# Role of the Accuracy of Machine Learning in Predicting the Outcome of Methotrexate Treatment in Plaque Psoriasis

**Sama Asim Sahib** MBChB**, Azzam Muhsin Al-Salami*** FIBMS**,**
**Nada Fadhil Al-Tajer**** CABD**, Sabeeh A Al-Mashhadani**** MSc

## ABSTRACT

**Background***:* Machine learning algorithm can be simply defined as giving data to the computer to answer a question, which was used recently by many medical branches including dermatology to predict the outcome of different treatment modalities, morbidity and mortality of several diseases with good results.
**Objective**: Evaluation the accuracy of Machine learning in predicting the outcome of Methotrexate treatment in plaque psoriasis.
**Patients and methods**: This is an observational analytic study that was conducted at the center of dermatology and venereology, Medical City in Baghdad, Iraq, from October 2018 to March 2020. Patients with plaque psoriasis who attended the outpatient clinic taking methotrexate drug were included in the study. Assessment of patient's demography, disease parameters, and drug parameters were done. All patients were followed up monthly for 3 months. Investigations, Psoriatic area and severity index score and side effect of medication were recorded at each visit and all of these data were arranged in Excel program and then introduced to machine learning program.
**Results**: A total of 208 patients, 140 (67.3%) male and 68(32.7%) females. Dataset processing was done by Amazon Machine learning and using binary model, it was found that if the error threshold is 0.5 then the accuracy of machine training phase is about 80% with sensitivity about 76% and specificity 82%, and when we increase the error threshold to 0.6, the accuracy reaches 84% with sensitivity about 71% and specificity about 93%.
**Conclusions**: Machine learning algorithm can be used to predict the response of methotrexate treatment in plaque psoriasis patients with accuracy exceeding 80%.
**Keywords: Binary model, Error threshold, Machine learning, Methotrexate, Psoriasis.**

***Iraqi Medical Journal Vol. 68, No. 2, July-Dec 2022; p. 73-82.***

Psoriasis vulgaris is a chronic inflammatory skin disease that classically presents with well demarcated, red plaques with silvery scale, commonly involving the scalp, elbows, knees, and presacral region, though any area of the skin may be involved, including the palms, soles, nails, and genitalia.(1)

Methotrexate (MTX) is a competitive inhibitor of dihydrofolate reductase, decreasing folate co-factors required for the synthesis of nucleic acids(2).

In addition, polyglutamate derivatives of methotrexate are potent inhibitors of 5-aminoimidazole-4-carboxamide ribonucleotide transformylase, an effect that results in increased amounts of endogenous adenosine, an anti-inflammatory molecule.(3,4)

These inhibitory actions have a large effect on rapidly dividing cells. In psoriasis, a theoretic inhibitory effect on keratinocytes was originally suspected to be responsible for its therapeutic benefit.(5)

Artificial intelligence (AI) is a branch of computer science that uses machines and programs to simulate intelligent human behavior.(6)

*Dept. of Otolaryngology, College of Medicine in Baghdad University.
**Dept. of Dermatology and Venereology, College of Medicine in Baghdad University.
Corresponding author: Dr. Sama Asim Sahib
E-mail: 18100049@arab-board.education

In 1970s, software engineers had created algorithms with explicit rules for computers on how processing data. Although, the human decision making in medicine were not easy to transferred to program into fixed rules. Machine learning (ML) is a tool that enables the goals of artificial intelligence to be achieved. Today, ML has aroused attention for its broad range of uses in our daily life from personalized online recommendations for videos and news to self-driving cars.[7]

Machine learning approaches can be divided into three categories:[8]

1. Supervised machine learning: The program is "trained" on a predefined set of "training examples", when we use labeled data (data with the outcome), which then facilitate its ability to reach an accurate conclusion when given new unlabeled data (data without the outcome), it most common type used in dermatology.

2. Unsupervised machine learning: The program is given a bunch of unlabeled data and must find patterns and relationships between them.

3. Reinforcement learning: is a hybrid of both supervised and unsupervised learning which learns by trial and error and input data.

