية الدموية =٠.٦٩r، بينما بلغ ارتباط الوفيات بالملاريا =٠.٥٦r وبالسكري =٠.٥٥r. أوضح نموذج الانحدار قدرة تفسيرية مرتفعة نسبيًا (R^2=٠.٨٩٥، و$Adjusted\ R^2$=٠.٨١٩، و=F١١.٨٦٨، و$p\approx$٠.٠٠٠٣)، وكانت الملاريا والتسمم الغذائي والسكري وأمراض القلب والأوعية الدموية ذات معاملات دالة إحصائيًا عند مستوى ٠.٠٥، في حين لم يكن معامل التهاب الكبد الفيروسي دالًا. كما بلغ RMSE=٢٢.٠٦٩ و=MAE١٢.٨١٢ و=MAPE٩.٠١٦%. تشير النتائج إلى تقلبات زمنية واضحة في العبء المرضي وإلى أهمية دمج

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

المؤشرات الصحية والديمغرافية عند التخطيط للصحة العامة، مع ضرورة تفسير العلاقات بوصفها ارتباطات إحصائية لا علاقات سببية.

الكلمات المفتاحية: السلاسل الزمنية؛ التنبؤ الصحي؛ الصحة العامة؛ الأمراض المعدية؛ الأمراض المزمنة؛ الانحدار الخطي المتعدد.

Health Forecasting Using Time Series Analysis in Biomedicine and Public Health: A Case Study of Salah al-Din Governorate

Assis .Lec. Aida Abdulhssein Mohammed
Salah al-Din General Directorate of Education / Open Educational College, Balad Branch, Iraq
Dr.ayeda.mohammed@st.tu.edu.iq
Intisar Ahmed Mohammed
Salah al-Din General Directorate of Education / Balad Education Department, Balad Preparatory School for Girls, Iraq
eahmed.math@gmail.com | ORCID

Abstract

This study analyzes temporal trends in selected health and demographic indicators in Salah al-Din Governorate during 2010–2022, examines their associations with population growth and mortality, and develops a model explaining variation in mortality using the available disease indicators. The analysis used 13 annual observations for viral hepatitis, malaria, food poisoning, diabetes, cardiovascular disease, population growth, and mortality. Descriptive statistics, Pearson correlation, multiple linear regression, and predictive-fit accuracy measures were applied. Cardiovascular disease had the highest annual mean (6265.692), followed by viral hepatitis (3397.846). Population growth correlated positively with diabetes (r=0.55) and cardiovascular disease (r=0.69), while mortality correlated with malaria (r=0.56) and diabetes (r=0.55). The regression model showed relatively high explanatory power (R^2=0.895; adjusted R^2=0.819; F=11.868; p≈0.003). Malaria, food poisoning, diabetes, and cardiovascular disease had statistically significant coefficients at the 0.05 level, whereas viral hepatitis was not significant. In-sample fit measures were

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

RMSE=22.069, MAE=12.812, and MAPE=9.016%. The findings reveal substantial temporal variability in disease burden and support integrating health and demographic indicators in public-health planning, while interpreting the observed relationships as statistical associations rather than causal effects.

Keywords: Time Series Analysis; Health Forecasting; Public Health; Infectious Diseases; Chronic Diseases; Multiple Linear Regression.

١. المقدمة

يمثل تحليل السلاسل الزمنية أداة أساسية في دراسات الصحة العامة لأنه يتيح تتبع تغير مؤشرات المرض والوفيات عبر الزمن، والكشف عن الاتجاهات العامة والتقلبات غير المنتظمة، وتقييم مدى قدرة النماذج الإحصائية على وصف القيم المستقبلية أو المقدرة. ويكتسب هذا النوع من التحليل أهمية أكبر عندما تتغير أنماط الأمراض بسرعة أو عندما تتزامن التغيرات الصحية مع تحولات سكانية وبيئية. وتؤكد الأدبيات الحديثة أن اختيار نموذج التنبؤ يجب أن يعتمد على خصائص السلسلة الزمنية وطولها وتكرارها وطبيعة المتغيرات المرافقة، لا على تفضيل نموذج واحد بصورة مسبقة [١–٧].

أظهرت دراسات حديثة في الصحة العامة تنوعًا واضحًا في طرائق التنبؤ. فقد قارن Rahman وزملاؤه بين ARIMA وXGBoost في بيانات كوفيد–١٩ في بنغلادش، وأكدوا أهمية تقييم الدقة باستخدام MAE وRMSE و MAPE[١]. كما أوضح Somyanonthanakul وزملاؤه أن إدخال متغيرات تفسيرية في نموذج ARIMAX يمكن أن يحسن التنبؤ بعدد الحالات مقارنة بنموذج ARIMA الأحادي [٢]. وفي دراسة لانتشار جدري القردة، أظهر Wei وزملاؤه أن أداء ARIMA والتنعيم الأسي والنموذج الرمادي وLSTM يختلف باختلاف الدولة وخصائص البيانات [٣].

وفي مجال العوامل البيئية، بينت دراسة Zheng وزملائه أن دمج مؤشرات بيئية متأخرة زمنيًا داخل ARIMAX حسّن التنبؤ بحالات الإنفلونزا [٤]. كما قدم Jain وزملاؤه نموذجًا هجينًا LSTM–ARIMA حقق دقة تنبؤية أفضل من عدة نماذج منفردة في بيانات كوفيد–١٩ [٥]. وأظهرت دراسة Deng وزملائه في عام ٢٠٢٥ أن مقارنة النماذج التقليدية والحديثة ضرورية عند التنبؤ بوفيات أمراض القلب، لأن الأداء يتغير تبعًا لبنية البيانات وأفق التنبؤ [٦].

