PMAC
PRINCE MAHIDOL AWARD CONFERENCE 2025

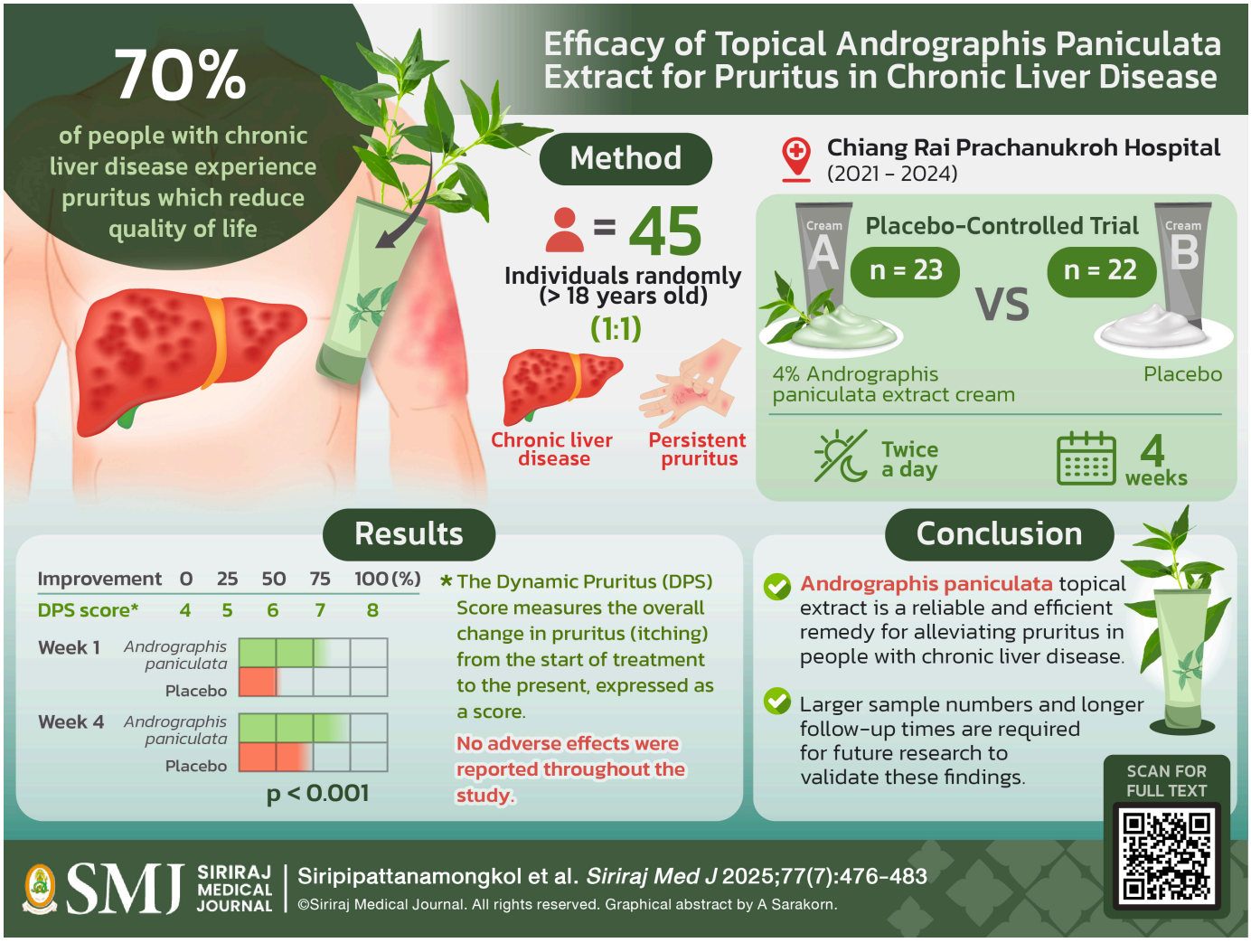
E-ISSN 2228-8082
Volume 77, Number 7, July 2025

# SMJ
## Siriraj Medical Journal

The world-leading biomedical science of Thailand

MONTHLY

ORIGINAL ARTICLE REVIEW ARTICLE




Indexed by Scopus® DOAJ

https://he02.tci-thaijo.org/index.php/sirirajmedj/index
E-mail: sijournal92@gmail.com

SMJ
Siriraj Medical Journal
The world-leading biomedical science of Thailand
Volume 77 Number 7
July 2025