Neural network and deep learning:[6] Artificial neural networks defined as collections of units, neurons that are similar to the network of neurons of human brain, which arranged in linear arrays called layers, so when creating artificial neural networks, we need to choose the number of layers, the number of the nodes in each layer and path of connection among the nodes, an example classification of pigmented lesion into melanoma and nevus.

While deep learning is a part of ML that uses computational models that are composed of multiple processing layers to learn the represented data.

Precision medicine used to develop targeted treatments based on data from multiomics platforms or other phenotypic or psychosocial characteristics this improve clinical outcomes and reduce unnecessary side effects for those less likely to respond to a certain treatment.[9]

In psoriasis several studies have created ML prediction models, for example a model to determine the long-term treatment response to biologics by assessment of baseline samples of gene expression, pathway associated with tumor necrosis factor (TNF) and major histocompatibility complex from skin biopsies[10]. Although it is invasive procedure by using skin biopsies for assessing gene expression, it is still a promising method to predict psoriasis response to anti-TNF biological drugs. For this, Tomalin et al created a ML predictive model from blood biochemical measurements rather than skin biopsies.[11]

In psoriatic arthritis, ML could be useful in developing quantitative assessment for psoriatic arthritis risk among psoriasis patients before symptoms appear.[10]

In May 2020, a study done by S Emam et al for predicting the long term outcome of biological therapy in patients with psoriasis using ML, when he used a data of 681 patients with psoriasis from the Danish registry cohort, DERMBIO.[12]

The aim of this study is to evaluate the accuracy of machine learning in predicting the response of methotrexate therapy in plaque psoriasis.

## ——————————Methods

This is an observational analytic retrospective study that was carried out at the Center of Dermatology and Venereology, Medical City in Baghdad, Iraq, from October 2018 to March 2020.

Two hundred seventy patients who attended the outpatient clinic with plaque psoriasis for whom MTX treatment was described in the period of the study, from those 62 patients were excluded from the study either because they lost from follow up or they had incomplete data due to discontinuation of the medication because of side effects or because they were not cooperative and refused the examination.

The datasets of all the patients were collected. Therapy outcome measures were categorized based on psoriasis area and severity index (PASI) score improvement from the baseline.

Inclusion criteria: All patients taking MTX treatment who had moderate to severe plaque psoriasis with one or both of the following: PASI score > 10 and/or body surface area > 20%.

Exclusion criteria: Patients who had or develop any one of the contra-indications to MTX therapy which include: Pregnancy, lactation, alcoholic, obese patients, liver or renal toxicity, patients with peptic ulcer disease, men or women currently planning to have a child, potentially serious infection that could be reactivated (such as TB, hepatitis), patients who use other medications (such as topical treatment or biological therapy) in combination with MTX and patients with incomplete data or lost from the follow up.

For every patient involved in the study: A thorough history was taken which included gender, age, residency, occupation, duration of the disease, previous treatment taken with emphasis about the reason of discontinuation. Review of other systems were collected especially the gastrointestinal tract system, genitourinary system and musculoskeletal system.

Past medical history, past surgical history, social history, marital status, family history and drug history, including the previous drug regimens used by the patient for psoriasis (type, dose, duration, any reaction or side effect), drug used for other diseases, any drug allergy or side effects were also recorded.

General examination: looking for pallor, jaundice, body temperature, any signs of bleeding tendency (ecchymosis, petechia), body weight and body mass index. Examination of the lesions: by naked eyes and sometimes using magnifying lens looking for erythema, scaling, and estimating the thickness of the lesions by palpation. Measuring the baseline PASI score, in addition to examination of other areas of predilection for psoriasis (e.g., scalp, nail, tongue) looking for the distribution of the lesions, any kobnerization and measuring body surface area involved and examination of other systems especially related systems such as joints for any signs of inflammation (tenderness, redness, warmth, and swelling). Psoriasis area and severity index score were calculated to all patients.

All patients were sent for the following laboratory tests: Complete blood count (CBC), liver function test (LFT): serum bilirubin, aspartate transaminase (AST) and alanine transaminase (ALT), renal function test (RFT): serum urea and blood creatinine level, virology screen: (hepatitis B, hepatitis C, and HIV) and pregnancy test for married females.