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

على الرغم من هذا التطور، ما تزال الأدلة المحلية التي تجمع في تحليل واحد بين الأمراض المعدية والأمراض المزمنة والمؤشرات الديمغرافية في محافظة صلاح الدين محدودة. وتتمثل الفجوة التي تعالجها الدراسة الحالية في تحليل سلسلة سنوية موحدة للفترة ٢٠١٠–٢٠٢٢، وربط التغير في عدد من مؤشرات المرض بالنمو السكاني والوفيات، ثم قياس القدرة التفسيرية والملاءمة التنبؤية لنموذج انحدار متعدد. ويسمح هذا النهج بتقديم قراءة كمية مباشرة للبيانات المتاحة مع تجنب المبالغة في تفسير الارتباطات على أنها علاقات سببية.

١.١ مشكلة البحث والفجوة البحثية

تتمثل مشكلة البحث في وجود تقلبات كبيرة من سنة إلى أخرى في أعداد الإصابات بالأمراض المعدية والمزمنة وفي المؤشرات الديمغرافية بمحافظة صلاح الدين، من دون عرض تكاملي يوضح مقدار هذه التقلبات وعلاقاتها المتبادلة وأثرها الإحصائي في معدل الوفيات. لذلك تسعى الدراسة إلى الإجابة عن سؤال رئيس: إلى أي مدى ترتبط مؤشرات المرض بالنمو السكاني والوفيات خلال ٢٠١٠–٢٠٢٢، وما مقدار التغير في معدل الوفيات الذي يمكن تفسيره عند إدخال مؤشرات المرض معًا في نموذج انحدار متعدد؟

١.٢ أهداف البحث

١. وصف الاتجاهات السنوية لمؤشرات التهاب الكبد الفيروسي والملاريا والتسمم الغذائي والسكري وأمراض القلب والأوعية الدموية خلال ٢٠١٠–٢٠٢٢.

٢. قياس التغير في النمو السكاني والوفيات خلال الفترة نفسها.

٣. تقدير معاملات الارتباط بين المؤشرات الصحية وبين النمو السكاني والوفيات.

٤. بناء نموذج انحدار خطي متعدد لتفسير معدل الوفيات اعتمادًا على مؤشرات المرض الخمسة.

٥. تقييم الملاءمة التنبؤية للنموذج باستخدام RMSE وMAE وMAPE وsMAPE ومعامل ثايل.

١.٣ فرضيات البحث

H_1: توجد علاقات ارتباطية بين بعض مؤشرات الأمراض وبين التغير الديمغرافي في محافظة صلاح الدين خلال ٢٠١٠–٢٠٢٢.

H_2: تسهم مؤشرات الأمراض المعدية والمزمنة مجتمعة في تفسير جزء دال إحصائيًا من التباين في معدل الوفيات.

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

٢. الدراسات السابقة

Rahman et al.(٢٠٢٢): استخدموا بيانات يومية لحالات ووفيات كوفيد-١٩ في بنغلادش لمقارنة ARIMA وXGBoost، وقيموا الأداء باستخدام MAE وMPE وRMSE وMAPE. أظهرت النتائج أن الأداء يعتمد على بنية السلسلة ومجموعة الاختبار، وهو ما يؤكد ضرورة استخدام أكثر من مقياس للحكم على دقة التنبؤ [١].

Somyanonthanakul et al.(٢٠٢٢): حللوا بيانات ٣٦٨٥ مريضًا بكوفيد-١٩، وقارنوا ARIMA مع ARIMAX بعد إضافة عوامل تنبؤية مختارة. انخفض RMSE من ٢٩.٧٦٠٥ إلى ٢٧.٧٥٠٨، وانخفض MAE من ٢٧.٥١٠٢ إلى ٢٣.٤٦٤٢ عند استخدام ARIMAX، مما يوضح فائدة المتغيرات الخارجية في تحسين التنبؤ [٢].

Wei et al.(٢٠٢٣): قارنوا ARIMA والتنعيم الأسي و(GM(١,١ وLSTM لتوقع اتجاهات وباء جدري القردة. بينت النتائج أن النموذج الأفضل اختلف بين السلاسل المدروسة، وهو ما يدعم مبدأ مواءمة الطريقة مع خصائص البيانات [٣].

Zheng et al.(٢٠٢٤): درسوا ٢١,٥٧٣ حالة إنفلونزا وربطوا الانتشار بمؤشرات بيئية باستخدام ARIMAX. أدى إدخال PM١٠ المتأخر زمنيًا إلى تحسين الأداء التنبؤي، مع RMSE بلغ ١٢.٠٣٣ في النموذج الأفضل [٤].

Jain et al.(٢٠٢٤): اقترحوا نموذجًا هجينًا يجمع ARIMA وLSTM لتوقع حالات كوفيد-١٩، وحقق النموذج الهجين MAPE قدره ٢.٤% مقابل ٣.٦% لنموذج LSTM، ما يشير إلى إمكانية الاستفادة من الدمج بين البنية الخطية وغير الخطية [٥].

Deng et al.(٢٠٢٥): قارنوا نماذج ARIMA وProphet وXGBoost وLSTM في التنبؤ بوفيات مرض القلب الناتج عن ارتفاع ضغط الدم لدى النساء في سن الإنجاب. وأكدت نتائجهم أن تقييم النماذج على فترة تحقق مستقلة أكثر أهمية من الاعتماد على الملاءمة داخل العينة وحدها [٦].

توضح هذه الدراسات أن التنبؤ الصحي الحديث يتجه إلى المقارنة بين نماذج متعددة وإدخال متغيرات تفسيرية عند توافرها. وتمتاز الدراسة الحالية بأنها لا تركز على وباء واحد، بل تربط عدة مؤشرات مرضية معدية ومزمنة بمتغيرات ديمغرافية ضمن سياق محلي واحد، ثم تقيس الارتباط والانحدار والملاءمة التنبؤية بصورة متكاملة.