## ORIGINAL ARTICLE



## REVIEW ARTICLE

# SIRIRAJ MEDICAL JOURNAL

https://he02.tci-thaijo.org/index.php/sirirajmedj/index



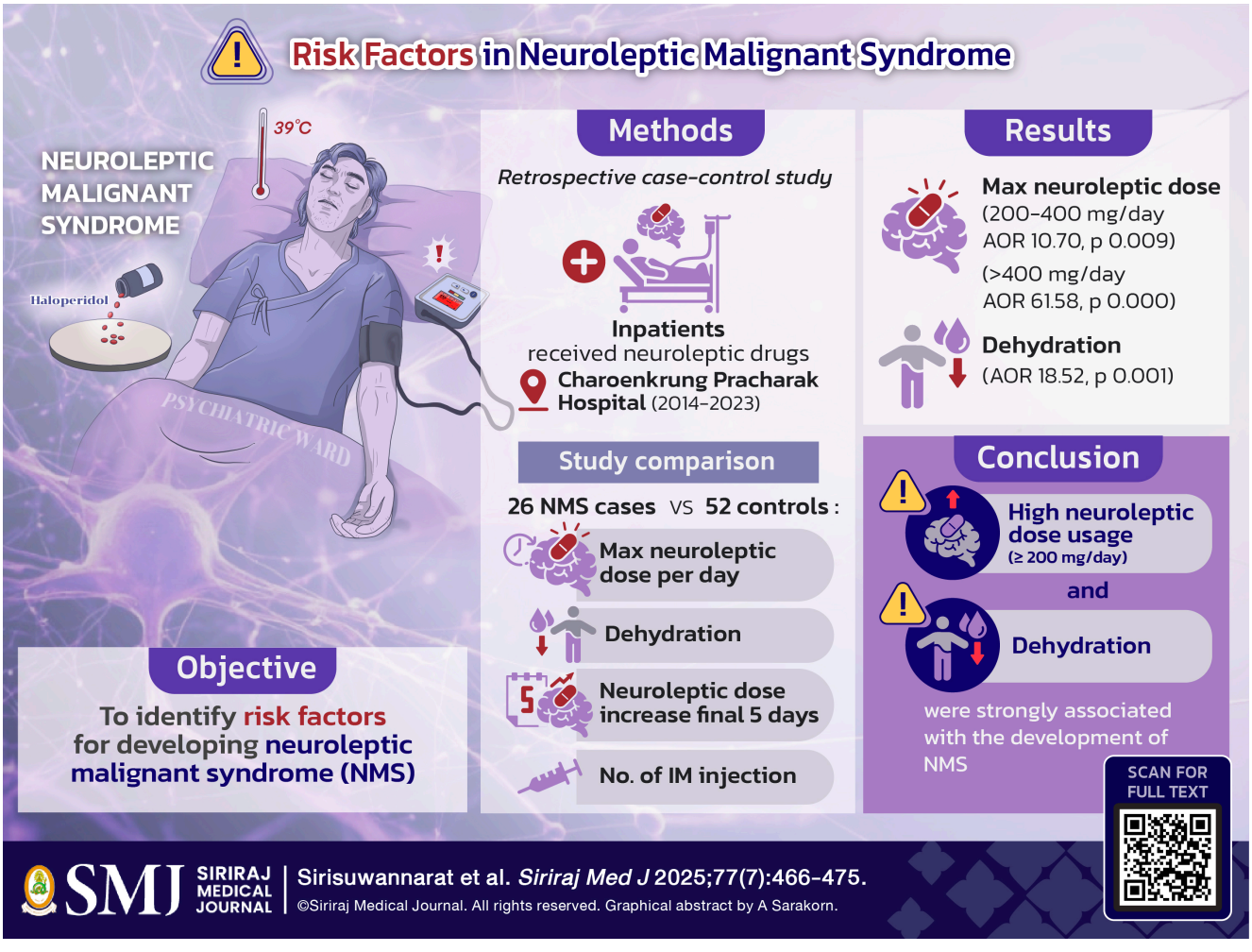


Office: His Majesty the King's 80th Birthday Anniversary 5th December 2007 Building (SIMR), 2nd Fl., Room No.207
Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand
E-mail: sijournal92@gmail.com

# Risk Factors in Neuroleptic Malignant Syndrome: A 10-year Case-control Study of Neuroleptic Use in Patients in Charoenkrung Pracharak Hospital

**Sarinthip Sirisuwannarat, M.D.**

*Unit of Psychiatry, Charoenkrung Pracharak Hospital, Bangkok, Thailand.*




*Corresponding author: Sarinthip Sirisuwannarat*
*E-mail: sarinthip.gift@gmail.com*
*Received 19 February 2025 Revised 14 March 2025 Accepted 16 March 2025*
*ORCID ID:http://orcid.org/0009-0000-0340-7955*
*https://doi.org/10.33192/smj.v77i7.273739*

## ABSTRACT

**Objective:** To identify the risk of using high doses of neuroleptics in the development of neuroleptic malignant syndrome (NMS). Additionally, we examined other potential risk factors, including age, psychiatric diagnosis, route of neuroleptic administration, dose escalation over a short period, psychomotor agitation, dehydration, and electrolyte imbalance.

**Materials and Methods:** A case-control study was performed, comprising 26 NMS cases matched with 52 controls by sex and time of admission to the hospital over a 10-year data-collection period. A retrospective chart review was conducted to compare the two groups using conditional logistic regression analysis.

**Results:** The maximum neuroleptic dose (adjusted OR 10.70, 95%CI 1.79–64.00 for a neuroleptic dose of 200–400 mg/day and adjusted OR 61.58, 95%CI 6.87–552.19 for a neuroleptic dose > 400 mg/day) and dehydration (adjusted OR 18.52, 95%CI 3.22–106.62) were found to be significant risk factors for developing NMS.

**Conclusion:** The risk factors for developing NMS were found to be receiving a high dose of neuroleptics per day and dehydration.