We needed to send some patients for consultations to other medical branches like rheumatology, internal medicine and psychiatry. Such as in cases of psoriatic arthritis, or in order to manage MTX side effects like severe gastrointestinal symptoms (nausea, vomiting, abdominal pain and diarrhea), or psychiatric symptoms (depression and anxiety) and to make sure that continuous use of MTX was safe in some patients who developed abnormal laboratory readings (elevation in liver function test and severe anemia).

All patients received MTX 15 mg in three divided doses over 24-hours period each week, and folic acid tab 5 mg per day in all days except the 24-hours of MTX therapy. They were followed monthly for three months, at each follow up visit we looked for response to treatment by measuring their PASI score, investigations were done LFT and CBC and any side effects from the medication were recorded and managed accordingly.

The dataset divided into patient demographics, disease parameters and drug parameters.

Data processing: In order to create ML algorithm and obtain the result from ML, we had performed the following steps: Changing the data types/formats many

times such as removing the sign (%) from the column of BSA and change the residency from Baghdad for example to number 1 and Babel 2 and so on. Rescoring class variables such as PASI scoring 0 and 1, in which 1 indicates a PASI score equal or greater than 75% and 0 indicates a PASI score less than 75%. Removing any rows with missing data. PASI score and investigation at 1 month and at 2 month columns were removed from dataset.

Machine learning: First of all, we split our data into two groups (D1 and D2) the first group which was about 70% of the data that contain the final outcome for learning the machine and the second group D2 which was 30% of the data in which we've deleted the final outcome to allow the machine to predict it.

The idea from dividing the data into two separate groups because dichotomization ease the analysis. Overfitting occurs when the ML model performs well on the training data but unable to generalize it on the testing phase because the model is memorizing the data it has seen and is unable to generalize to unseen examples, this occur when the same data used for training and testing phase, while under-fitting occur when ML model perform poorly in the training model because the model unable to capture the relationship between the input data (often called X) and the target value (often called Y).[7]

**Table 1: Demographic dataset of the patients.**

| Patient demographics | Mean | Standard deviation |
|---|---|---|
| Age (years) | 36.4 | 13.89 |
| Duration (years) | 9.69 | 7.94 |
| Baseline BSA (%) | 43 | 21.28 |
| Baseline PASI | 20.26 | 8.30 |

**Table 2: Disease parameters dataset of the patients.**

| Disease parameters | Number | Percentage |
|---|---|---|
| Psoriatic arthritis | Yes: 47 | 22.6 |
| Nail psoriasis | Yes: 110 | 52.88 |
| Previous therapy (other than MTX, and number of drugs used) | Yes: 200 (1: 140 2:47 3: 8 4: 5) | 96.15 |
| PASI 75% (1:>=75% PASI, 0:<75% PASI after 3 months) | 1:73<br>0:171 | 35<br>65 |
| Distribution | Scalp: 133<br>Genitalia: 78<br>Upper and lower limbs:18<br>Trunk :4<br>Trunks and limbs: 183<br>Palmoplanter: 56 | 63.94<br>37.5<br>8.6<br>2<br>88<br>26.92 |
| Family history | Yes: 54 | 25.96 |
| Vitiligo | Yes: 5 | 2.4 |
| Number of co-morbidities (diabetes mellitus, hypertension, smoking) | None (0): 121<br>1: 64<br>2:17<br>3: 6 | 58.17<br>30.76<br>8.17<br>2.88 |

**Table 3: Drugs parameters dataset of the patients.**

| Drug parameters | Number | Percentage |
|---|---|---|
| Abnormal investigation | Elevated liver enzyme: 28<br>Abnormal CBC: 9 | 13.45<br>4.32 |
| Patients with side effects | Nausea: 115<br>Vomiting: 49<br>Diarrhea: 7<br>Abdominal pain: 15<br>Oral ulcer: 2<br>Mood change: 4<br>Hair loss: 5 | 55.28<br>23.55<br>3.36<br>7.211<br>0.96<br>1.92<br>2.4 |

## Results

In this study, those patients who got 75% or more PASI score improvement were labeled as 1, while those with less than 75% improvement labeled as 0.[13]

There are three phases in ML model: Training phase, real time prediction and batch prediction.