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

٣. المنهجية الإحصائية

٣.١ تصميم الدراسة والبيانات

اعتمدت الدراسة تصميمًا تحليليًا زمنيًا استعاديًا على مستوى المحافظة. تتكون قاعدة التحليل من ١٣ مشاهدة سنوية للفترة ٢٠١٠–٢٠٢٢. وتشمل المتغيرات: التهاب الكبد الفيروسي (HV)، والملاريا (MAL)، والتسمم الغذائي (FP)، ومرض السكري (DM)، وأمراض القلب والأوعية الدموية (CVD)، ومعدل النمو السكاني، ومعدل الوفيات (DR). استخدمت القيم السنوية المتاحة في ملف الدراسة كما هي، مع توحيد سنوات المقارنة بين جميع السلاسل.

٣.٢ الإحصاء الوصفي

استخدم المتوسط الحسابي لقياس المستوى المركزي لكل سلسلة، والانحراف المعياري لقياس مقدار التشتت بين السنوات، إضافة إلى الحدين الأدنى والأعلى لتحديد مدى التقلب الزمني. حُسب المتوسط والانحراف المعياري وفق الصيغتين الآتيتين:

$$\bar{x} = (1/n)\ \Sigma_{t=1}^{n}\ x_t$$

$$s = \sqrt{[\ \Sigma_{t=1}^{n}\ (x_t - \bar{x})^2 / (n - 1)\ ]}$$

٣.٣ معامل الارتباط

استُخدم معامل ارتباط بيرسون لقياس اتجاه وقوة العلاقة الخطية الثنائية بين كل مؤشر مرضي وكل من النمو السكاني والوفيات. تتراوح قيمة r بين −١ و+١، وتدل الإشارة على اتجاه العلاقة بينما يدل المقدار المطلق على قوتها.

$$r_{xy} = \Sigma(x_t - \bar{x})(y_t - \bar{y}) / \sqrt{[\Sigma(x_t - \bar{x})^2\ \Sigma(y_t - \bar{y})^2]}$$

٣.٤ نموذج الانحدار الخطي المتعدد

استُخدم معدل الوفيات بوصفه المتغير التابع، وأُدخلت مؤشرات الأمراض الخمسة بوصفها متغيرات مستقلة. يسمح هذا النموذج بتقدير الأثر الجزئي لكل مؤشر بعد تثبيت بقية المؤشرات، وهو ما يختلف عن الارتباط الثنائي. صيغ النموذج كما يأتي:

$$DR_t = \beta_0 + \beta_1 HV_t + \beta_2 MAL_t + \beta_3 FP_t + \beta_4 DM_t + \beta_5 CVD_t + \varepsilon_t$$

قُيمت دلالة النموذج الكلية باختبار F، ودلالة كل معامل باختبار t عند مستوى α=٠.٠٥، وعُرضت قيم R^2 و$Adjusted\ R^2$ لقياس نسبة التباين المفسر مع مراعاة عدد المتغيرات. وبسبب

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

محدودية عدد المشاهدات السنوية، تفسر النتائج بحذر وتستخدم أساسًا لوصف العلاقات داخل الفترة المدروسة.

٣.٥ تقييم الملاءمة التنبؤية

قورنت القيم المقدرة من نموذج الانحدار بالقيم الفعلية لمعدل الوفيات داخل العينة. استُخدمت RMSE وMAE وMAPE وsMAPE ومعامل ثايل وTheil U٢. تشير القيم الأصغر عمومًا إلى تقارب أكبر بين القيم المقدرة والفعلية. ويشار إلى هذه النتائج في البحث بوصفها ملاءمة تنبؤية داخل العينة، وليست اختبارًا خارجيًا لتنبؤات مستقبلية مستقلة.

$$RMSE = \sqrt{[(1/n)\ \Sigma^{t}_{=1}{}^{n}\ (y_t - \hat{y}_t)^2]}$$

$$MAE = (1/n)\ \Sigma^{t}_{=1}{}^{n}\ |y_t - \hat{y}_t|$$

$$MAPE = (100/n)\ \Sigma^{t}_{=1}{}^{n}\ |(y_t - \hat{y}_t)/y_t|$$

٤. الأساس النظري لتحليل السلاسل الزمنية

السلسلة الزمنية هي مجموعة مشاهدات مرتبة بحسب الزمن، ويكون ترتيبها جزءًا من المعلومة الإحصائية نفسها. وقد تتضمن السلسلة اتجاهًا طويل الأمد، أو تغيرات موسمية، أو دورات، أو تقلبات عشوائية. عندما تكون البيانات سنوية وعدد السنوات محدودًا، تزداد أهمية الفحص الوصفي والرسم البياني قبل الانتقال إلى نماذج أكثر تعقيدًا [٨].

في تطبيقات الصحة العامة، لا يكفي وصف متوسط عدد الحالات فقط، لأن ارتفاع الانحراف المعياري أو وجود قمم حادة في سنوات معينة قد يكشف عبئًا صحيًا غير مستقر. كما يساعد الارتباط في تحديد المتغيرات التي تتحرك معًا، بينما يسمح الانحدار المتعدد بدراسة العلاقة الجزئية لكل متغير مع النتيجة بعد التحكم الإحصائي في المتغيرات الأخرى. ومع ذلك، لا يثبت الارتباط أو الانحدار السببية ما لم يدعم التصميم البحثي ذلك.