**Keywords:** Neuroleptic malignant syndrome; maximum neuroleptic dose; neuroleptic increase final 5 days; delirium; dehydration; electrolyte imbalance (Siriraj Med J 2025; 77: 466-475)

## INTRODUCTION

Neuroleptic malignant syndrome (NMS) is a rare fatal adverse effect of taking neuroleptic drugs. This syndrome is characterized by three main symptoms: generalized muscle rigidity with akinesia, autonomic nervous system (ANS) dysregulation, and alteration in mental status.[1] The incidence rate ranged from 0.02% to 3.2%.[2] The mortality rate was reported as high as 10%–55%.[3] Currently, the most effective approach to avoid NMS is prevention through raising awareness of the potential risks. Identifying the risk factors has been of great interest.

Previous research into the risk factors for developing NMS has been done by the first case series of Addonizio et al. U.S. in 1986.[4] Later, the first case-control study was published by Keck et al. U.S. in 1989[5], who concluded that the risk factors correlated with NMS were the use of a high dose of neuroleptic drugs within a short period of time and intramuscular neuroleptic injection; however, that study had some limitations, including its small sample size, which limited the analysis of the statistically significant independent variables.

Since then, there have been many studies performed in various formats, including case series, cross-sectional, and case-control methodologies, aimed at determining the potential risk factors associated with NMS. One good quality study was done by Sachdev et al. in 1997[6] with an Australian population and an adequate sample size (25 cases and 50 controls) that showed that receiving a high neuroleptic dose per day was a risk factor for developing NMS. This study also found that dehydration was another key risk factor like in a study of Chen et al.[7] which suggested that dehydration was the risk factor for NMS.

Many experts have studied about risk factors for developing NMS. "Maximum neuroleptic dose" (Maximum dose of neuroleptic per day) was one of the most interesting factors. The previous studies[5,6,8,9] found that maximum neuroleptic dose was the associated factor to cause NMS but all these studies emphasize only increasing dose was the risk but did not study which dose level that being the risk. The study by Su YP et al.[10] and Guinart et al.[11] were the large and most recent studies that studied more about which dose level that might be the risk but their finding had different outcomes. SU YP et al. concluded that maximum neuroleptic dose was not associated with increasing the risk of NMS but Guinart et al. concluded that neuroleptic dose more than 2 DDDs (defined daily dose) increase the risk of developing NMS.

About the class and potency of neuroleptic drugs, we found three large case-control studies which had results in the same direction that likelihood of developing NMS did not differ by neuroleptic class (first-generation antipsychotics VS second-generation antipsychotics).[7,11,12] About the potency of neuroleptic drug, Nielsen et al.[13] found that high- or mid-potency first-generation antipsychotics and second-generation antipsychotics indicated an increased risk of NMS. However, since it did not directly compare the two potencies, we cannot conclusively determine whether potency itself is a risk factor for developing NMS.

Regarding the form of neuroleptic use, intramuscular neuroleptic injection was frequently reported as a significant risk factor for developing NMS in most studies.[5,9] Only

one study[6] reported non-significant results. Concerning long-acting neuroleptic injection, the majority of studies found no statistically significant association with NMS.[9-12,14]

In terms of "increasing the neuroleptic dose over a short period", many studies have emphasized this factor.[5,6,8,15] However, the outcomes have been controversial. Keck et al.[5] and Berardi et al.[8] reported positive findings supporting this as a risk factor, whereas Sachdev et al.[6] and Langan et al.[15] found no significant association. Langan et al. specifically concluded that "increasing the neuroleptic dose over a short period" is not a risk factor for developing NMS.

In Thailand, the focus area of this research, there have been prior case reports and case series about NMS, Taemeeyapradit et al.,1989[16] and Kooptiwoot et al.,1999[17] were the first and second case reports in Thailand but, to the best of our knowledge, no case-control study has yet been performed in Thai population. However, four big case series studies were performed. The first case series was by Wae-alee et al. in 1996[18], who found that changes in the amount or type of neuroleptic drug used were associated factors for developing NMS. The second case series was performed by Tantiphlachiva in 1999[19], who found that the associated factors for developing NMS were an agitated state, dehydration, long-acting neuroleptic injection, and intramuscular neuroleptic injection. The third case series was performed by Kasantikul et al. in 2006[20] who found that the associated factor for developing NMS was dehydration. The fourth, and most recent, case series was performed by us in 2020.[21] In our study, we postulated that old age, delirium, and alcohol dependency were associated risk factors for developing NMS.

In conclusion, from reviewing the past literature, the following factors have been put forward as risk factors for developing NMS: male sex[7], old age[21], psychiatric diagnosis[22-26], delirium[7,8], alcohol dependence[21], intravenous neuroleptic injection[8], intramuscular neuroleptic injection[5,9], long-acting neuroleptic injection[19], maximum neuroleptic dose[5,6,8,9,11,15], increase in neuroleptic dose in a short period[5,8], changes in the amount or type of neuroleptic drug[18], psychomotor agitation[5,6,8,9,19], dehydration[6,7,19,20], mechanical restraint or locked in open seclusion[6], extra pyramidal symptoms[7,8], and electrolyte imbalance.[27]

Nevertheless, despite significant advances in the literature regarding risk factors for NMS, there remain some limitations. For example, smaller studies often provide detailed information on each case and control but they lack sufficient statistical power to reach definitive conclusions. Conversely, larger studies may have adequate statistical power but frequently lack detailed data essential for ensuring reliability. Additionally, the two largest and most recent studies examining the maximum neuroleptic dose reported conflicting outcomes, resulting in uncertainty about whether the maximum neuroleptic dose is a definitive risk factor for NMS. Furthermore, most previous case-control studies were conducted in Western populations. Upon reviewing the literature, we found that non-white ethnic groups have a twofold increased risk of NMS[10], potentially related to genetic differences in the CYP2D6 gene, as Caucasian populations typically possess more normal-function CYP2D6 alleles compared to Asian, African, or African American populations.[28,29] Due to these varying outcomes and genetic differences between Caucasian and non-white populations (including Asians), it was essential to perform a case-control study that has adequate sample size for the main objective and detailed data collection specifically targeting the Thai population to enhance the reliability of our findings.