In the training phase, we used 70% of the data (D1 group) for training, accuracy metric for binary model called area under curve (AUC) is used by Amazon ML tool to measure ML model accuracy. The AUC metric returned the result as real number between 0 and 1.

From our data, the AUC is 0.92 which reflect the accuracy of machine learning, (Figure 1).

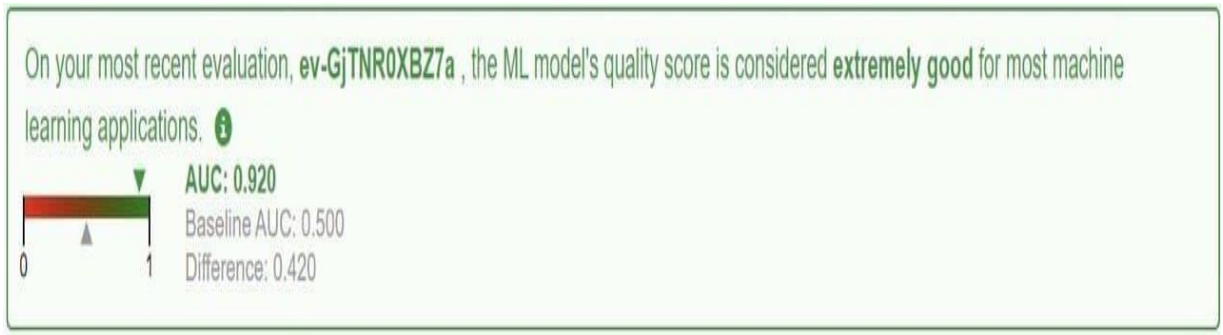


**Figure 1: Machine learning performance metric.**

If the number was near 1, this indicated that the model was trained and accurate. If the number was near 0.5, this indicated that the model was not trained. Numbers near 0 indicated that there was a problem with the data. The baseline AUC metrics for binary model was by default equal to 0.5, the AUC metric should not equal 0.5, in other word; the binary ML model should perform better than the baseline value.[7]

When we used trade-off based on score threshold 0.5,

**80% are correct** **20% are errors**
32.7% true positive 10.2% false positive
46.9% true negative 10.2% false negative
*43% of the records are predicted "1".
*57% of the records are predicted "0".

The accuracy is about 80% with sensitivity equal to 76% and specificity equal to 82%, calculated using the following formulas:

Sensitivity = TP / TP + FN
Specificity = TN / TN + FP
Accuracy = TP+TN / TP+FP+FN+TN