تستخدم مقاييس الخطأ التنبؤي لتقييم مقدار الفرق بين القيم الفعلية والمقدرة. ويعد RMSE أكثر حساسية للأخطاء الكبيرة بسبب تربيع الخطأ، بينما يقدم MAE متوسطًا مباشرًا لحجم الخطأ، ويعرض MAPE الخطأ بصورة نسبية مئوية. ولهذا فإن استخدام مجموعة من المقاييس يعطي صورة أوضح من الاعتماد على مقياس واحد [١,٨].

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

٥. النتائج

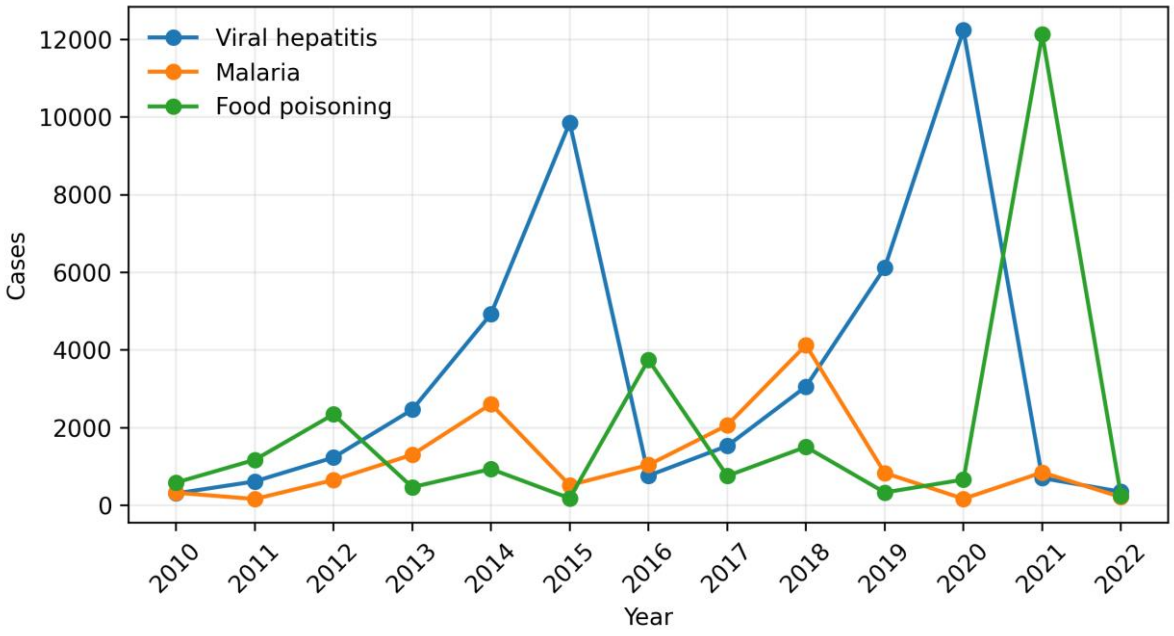
تعرض النتائج أولًا الخصائص الوصفية للسلاسل السنوية، ثم معاملات الارتباط، ثم نموذج الانحدار وملاءمته التنبؤية. جرى توحيد ترقيم الجداول والأشكال والإشارة إليها داخل النص.

الجدول (١). الإحصاء الوصفي للأمراض المعدية في محافظة صلاح الدين خلال ٢٠١٠–٢٠٢٢

المرض	المتوسط	الانحراف المعياري	أدنى قيمة (السنة)	أعلى قيمة (السنة)
التهاب الكبد الفيروسي	3397.846	3868.435	308 (2010)	12240 (2020)
الملاريا	1143.923	1161.701	163 (2011)	4128 (2018)
التسمم الغذائي	1928.615	3222.194	182 (2015)	12128 (2021)

يبين الجدول (١) أن التهاب الكبد الفيروسي سجل أعلى متوسط بين الأمراض المعدية المدروسة، مع انحراف معياري يتجاوز المتوسط، وهو ما يعكس تباينًا سنويًا كبيرًا. وتظهر السلسلة قمة واضحة في عام ٢٠٢٠. كما اتسم التسمم الغذائي بتقلب شديد بسبب الارتفاع الكبير في عام ٢٠٢١، في حين بلغ أعلى عدد لحالات الملاريا في عام ٢٠١٨. لذلك فإن المتوسطات وحدها لا تكفي لوصف هذه السلاسل، ويجب تفسيرها إلى جانب التشتت والشكل الزمني.

الشكل (١). التغير السنوي في مؤشرات الأمراض المعدية خلال ٢٠١٠–٢٠٢٢



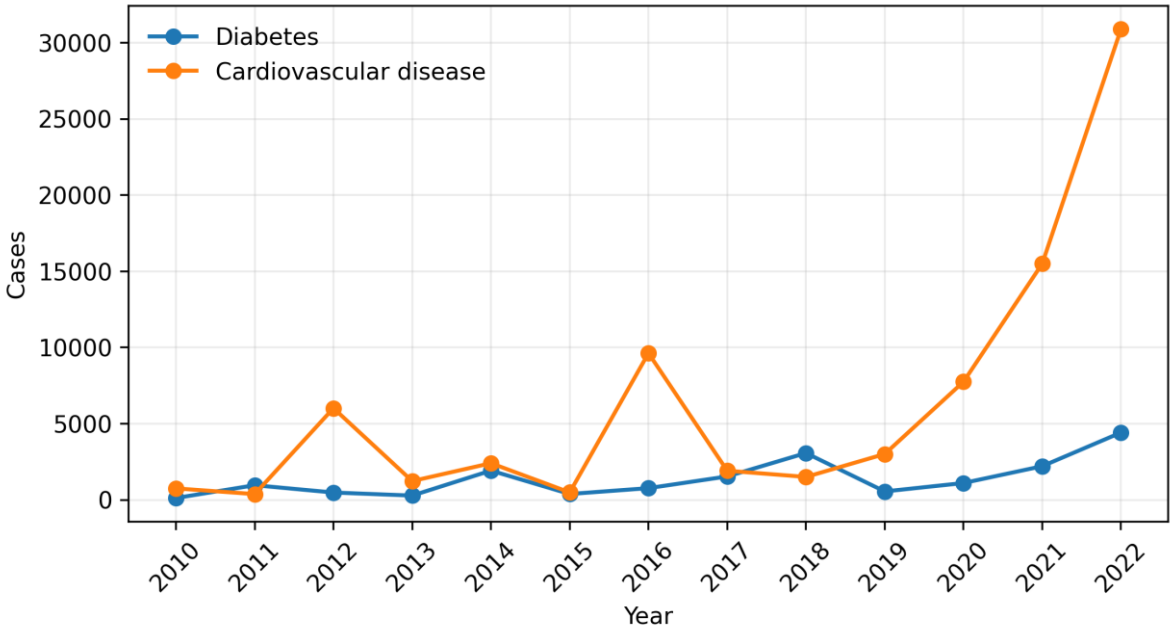
JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