In our opinion, we consider that the "maximum neuroleptic dose at a high level" may be the greatest risk factor for developing NMS based on our review of the literature, whereby this factor has the strongest evidence to support this as a risk factor. Also, previous studies mostly did not identify the exact dose that would represent a risk factor for NMS. Consequently, these inspired us to set the primary outcome as "the level of neuroleptic dose that would represent a risk factor for developing NMS". Our secondary outcomes were the other factors that might be risk factors for NMS, such as age, psychiatric diagnosis, intramuscular neuroleptic injection, long-acting neuroleptic injection, neuroleptic increase final 5 days (increase in neuroleptic dose in a short period), psychomotor agitation, dehydration, and electrolyte imbalance.

## MATERIALS AND METHODS

The study protocol was approved by the Bangkok Metropolitan Administration Human Research Ethics Committee (BMAHREC); Approval number: S015hc/67_EXP.

We initially identified all the neuroleptic-treated inpatients in our hospital from 2014 to 2023 who were diagnosed with NMS according to ICD-10 code G210, then selected only the cases that matched the DSM-5 criteria.[30] After all the cases (patients with neuroleptic use who were diagnosed with NMS by DSM-5 criteria) had been identified, systematic sampling (as shown in Fig 1) was performed to select 26 NMS cases (as the calculated required sample size, see below) for the statistical analysis. We chose 52 control patients (matched to the cases by