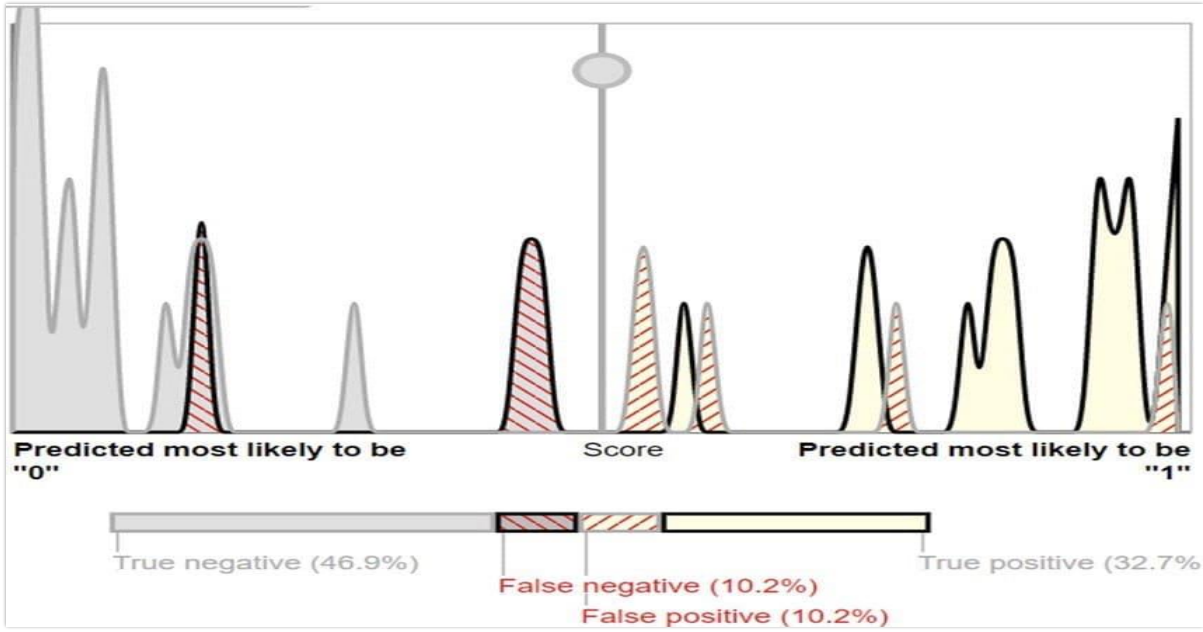


**Figure 2: Details for error score 0.5.**

Then we adjusted the error score to 0.6:

**84% are correct** **16% are errors**
30.6% true positive 4.1% false positive
53.1% true negative 12.2% false negative
*35% of the records are predicted as "1".
*65% of the records are predicted as "0".

The accuracy of 0.6 error score equal to 84% with sensitivity about 71% and specificity equal to 93% as calculated by the following formulas:

Sensitivity = TP / TP + FN
Specificity = TN / TN + FP
Accuracy = TP+TN / TP+FP+FN+TN

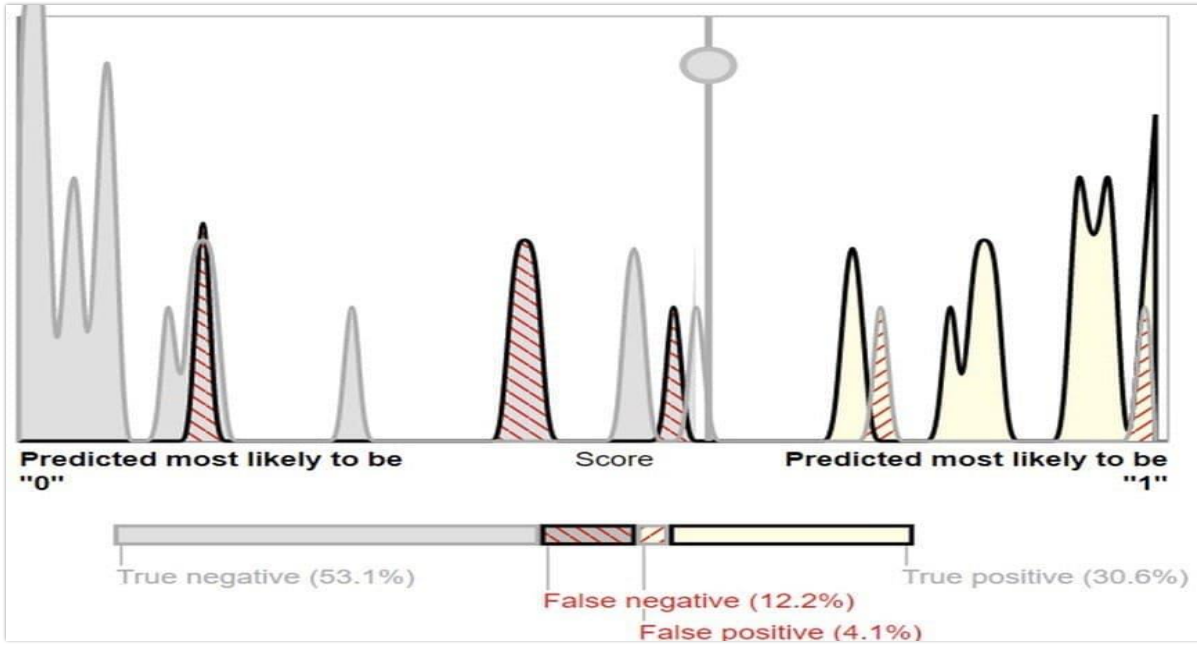


**Figure 3: Details of error 0.6.**

Real time prediction: In this step we enter each case individually from group D2 (30% of our data) which is used for testing the machine as we know the outcome already, below is an example about the first patient in the group D2 where the prediction must be 0, and it gave us a correct answer, i.e., 0, (Figure 4).