الجدول (٢). الإحصاء الوصفي لمرض السكري وأمراض القلب والأوعية الدموية خلال ٢٠١٠–٢٠٢٢

المرض	المتوسط	الانحراف المعياري	أدنى قيمة (السنة)	أعلى قيمة (السنة)
مرض السكري	1369.231	1258.057	120 (2010)	4416 (2022)
أمراض القلب والأوعية الدموية	6265.692	8641.795	376 (2011)	30892 (2022)

يوضح الجدول (٢) أن أمراض القلب والأوعية الدموية تمثل السلسلة الأعلى متوسطًا والأكثر تشتتًا بين المؤشرات المرضية المدروسة، وقد بلغت أعلى قيمة لها ٣٠,٨٩٢ في عام ٢٠٢٢. كذلك سجل السكري أعلى قيمة في ٢٠٢٢. وتشير هذه الزيادات المتزامنة في نهاية الفترة إلى ضرورة متابعة العبء المزمن مع التحولات السكانية، مع عدم تفسير التزامن وحده كعلاقة سببية.

الشكل (٢). التغير السنوي في مرض السكري وأمراض القلب والأوعية الدموية خلال ٢٠١٠–٢٠٢٢



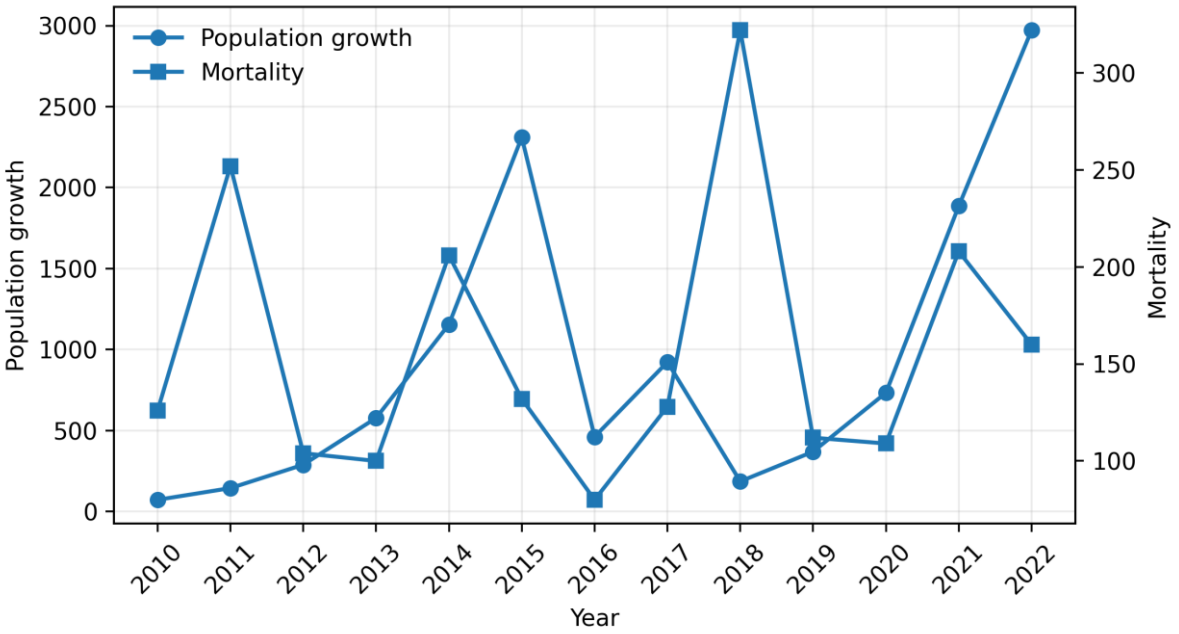
JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

الجدول (٣). الإحصاء الوصفي للمؤشرات الديمغرافية خلال ٢٠١٠–٢٠٢٢

المؤشر	المتوسط	الانحراف المعياري	أدنى قيمة (السنة)	أعلى قيمة (السنة)
معدل النمو السكاني	929.239	915.435	72.21 (2010)	2972.16 (2022)
معدل الوفيات	156.846	70.713	80 (2016)	322 (2018)

يشير الجدول (٣) إلى تباين كبير في معدل النمو السكاني خلال الفترة، إذ بلغ أعلى مستوى في ٢٠٢٢. أما الوفيات فقد بلغت أعلى قيمة في ٢٠١٨ وأدنى قيمة في ٢٠١٦. ويبين الشكل (٣) أن مساري النمو والوفيات لا يتحركان دائمًا في الاتجاه نفسه، وهو ما يتفق مع معامل الارتباط الضعيف بينهما.