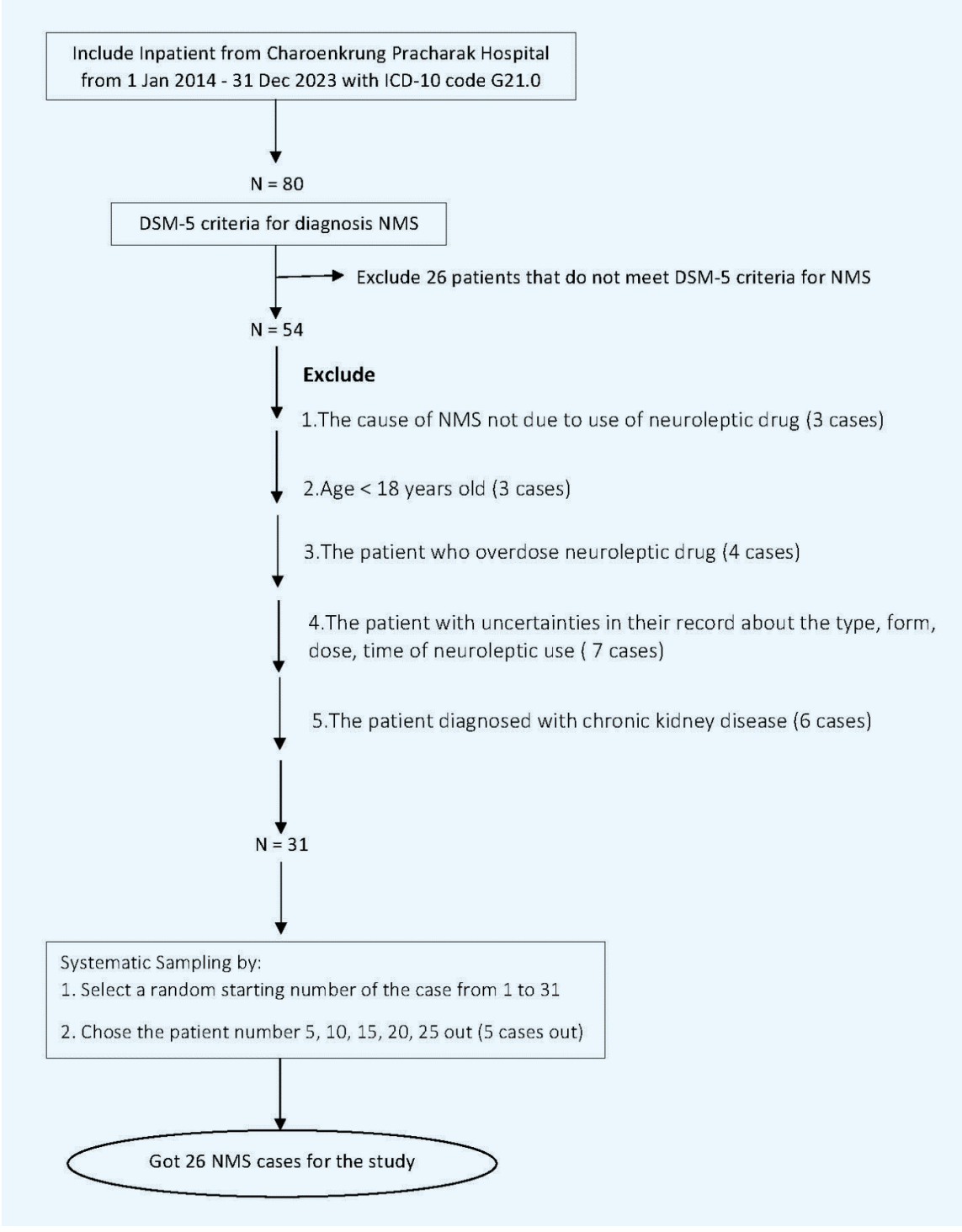



**Fig 1.** Systematic sampling flowchart for including NMS cases to this study.

gender and date of admission to the hospital). All the patients were treated at the inpatient unit of Charoenkrung Pracharak Hospital. We excluded patients who were aged <18 years old, who did not use neuroleptic drugs, intentionally took an overdose of neuroleptic drugs, patients with missing laboratory data or a psychiatric diagnosis, and patients with uncertainties in their records about the type, form, dose, and time of neuroleptic use.

The definitions of some variables are as follows:

1. Maximum neuroleptic dose (mg per day): the dose of neuroleptics used on the day with the highest dose within 2 weeks preceding the end point[31] (see definition of "the end point" below). The total neuroleptic dose was converted to chlorpromazine equivalents dose (using Davis's schedules[32]).

- For intravenous neuroleptic use: doses were

converted to chlorpromazine equivalents by estimating the intravenous administration to be 3 times as potent as oral dosing, as determined by Davis's schedules.[32]

- For intramuscular neuroleptic use: doses of haloperidol 5 mg intramuscular were converted to haloperidol 8.3 mg oral (reference from Hertfordshire Partnership University NHS Foundation Trust).

- For long-acting neuroleptic use:

- haloperidol decanoate 100 mg intramuscular was considered equivalent to 10 mg oral haloperidol[33];
- fluphenazine decanoate 25 mg intramuscular was considered equivalent to 300 mg oral chlorpromazine.[9]

2. Neuroleptic increase final 5 days: neuroleptic dose increases during the 5 days preceding the end point.

- For a dose increase compared to the starting point: use '+' symbol (e.g., +50 mg);

- For a dose decrease compared to the starting point: use '-' symbol (e.g., -50 mg).

3. Number of intramuscular neuroleptic injections: number of intramuscular neuroleptic injections received during the 2 weeks preceding the end point.[31]

4. Long-acting neuroleptic injection: "use" or "no use" of long-acting neuroleptics within 1 month preceding the end point.[9]

5. Psychiatric diagnosis: defined following the ICD-10 codes.

6. Delirium: defined as "delirium" following the ICD-10 code "F05" within 2 weeks preceding the end point.[31]

7. Psychomotor agitation: excessive and purposeless motor activity, requiring medication or restraint or seclusion within 2 weeks preceding the end point.[31]

8. Dehydration: 1) hematocrit or blood urea nitrogen concentration (BUN) > 50% of the normal range or 2) serum creatinine > 1.2 mg/dL[8] preceding the end point (rules out patients diagnosed with chronic kidney disease who might show a permanent BUN or serum creatinine increase to prevent misinterpreting this as dehydration).

9. Electrolyte imbalance: hyper-hyponatremia with blood sodium level > 145 or < 136 mmol/L, and hyper-hypokalemia with blood potassium level > 5.1 or < 3.5 mmol/L preceding the end point.

The definition of "the end point" (for variable numbers 1–9 above)

- for cases: the onset of NMS[8];

- for controls: the hospital day corresponding to the day of onset of NMS in the cases.[8]

The date used for recording the laboratory data for variable numbers 8–9 was recorded within ±3 days, when available.[8]

**Sample size**

We calculated the required sample size for our study based on the study by Sachdev et al.[6] as a reference (see Table 1 in Sachdev et al.: "Case-control study of neuroleptic malignant syndrome"). The "maximum neuroleptic dose" was used to calculate the required number of cases and controls as follows.

$$n_1 = \frac{(z_{1-\frac{\alpha}{2}} + z_{1-\beta})^2 \left[\sigma_1^2 + \frac{\sigma_2^2}{r}\right]}{\Delta^2}$$

$$r = \frac{n_2}{n_1},\ \Delta = \mu_1 - \mu_2$$

$$\mu_1 = mean\ in\ group1$$
$$\mu_2 = mean\ in\ group2$$
$$\sigma_1 = S.D.\ in\ group1$$
$$\sigma_2 = S.D.\ in\ group2$$
$$\mu_1 = 958.3$$
$$\mu_2 = 488.1$$
$$\sigma_1 = 821.1$$
$$\sigma_2 = 334.4$$

n1 = cases n2 = controls

$$\gamma = ratio = \frac{2}{1}$$
$$\alpha = 0.05\ ;\ \beta = 0.2$$
$$n_1 = \frac{[Z_{1-\frac{0.05}{2}} + Z_{1-0.2}]^2[(821.1)^2 + \frac{334.4^2}{2}]}{(958.3 - 488.1)^2}$$

$$n_1 = 26$$
$$n_2 = 52$$

The sample size for cases was 26, controls was 52 based on the calculated formula as above.