Another four examples tested three of them gave correct predication and one false.

Batch prediction: This used when we want to generate predictions for set of observations all at once, and then action on a certain percentage or number of observations. In our model we enter all the D2 group cases at the same time but this step did not work in such small data and needs very big data.

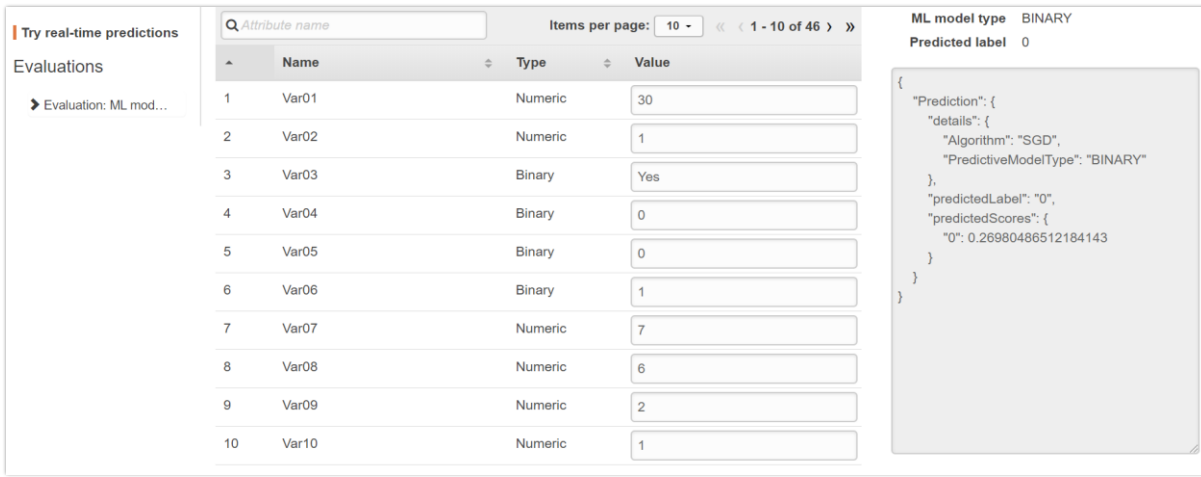


**Figure 4: Machine learning testing cases individually.**

## ——————————Discussion

Despite the vast amounts of data on the efficacy of drugs used in psoriasis management, therapeutic decision-making (decision which treatment to administer to each individual patient) is still based on a trial-and-error approach. There is a wide variety in the response degree and the side effects among the patients so each one need a personalized medicine approach depending on many parameters including patients parameters (age, gender, occupation, residency, social status, duration of disease, body surface area, baseline PASI score), disease parameters (type of psoriasis, lesion distribution, history of previous treatment and the degree of response to it, psoriatic co morbidities, family history of psoriasis and associated other dermatological diseases), drug parameters (type of medication, regime, investigations, and side effects of medication).

As methotrexate is one of the main drugs used for treatment of psoriasis for many years, we concentrated on this drug in this research trying to access the accuracy of ML in predicting who can get benefit from it with minimal side effects and better outcome, to apply this method for other newer drugs like biological therapy.

Machine learning was known to be used in the last few years by other medical branches (internal medicine, oncology, psychiatry, emergency medicine and pharmacy, etc) to predict the outcome of different modalities of treatment or even the morbidity and mortality of different diseases with promising and encouraging results, in

this study we tried to apply this method as a pioneer in our field.