الشكل (٣). التغير السنوي في النمو السكاني ومعدل الوفيات خلال ٢٠١٠–٢٠٢٢



JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

الجدول (٤). معاملات ارتباط بيرسون بين مؤشرات المرض والنمو السكاني والوفيات

المؤشر المرضي	الارتباط مع النمو السكاني r	الارتباط مع الوفيات r
التهاب الكبد الفيروسي	0.12	−0.18
الملاريا	−0.22	0.56
التسمم الغذائي	0.18	0.18
مرض السكري	0.55	0.55
أمراض القلب والأوعية الدموية	0.69	−0.06

تظهر النتائج في الجدول (٤) أن أقوى ارتباط للنمو السكاني كان مع أمراض القلب والأوعية الدموية (r=٠.٦٩)، يليه السكري (r=٠.٥٥). أما الوفيات فكان ارتباطها موجبًا متوسط القوة تقريبًا مع الملاريا (r=٠.٥٦) والسكري (r=٠.٥٥)، بينما كان ارتباطها بأمراض القلب والأوعية الدموية قريبًا من الصفر (r=−٠.٠٦). وتؤكد هذه النتائج أن اتجاه العلاقة الثنائية قد يختلف عن معامل الانحدار الجزئي بعد إدخال بقية المتغيرات في نموذج واحد.

مجلة العلوم الاساسية
للعلوم التربوية والنفسية وطرائق التدريس للعلوم الاساسية

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

الجدول (٥). نتائج الانحدار الخطي المتعدد لمعدل الوفيات

المتغير		المعامل B	الخطأ المعياري	t	p
الثابت		20.124	133.076	6.613	<0.001
التهاب الكبد الفيروسي		0.002	−0.0021	−0.846	0.426
الملاريا		0.015	−0.0372	−2.560	0.038
التسمم الغذائي		0.003	0.0082	2.702	0.031
مرض السكري		0.018	0.1074	5.945	0.001
أمراض القلب والأوعية الدموية		0.003	−0.0143	−5.248	0.001

ملخص النموذج: R^2=٠.٨٩٥، Adjusted R^2=٠.٨١٩، F(٥,٧)=١١.٨٦٨، p≈٠.٠٠٠٣. وبذلك يفسر النموذج نحو ٨٩.٥% من التباين المرصود في معدل الوفيات داخل العينة، مع بقاء ١٠.٥% تقريبًا غير مفسر بواسطة المتغيرات المدرجة. كانت معاملات الملاريا والتسمم الغذائي والسكري وأمراض القلب والأوعية الدموية دالة عند مستوى ٠.٠٥، بينما لم يكن التهاب الكبد الفيروسي دالًا (p=٠.٤٢٦).

تأخذ المعادلة المقدرة الشكل الآتي:

$$\widehat{DR}_t = 133.076 - 0.0021HV_t - 0.0372MAL_t + 0.0082FP_t + 0.1074DM_t - 0.0143CVD_t$$

لا ينبغي تفسير إشارات المعاملات بوصفها آثارًا سببية مباشرة. فالمعامل السالب للملاريا أو أمراض القلب يمثل علاقة جزئية مشروطة ببقية المتغيرات داخل عينة صغيرة من ١٣ سنة، وقد يتأثر بالتداخل بين السلاسل وبالتقلبات الحادة في بعض السنوات.

مجلة العلوم الاساسية
للعلوم التربوية والنفسية وطرائق التدريس للعلوم الاساسية

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

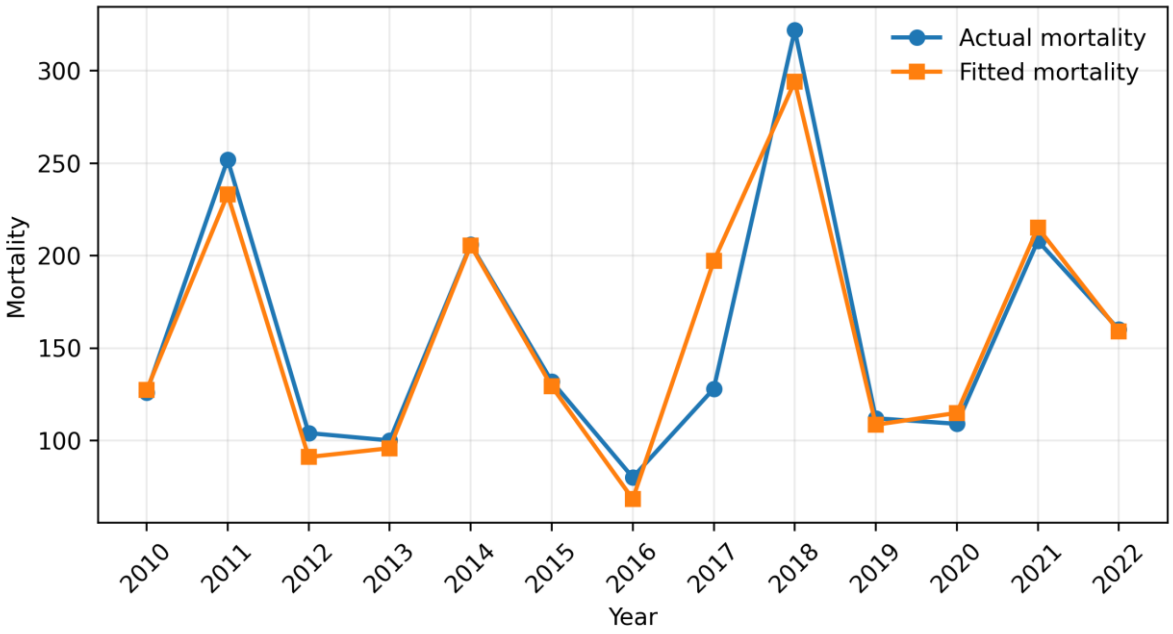
الجدول (٦). مقاييس الملاءمة التنبؤية لمعدل الوفيات داخل العينة