**Statistical analysis**

Comparative analyses of the two groups were conducted using different statistical methods based on the type of data. For continuous variables with a normal distribution, the T-test was employed, while the Wilcoxon signed-rank test was utilized for data that were non-normally distributed or for noncontinuous discrete variables. Fisher's exact test was applied to assess the differences in dichotomous variables. The data are presented herein as frequencies and as mean values accompanied by their standard deviations (mean ± SD). Relationships between variables were analyzed through conditional logistic regression, which is known to be suitable for case-control studies where one case is matched to two control subjects (A 1:2 case-control ratio was chosen to improve statistical power while

**TABLE 1.** Differences in risk factors for developing neuroleptic malignant syndrome between patients with the syndrome (N=26) and the controls (N=52): continuous variables.

| Variables | Case | | Control | | Crude Odds Ratio | | | | Adjusted Odds Ratio | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Continuous | Mean | SD | Mean | SD | Odds ratio | P-value | 95% Confidence interval | | Odd Ratio | P-value | 95% Confidence interval | |
| Age | 57.23 | 18.02 | 63.96 | 17.97 | 0.98 | 0.098 | 0.95 | 1.00 | | | | |
| Maximum neuroleptic dose (mg/day) | 534.32 | 314.90 | 114.31 | 314.91 | 1.01 | 0 | 1.00 | 1.01 | | | | |
| - Neuroleptic level 0–200 mg/day | 106.80 | 54.80 | 67.63 | 44.31 | NS | NS | NS | NS | | | | |
| - Neuroleptic level 200–400 mg/day | 254.92 | 127.49 | 92.43 | 70.89 | 8.20* | 0.002 | 2.13 | 31.63 | 10.70* | 0.009 | 1.79 | 64.00 |
| - Neuroleptic level >400 mg/day | 795.24 | 363.92 | 471.67 | 49.07 | 35.53* | 0 | 7.46 | 169.33 | 61.58* | 0.000 | 6.87 | 552.19 |
| No. of intramuscular injections | 0.62 | 0.77 | 0.15 | 0.78 | 2.47* | 0.037 | 1.06 | 5.79 | 0.85 | 0.712 | 0.37 | 1.97 |
| Neuroleptic dose increase final 5 days | 46.80 | 147.92 | -3.80 | 147.92 | 1.00 | 0.170 | 1.00 | 1.01 | | | | |
| - Neuroleptic increase < 0 mg | 33.04 | 147.18 | -70.45 | 55.00 | NS | NS | NS | NS | | | | |
| - Neuroleptic increase ≥ 0 mg | 95.51 | 186.97 | 28.57 | 83.00 | 3.72* | 0.050 | 0.98 | 14.16 | 9.65 | 0.070 | 0.83 | 111.84 |

* = statistically significant

**Abbreviations:** N = Number of patients; NS = not statistically significant; OR = Odds Ratio; SD = Standard deviation

maintaining feasibility in data collection.) The results are reported as odds ratio with 95% confidence intervals (95% CI) to indicate statistical significance. The analyses were performed using STATA version 17.0.

## RESULTS

In total, 26 inpatients from Charoenkrung Pracharak Hospital who were diagnosed with NMS during January 1, 2014, to December 31, 2023 were added in our study, in line with the required sample size determined above. The data of these patients were compared to those of the control group, comprising 52 inpatients. The mean age for the cases was 57.23 ± 18.02 years old, and for the controls was 63.96 ± 17.97 years old, with the proportions of males and females as 53.85% males and 46.15% females in both groups. The main psychiatric diagnosis for both the cases and controls was delirium (cases = 57.69%, controls = 65.38%). The second main psychiatric diagnosis was schizophrenia for both the cases and controls (cases = 38.46%, controls = 25.00%). There was no statistically significant difference found for all types of psychiatric diagnosis.

Table 1 and Table 2 present comparisons of the cases and controls for different variables. Four variables showed statistical significance: 1) maximum neuroleptic dose (categorized into 200–400 mg/day: crude OR = 8.2,

**TABLE 2.** Differences in risk factors for developing neuroleptic malignant syndrome between patients with the syndrome (N=26) and the controls (N=52): categorical variables.

| Variables Categorical | | Case N | Case % | Control N | Control % | Crude Odds Ratio | | | | Adjusted Odds Ratio | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Odds ratio | P-value | 95% Confidence interval | | Odds ratio | P-value | 95% Confidence interval | |
| Sex | Male | 14 | 53.85 | 28 | 53.85 | Matched variable | | | | | | | |
| | Female | 12 | 46.15 | 24 | 46.15 | | | | | | | | |
| Psychiatric diagnosis | Schizophrenia | 10 | 38.46 | 13 | 25.00 | 2.44 | 0.083 | 0.89 | 6.72 | | | | |
| | Bipolar disorder | 2 | 7.69 | 1 | 1.92 | 4.25 | 0.247 | 0.37 | 49.20 | | | | |
| | Major depressive disorder | 1 | 3.85 | 1 | 1.92 | 2.04 | 0.619 | 0.12 | 33.98 | | | | |
| | Organic mental syndrome | 3 | 11.54 | 9 | 17.31 | 0.62 | 0.508 | 0.15 | 2.53 | | | | |
| | Anxiety | 1 | 3.85 | 1 | 1.92 | 2.04 | 0.619 | 0.12 | 33.98 | | | | |
| | Substance use disorder | 0 | 0.00 | 3 | 5.77 | NS | NS | NS | NS | | | | |
| | Delirium | 15 | 57.69 | 34 | 65.38 | 0.89 | 0.843 | 0.27 | 2.89 | | | | |
| | Mental retardation | 3 | 11.54 | 0 | 0 | NS | NS | NS | NS | | | | |
| | Alcohol dependence | 6 | 23.08 | 11 | 21.15 | 6.65 | 0.109 | 0.66 | 67.43 | | | | |
| Pre-NMS conditions | Agitation | 13 | 50.00 | 31 | 59.62 | 0.79 | 0.627 | 0.31 | 2.04 | | | | |
| | Dehydration | 14 | 53.85 | 6 | 11.54 | 8.94* | 0.000 | 2.84 | 28.20 | 18.52* | 0.001 | 3.22 | 106.62 |
| | Electrolyte imbalance | 18 | 69.23 | 31 | 59.62 | 1.41 | 0.505 | 0.16 | 3.83 | | | | |
| Others related to treatment | Long-acting neuroleptic injections | 3 | 11.54 | 0 | 0 | 0.67 | 0.414 | 0.26 | 1.75 | | | | |