No similar study about methotrexate treatment prediction in psoriatic patient using ML was found in the literature, only one study about biological treatment in psoriasis published in British journal of dermatology in May 2020 by S Emam et al[14]. The number of patients in the present study was 208, 140(67.3%) males and 68(32.7%) females with a mean of 36.4 (±13.89) years, while in the study done by S Emam et al, the number of the patients was larger 681. The difference in the number of the patients between the two studies is related to the method of data collection, in the present study, we depended on the newly registered patients in one center for 16 months duration while in S Emam et al study they depended on data from Danish registry, Dermbio, for longer period from 2003-2013.

In S Emam et al study, males' number was 375 and the females' number was 306 which doesn't go with our results (more male predominance in the present study); with a median age of 42.8 year, which was slightly higher than the present study.

The duration of the disease in the present study ranged from 0.5-30 years with a mean of 9.69 (±7.95) years, while in S Emam et al study it was much longer with mean about 25.84 year; this may be related to the use of the biological treatment which often not used as the first line treatment for psoriasis unless other non-biological topical and systemic treatments failed.

In the present study, the baseline body surface area mean is about 43.69% (± 20.28) and baseline PASI score mean was 20.46 (±8.10), while in S Emam et al study, they were much smaller, the body surface area mean was 13.56% and baseline PASI mean was 10.5, probably because most of the patients used methotrexate or acitretin or other systemic non-biological medications prior to the treatment with biological therapy and this might contribute to lower the severity of the disease and low BSA and PASI score at the baseline in the of S Emam et al study.

The number of the patients with psoriatic arthritis in this study was about 47 (22.6%), while in S Emam et al study it was larger, 454 (66.6%), this could be explained by the longer duration of the disease in the patients in the present study.

In the present study, about 58.1% of patients had no co-morbidities and about 41.9% with one or more comorbidity. This almost exactly goes with S Emam et al study, in which, 58.7% of the patients had no comorbidity and 41.3% with one or more comorbidity.

Before the prescription of methotrexate almost all of the patients had history of taking one or more drugs like topical steroid (clobetasol ointment or betamethason ointment), phototherapy, systemic steroids and herbal medication but showed no benefit or immediate relapse after stopping the treatment.

We used lesion distribution; nail, scalp, palmoplanter, limbs only, trunk only or both and genitalia as variables that might affect the results while this was not taken in consideration in S Emam et al study.

Thirty-seven patients developed abnormal laboratory investigations, 28 had mild elevation of liver enzymes (those patients who had three folds or more elevation from baseline were excluded from the study). Nine patients developed abnormal CBC: four patients had decreased Hb and three had deceased WBC count and two showed decreases in both, all of them the decrease less than 20% from the baseline reading. All those patients sent for internal medicine consultation and no one of them advised to discontinue MTX therapy.

Other side effects: the most common was nausea in 115 (55.28%) patients, vomiting 49 (23.55%) and managed by giving metoclopramide or ondansetron. Seven patients developed mild diarrhea and 15 abdominal pain both treated symptomatically. Two patients developed minor oral ulcers; both of them had irregular

folic acid therapy. They advised to take it regularly and followed up at closer interval of two weeks, the ulcers disappeared without discontinuation of MTX. Four patients developed mood change, psychiatric consultations were done for them and treated accordingly.

Data processing: In order to make the ML able to understand our data and give the results we had to expand extra-time cleaning the data and changing its format and deleting some unnecessary variables such as the in between follow up information (PASI or investigations between the baseline and the final outcome). A similar way of data management done by S Emam et al.

Machine learning: There are different models used by ML programs for precision medicine like general linear model, deep learning, support vector machine, decision tree, gradient-boosted trees and others. The binary model which is applied in the present study predicted the response of psoriatic patients to methotrexate treatment with an accuracy exceeding 80%, even after we change the error threshold score from 0.5 to 0.6, which was surprising, as the dataset is relativity small which means that the machine is learning from the data and can be used to predict which patient will respond to MTX treatment with sensitivity about 76% and specificity about 82%. This can be expanded by using bigger data with more additional variables. Although we used only one model (binary model) which is general linear model, S Emam et al used seven different models but the accuracy was almost the same (close to 80 for each model).