المقياس	القيمة
RMSE	22.06870
MAE	12.81195
MAPE	9.015764%
sMAPE	8.335723%
Theil inequality coefficient	0.064827
Theil U2	0.352491
Bias proportion	0.000000
Variance proportion	0.027870
Covariance proportion	0.972130

تعكس قيمة MAPE البالغة نحو ٩.٠٢% تقاربًا جيدًا نسبيًا بين القيم الفعلية والمقدرة داخل الفترة المدروسة. كما أن انخفاض معامل ثايل واقتراب نسبة الانحياز من الصفر يشيران إلى أن الخطأ لا ينتج أساسًا عن انحياز ثابت في التقدير. ومع ذلك، تبقى هذه المقاييس داخل العينة ولا تمثل تحققًا خارجيًا على سنوات جديدة.

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

الشكل (٤). مقارنة معدل الوفيات الفعلي بالقيم المقدرة من نموذج الانحدار



٦. المناقشة

تكشف النتائج عن تفاوت زمني كبير في مؤشرات الأمراض بمحافظة صلاح الدين. كان التشتت مرتفعًا بصورة خاصة في التهاب الكبد الفيروسي والتسمم الغذائي وأمراض القلب والأوعية الدموية، وظهرت قمم حادة في سنوات محددة. هذه السمة مهمة منهجيًا لأن السلاسل ذات القمم الحادة قد تعطي متوسطات مضللة إذا لم تُقرأ مع الانحراف المعياري والرسم الزمني. كما أنها تعزز الحاجة إلى بيانات أكثر تكرارًا، مثل البيانات الشهرية أو الفصلية، عندما يكون الهدف بناء نماذج تنبؤ مستقبلية مستقرة.

أظهر النمو السكاني ارتباطًا موجبًا مع السكري (r=٠.٥٥) وأمراض القلب والأوعية الدموية (r=٠.٦٩). ويمكن فهم ذلك بوصفه حركة متزامنة في السلاسل قد تعكس زيادة عدد السكان المعرضين للخطر أو تغير أنماط التسجيل والخدمات الصحية. إلا أن البيانات المجمعة سنويًا لا تسمح بإثبات أن النمو السكاني سبب مباشر لزيادة المرض. تتفق هذه القراءة الحذرة مع الأدبيات الحديثة التي تؤكد أهمية إدخال متغيرات مفسرة مناسبة عند توافرها بدل الاعتماد على الزمن وحده [٢,٤].

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

بلغ ارتباط الوفيات بالملاريا =r٠.٥٦، وبالسكري =r٠.٥٥، لكن معامل الملاريا أصبح سالبًا ودالًا في نموذج الانحدار المتعدد. لا يمثل ذلك تناقضًا إحصائيًا؛ فالارتباط يقيس العلاقة الثنائية، بينما يقدّر معامل الانحدار العلاقة الجزئية بعد تثبيت مؤشرات المرض الأخرى. ومع صغر العينة ووجود ارتباطات بين بعض المتنبئات، يمكن أن تتغير الإشارة والمقدار. لذلك يجب التركيز على قدرة النموذج الكلية وعلى عدم تحويل المعاملات الجزئية إلى استنتاجات سببية.

أظهر النموذج R^2=٠.٨٩٥ و$Adjusted\ R^2$=٠.٨١٩، وبلغ MAPE نحو ٩.٠٢%. تعكس هذه القيم ملاءمة قوية نسبيًا للبيانات المستخدمة، لكنها لا تعني بالضرورة أن النموذج سيحافظ على الدقة نفسها في سنوات لاحقة. وتبرز هنا أهمية المقارنة مع الدراسات الحديثة؛ فقد استخدمت دراسة Rahman وزملائه تقسيمات تدريب واختبار متعددة عند مقارنة ARIMA و XGBoost[١]، كما اعتمدت دراسة Deng وزملائه فترة تحقق مستقلة للمقارنة بين نماذج مختلفة [٦]. وبالتالي فإن الخطوة المنهجية التالية لهذه الدراسة هي اختبار النموذج على بيانات لاحقة أو باستخدام تحقق زمني مناسب.

تتفق النتائج العامة مع ما توصلت إليه دراسات Wei و Jain وغيرهما من أن أداء النماذج يعتمد على طبيعة السلسلة، وأن النماذج الهجينة أو الحديثة قد تقدم تحسينًا عندما تكون البيانات طويلة وكثيفة [٣,٥]. إلا أن السلسلة الحالية سنوية وتضم ١٣ مشاهدة فقط؛ لذلك كان استخدام نموذج خطي تفسيري محدود المعلمات أكثر ملاءمة من إدخال شبكة عميقة عالية التعقيد. كما تتفق أهمية العوامل الخارجية مع Zheng وزملائه، الذين أظهروا أن إدخال المؤشرات البيئية داخل ARIMAX حسن توقع حالات الإنفلونزا [٤].

تتمثل أهم محددات الدراسة في قصر السلسلة الزمنية، واعتماد بيانات سنوية مجمعة، وعدم توافر متغيرات بيئية واجتماعية وخدمية ضمن النموذج الحالي. كما أن التنبؤ المقيم هنا داخل العينة، ولا توجد مجموعة اختبار مستقلة. ولهذا تمثل النتائج أساسًا وصفيًا وتفسيريًا مهمًا للسياق المحلي، بينما تحتاج قابلية التعميم والدقة المستقبلية إلى بيانات أطول وتحقق خارجي.