* = statistically significant

**Abbreviations:** N = Number of patients; NS = Not statistically significant; OR = Odds ratio; S.D. = Standard deviation

95% CI = 2.13–31.63 and >400 mg/day: crude OR = 35.53, 95% CI = 7.46–169.33); 2) number of intramuscular injections (crude OR = 2.47, 95% CI = 1.06–5.80); 3) neuroleptic increase final 5 days (≥0 mg: crude OR = 3.72, 95% CI = 0.98–14.16); 4) dehydration (crude OR = 8.94, 95% CI = 2.84–28.20). Next, we performed conditional logistic regression analysis for these four significant variables. The results (Tables 1 and Table 2) revealed three significant variables (P<0.05): 1) maximum neuroleptic dose 200–400 mg/day (adjusted OR = 10.70, 95% CI = 1.79–64.00); 2) maximum neuroleptic dose exceeding 400 mg/day (adjusted OR = 61.58, 95% CI = 6.87–552.19); and 3) dehydration (adjusted OR = 18.52, 95% CI = 3.22–106.62). For our primary outcome, we found evidence that the greater the increase in neuroleptic dose per day, the stronger the statistical significance, implying that the factor "maximum neuroleptic dose" was a great risk for developing NMS.

For the secondary outcomes, only dehydration was statistically significant between the two groups (adjusted OR = 18.52, 95% CI = 3.22–106.62), while the factors: age, psychiatric diagnosis, delirium, alcohol dependence, neuroleptic increase final 5 days, intramuscular neuroleptic injection, long-acting neuroleptic injection, psychomotor agitation, and electrolyte imbalance did not differ significantly between the two groups.

## DISCUSSION

From our study, the most significant factors correlated with the risk of developing NMS were the **maximum neuroleptic dose** and **dehydration**. Our study suggests that NMS appears to occur in patients who have received high doses of neuroleptics or are dehydrated. These findings accorded with the trends reported in previous studies[5,6,8,9,11] who reported that receiving a high neuroleptic dose per day was a risk factor for NMS. We also found evidence that the higher the neuroleptic dose, the stronger the statistical results. Our findings implied that the factor "maximum neuroleptic dose" was the greatest risk factor for developing NMS [maximum neuroleptic dose of 200–400 mg/day had an adjusted OR of 10.7 (P = 0.009, 95% CI = 1.79–64.00) while a maximum neuroleptic dose exceeding 400 mg/day had an adjusted OR of 61.58 (P = 0.000, 95% CI = 6.87–552.19)]. Our study results were in line with a previous case series of 13 NMS cases reported[15], who found that a "rapid dose escalation of neuroleptic" might be less correlated to the development of NMS than a "higher daily cumulative neuroleptic dose."

Regarding the number of intramuscular neuroleptic injections, our study found a statistically significant difference between the two groups when we performed univariable logistic regression (crude OR = 2.47, 95% CI = 1.06–5.79) but was not statistically significant when we performed multivariable logistic regression (adjusted OR = 0.85, 95% CI = 0.37–1.97), which was in contrast to previous studies[5,9], who found that intramuscular neuroleptic injection tended to be a risk factor for NMS.

Among the variables, the number of intramuscular neuroleptic injections and an increase in neuroleptic dose in the final 5 days showed trends toward significance, but did not reach this when we considered the adjusted odds ratio. In particular, a quite strong adjusted odds ratio of 9.65 was found for the neuroleptic increase final 5 days, but its p-value was 0.07, exceeding the significance threshold of $p < 0.05$. Our outcome was probably due to the high correlation between the three variables; maximum neuroleptic dose, rate of dose increase, and route of neuroleptic usage in this small sample, which made it difficult to see the impact of each variable separately. Increasing the sample size, or studying only one variable and matching other variables that have a high correlation with it could improve the statistical reliability. For example, if we aimed to study intramuscular neuroleptic injection as a possible risk factor, we could match the maximum neuroleptic dose and the rate of dose increase.

Our data suggest that dehydration independently contributes to the risk of NMS, aligning with the findings in previous studies[6,7,19,20] which strongly suggested that dehydration is a significant risk factor for developing NMS.