In conclusion; Machine learning algorithm can predict the response of methotrexate treatment in plaque psoriatic patients with an accuracy exceeding 80%, based on small set of variables, and this can save time and effort in predicting which patients will respond to MTX therapy with minimal side effect and better outcome.

Author contributions

**Sama Asim Sahib:** Study concept, study design, data collection, data analysis, interpretation of data and drafting of manuscript.

**Azzam Muhsin Al-Salami**: Data collection, data interpretation and support.

**Nada Fadhil Al-Tajer:** Data collection and data interpretation.

**Sabeeh A Al-Mashhadani:** Supervisor the findings of this work.

## References

1. Menter A, Gelfand JM, Connor C, Armstrong AW, Cordoro KM, Davis DMR, et al. Joint American Academy of Dermatology–National Psoriasis Foundation guidelines of care for the management of psoriasis with systemic nonbiologic therapies. J Am Acad Dermatol 2020; 82(6):1445-86.
2. Elmets CA, Lim HW, Stoff B, Connor C, Cordoro KM, Lebwohl M, et al. Joint American Academy of Dermatology–National Psoriasis Foundation guidelines of care for the management and treatment of psoriasis with phototherapy. J Am Acad Dermatol 2019; 81(3):775-804.
3. Whittle SL, Hughes RA. Folate supplementation and methotrexate treatment in rheumatoid arthritis: A review. Rheumatol Oxf Engl 2004; 43(3):267-71.
4. Montesinos MC, Desai A, Delano D, Chen J-F, Fink JS, Jacobson MA, et al. Adenosine A2A or A3 receptors are required for inhibition of inflammation by methotrexate and its analog MX-68. Arthritis Rheum 2003;48(1):240-7.
5. Jeffes EW, McCullough JL, Pittelkow MR, McCormick A, Almanzor J, Liu G, et al. Methotrexate therapy of psoriasis: Differential sensitivity of proliferating lymphoid and epithelial cells to the cytotoxic and growth-inhibitory effects of methotrexate. J Invest Dermatol 1995; 104(2):183-8.
6. Turing AM. Computing machinery and intelligence. Mind. Dordrecht: Springer; Mind 1950; 49: 433-60.
7. Chan S, Reddy V, Myers B, Thibodeaux Q, Brownstone N, Liao W. Machine Learning in dermatology: Current applications, opportunities, and limitations. Dermatol Ther 2020;10(3):365-86.
8. Simon A, Deo M, Selvam V, Babu R. An overview of machine learning and its applications. Int J Electr Sci Eng 2016; 22-4.
9. Jameson JL, Longo DL. Precision medicine-personalized, problematic, and promising. N Engl J Med 2015; 372(23):2229-34.
10. Correa da Rosa J, Kim J, Tian S, Tomalin LE, Krueger JG, Suárez-Fariñas M. Shrinking the psoriasis assessment gap: Early gene-expression profiling accurately predicts response to long-term treatment. J Invest Dermatol 2017; 137(2):305-12.
11. Tomalin LE, Kim J, Correa da Rosa J et al. Early quantification of systemic inflammatory proteins predicts long-term treatment response to

Tofacitinib and Etanercept. Psoriasis Response Prediction 2019; 10(5): 503-18.
12. Emam S, Du AX, Surmanowicz P, Thomsen SF, Greiner R, Gniadecki R. Predicting the long term outcomes of biologics in patients with psoriasis using machine learning. Br J Dermatol 2020; 182(5):1305-7.
13. Gordon KB, Betts KA, Sundaram M, Signorovich JE, Li J, Xie M, Wu EQ, Okum MM. Poor early response to methotrexate portends inadequate long-term outcome in patients with moderate-to-severe psoriasis: Evidence from 2 phase 3 clinical trial. J Am Acad Dermatol. 2017; 77(6):1030-7.
14. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov IV. Artificial intelligence applications in dermatology: Where do we stand? Front Med 2020; 7:100.

**IMJ 2022; 68(2): 73-82.**