٧. الاستنتاجات

١. تتسم مؤشرات الأمراض في محافظة صلاح الدين خلال ٢٠١٠–٢٠٢٢ بتقلبات سنوية واضحة، وكان العبء المتوسط الأكبر لأمراض القلب والأوعية الدموية، في حين ظهرت قمم حادة في التهاب الكبد الفيروسي والتسمم الغذائي.

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ

٢. يرتبط النمو السكاني بصورة موجبة متوسطة إلى قوية نسبيًا مع السكري (r=٠.٥٥) وأمراض القلب والأوعية الدموية (r=٠.٦٩)، ما يبرز أهمية دمج المؤشرات الديمغرافية في المراقبة الصحية.

٣. بلغ ارتباط الوفيات بالملاريا r=٠.٥٦ وبالسكري r=٠.٥٥، بينما كانت علاقته بأمراض القلب والأوعية الدموية ضعيفة ثنائيًا (r=−٠.٠٦).

٤. حقق نموذج الانحدار المتعدد R^2=٠.٨٩٥ و$Adjusted\ R^2$=٠.٨١٩، وكان النموذج ككل دالًا إحصائيًا (F=١١.٨٦٨، p≈٠.٠٠٣).

٥. أظهرت مقاييس الملاءمة داخل العينة RMSE=٢٢.٠٦٩ وMAE=١٢.٨١٢ وMAPE=٩.٠١٦%، وهي نتائج تدعم فائدة النموذج في وصف الفترة المدروسة، لكنها تحتاج إلى تحقق خارجي قبل استخدامه للتنبؤ المستقبلي.

٨. التوصيات

١. تطوير قاعدة بيانات صحية زمنية موحدة في المحافظة مع تحديث شهري أو فصلي بدل الاقتصار على القيم السنوية.

٢. إدخال المتغيرات البيئية مثل درجة الحرارة وجودة الهواء والأمطار، والمتغيرات الديمغرافية والخدمية، في نماذج مستقبلية من نوع ARIMAX أو النماذج الديناميكية الأخرى.

٣. مقارنة الانحدار الحالي بنماذج ARIMA وETS وProphet وطرائق التعلم الآلي عند توافر سلسلة أطول، مع استخدام تقسيم زمني للتدريب والاختبار.

٤. التركيز في التخطيط الصحي على الأمراض المزمنة ذات المتوسطات والقمم المرتفعة، ولا سيما أمراض القلب والأوعية الدموية والسكري، مع الاستمرار في مراقبة موجات الأمراض المعدية.

٥. تجنب تفسير الارتباطات الإحصائية على أنها علاقات سببية، وإجراء دراسات وبائية فردية أو مكانية عند الحاجة إلى تحديد عوامل الخطر.

٩. المصادر والمراجع

1. Rahman MS, Chowdhury AH, Amrin M. Accuracy comparison of ARIMA and XGBoost forecasting models in predicting the incidence of COVID-19 in Bangladesh. PLOS Global Public Health. 2022;2(5):e0000495. doi:10.1371/journal.pgph.0000495.

2. Somyanonthanakul R, Warin K, Amasiri W, Mairiang K, Mingmalairak C, Panichkitkosolkul W, et al. Forecasting COVID-19 cases using time

series modeling and association rule mining. BMC Medical Research Methodology. 2022;22:281. doi:10.1186/s12874-022-01755-x.

3. Wei W, Wang G, Tao X, Luo Q, Chen L, Bao X, et al. Time series prediction for the epidemic trends of monkeypox using the ARIMA, exponential smoothing, GM(1,1) and LSTM deep learning methods. Journal of General Virology. 2023;104(4):001839.

4. Zheng X, Chen Q, Sun M, Zhou Q, Shi H, Zhang X, Xu Y. Exploring the influence of environmental indicators and forecasting influenza incidence using ARIMAX models. Frontiers in Public Health. 2024;12:1441240. doi:10.3389/fpubh.2024.1441240.

5. Jain S, Agrawal S, Mohapatra E, Srinivasan K. A novel ensemble ARIMA-LSTM approach for evaluating COVID-19 cases and future outbreak preparedness. Health Care Science. 2024;3:409–425. doi:10.1002/hcs2.123.

6. Deng Q, Wang S, Lyu J. Comparative effectiveness analysis of univariate time-series forecasting models for disease mortality rates in the global burden of disease database: a case study of global hypertensive heart disease among women of childbearing age. Frontiers in Public Health. 2025;13:1681569. doi:10.3389/fpubh.2025.1681569.

7. Sardar I, Akbar MA, Leiva V, Alsanad A, Mishra P. Machine learning and automatic ARIMA/Prophet models-based forecasting of COVID-19: methodology, evaluation, and case study in SAARC countries. Stochastic Environmental Research and Risk Assessment. 2023;37(1):345–359. doi:10.1007/s00477-022-02307-x.

8. Hyndman RJ, Athanasopoulos G. Forecasting: Principles and Practice. 3rd ed. Melbourne: OTexts; 2021.

9. Box GEP, Jenkins GM, Reinsel GC. Time Series Analysis, Forecasting and Control. 3rd ed. Englewood Cliffs, NJ: Prentice Hall; 1994.

١٠. عبد الغفور جاسم سالم. دراسة السلسلة الزمنية لعدد حالات الإصابة بمرض سرطان الرئة في مدينة الموصل (١٩٨٠–١٩٩٠). مجلة تكريت للعلوم الصرفة. ٢٠٠٦;١١(٢):٣٨–٤١.

JOBS
مجلة العلوم الأساسية
Journal of Basic Science
Print -ISSN 2306-5249
Online-ISSN 2791-3279
العدد السابع والاربعون
٢٠٢٦م /١٤٤٨هـ