For the factor "age" our study showed a different outcome from in the previous study[21] this might because our study mostly consists of elderly patients in both the cases and controls, so we could not identify a difference in this factor.

No statistically significant difference was found among all the types of psychiatric diagnosis (Table 2), which was a different outcome from the reviewed literature studies in refs[22-26] stating that psychiatric diagnosis was a risk factor for developing NMS. Then, we concluded that the psychiatric diagnosis is not being the risk factor for NMS. Surprisingly, we found no significant difference for the factor "delirium" and "alcohol dependence", unlike in the reviewed literature studies[7,8] that delirium was the risk for NMS and unlike in the reviewed literature study[21] that alcohol dependence was a risk factor for NMS. The reason that in our study these factors showed no statistical significance, which may have been due to different in characteristics of the sample group from normal population that in our study both the cases and controls were mostly under delirious and alcohol dependent state, so we could not identify a difference between the two groups.

In terms of "long-acting neuroleptic injection" as a factor, we found no statistical significance among the two groups, unlike in the case series.[19] But our study tends to has the same outcome as the study of SU YP et al.[10] which found that depot flupentixol has no significant association to be the cause of NMS, also in the same direction as previous studies[9,11,12,14] which found that long-acting neuroleptic injection was not associated with an increased frequency of NMS.

For the factor of psychomotor agitation, there was no statistical significance found in our study, unlike the results in previous studies[5,6,8,9,19], who all reported that psychomotor agitation was a risk factor for NMS. The difference might be because in our study both the cases and the controls were mostly in an agitated state and our definition of psychomotor agitation was less specific to detect a differentiation between the groups. Future studies should investigate this factor further, and should consider a more specific definition in rating the severity of agitation to rule out cases where agitation was present but not severe enough to cause NMS.

Regarding electrolyte imbalance, this could be due to the fact that all our controls were inpatients, who are generally in weakened states also suggested in previous

study.[27] As a result, nearly all the controls selected also had electrolyte imbalance, with high percentages of electrolyte imbalance found in both the cases (69.23%) and the controls (59.62%). Therefore, the results showed no statistical significance between the two groups in our study. To improve the representation, a more diverse control group could be included to ensure a broader variety of participants.

**Strengths and limitations of the present study**

The main strengths of this study were its design as a case-control study, which is considered the most reliable methodology for rare syndromes like NMS, and also our delving deeply into each factor with an adequate sample size (NMS cases=26) to address our main objective, which was to prove that a high neuroleptic dose is a risk factor for developing NMS. Specifically, we achieved this by evaluating "the maximum neuroleptic dose" as a risk factor for NMS. This study chose an inpatient department (IPD) population, which allowed us to easily obtain all the data needed for the study, especially details on the dose, the route, and the period of neuroleptic administrations. By using an adequate sample size, our results allowed us to determine the statistical significance in the main objective, and we were able to prove that receiving a high neuroleptic dose is indeed a risk factor for developing NMS. We also found that being in a state of "dehydration" is a risk factor for NMS too.

There are some limitations of our study to note with some factors, such as neuroleptic increase final 5 days, number of intramuscular neuroleptic injections, whereby the results showed trends toward significance, but did not show statistical significance, possibly due to an inadequate sample size. We advise that further study should be done with an increased sample size to create a greater possibility of accurately determining the correlation between these factors and NMS. Also, the sample group in this study mostly consists of elderly patients, most of whom have delirium and did not use long-acting neuroleptic injection, the findings may not be generalizable to other populations, such as younger patients. Regarding alcohol dependence, the lack of statistical significance could be attributed to the similar characteristics between the cases and controls, as both groups primarily included alcohol-related conditions. Consequently, we were unable to distinguish alcohol dependence as an independent risk factor between the two groups.

## CONCLUSION

According to our study, we found four statistically significant risk factors between the two groups (cases and controls) for developing NMS, which were a maximum neuroleptic level higher than 200 mg/day, dehydration, number of intramuscular injections, and neuroleptic increase final 5 days (≥0 mg). However, after calculating the adjusted odds ratio to smooth out the confounding factors, we found that only **a maximum neuroleptic dose higher than 200 mg/day and dehydration were risk factors for developing NMS**. Our findings suggest avoiding the use of high-dose neuroleptics and dehydration in patients administered with neuroleptic drugs.

### Data Availability Statement

The data supporting the findings of this original article are available within the article.

## ACKNOWLEDGEMENTS

The author would like to express our gratitude to Kuntapat Rachchaiya M.D., Head of Psychiatric Unit and Krairerk Sintavanuruk M.D., Head of Department of Research Development, Charoenkrung Pracharak Hospital for their wise counsel and useful suggestions. We also thank Ratirat Sayanhansa RN., Secretary of the Department of Research Development and Waraporn Netphrao, Statistician at the Department of Research Development, Charoenkrung Pracharak Hospital for advice and support with the statistical analysis.

## DECLARATIONS

### Grants and Funding Information

Our study received research grants from Charoenkrung Pracharak Hospital.

### Conflict of Interest

The author declares that they have no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

### Registration Number of Clinical Trial

None

### Author Contributions

Conceptualization and methodology, S.S.; Data collection, S.S.; Formal analysis, S.S.; Visualization and writing – original draft, S.S.; Writing – review and editing, S.S.

### Use of Artificial Intelligence

Artificial Intelligence tool was not used in this manuscript.

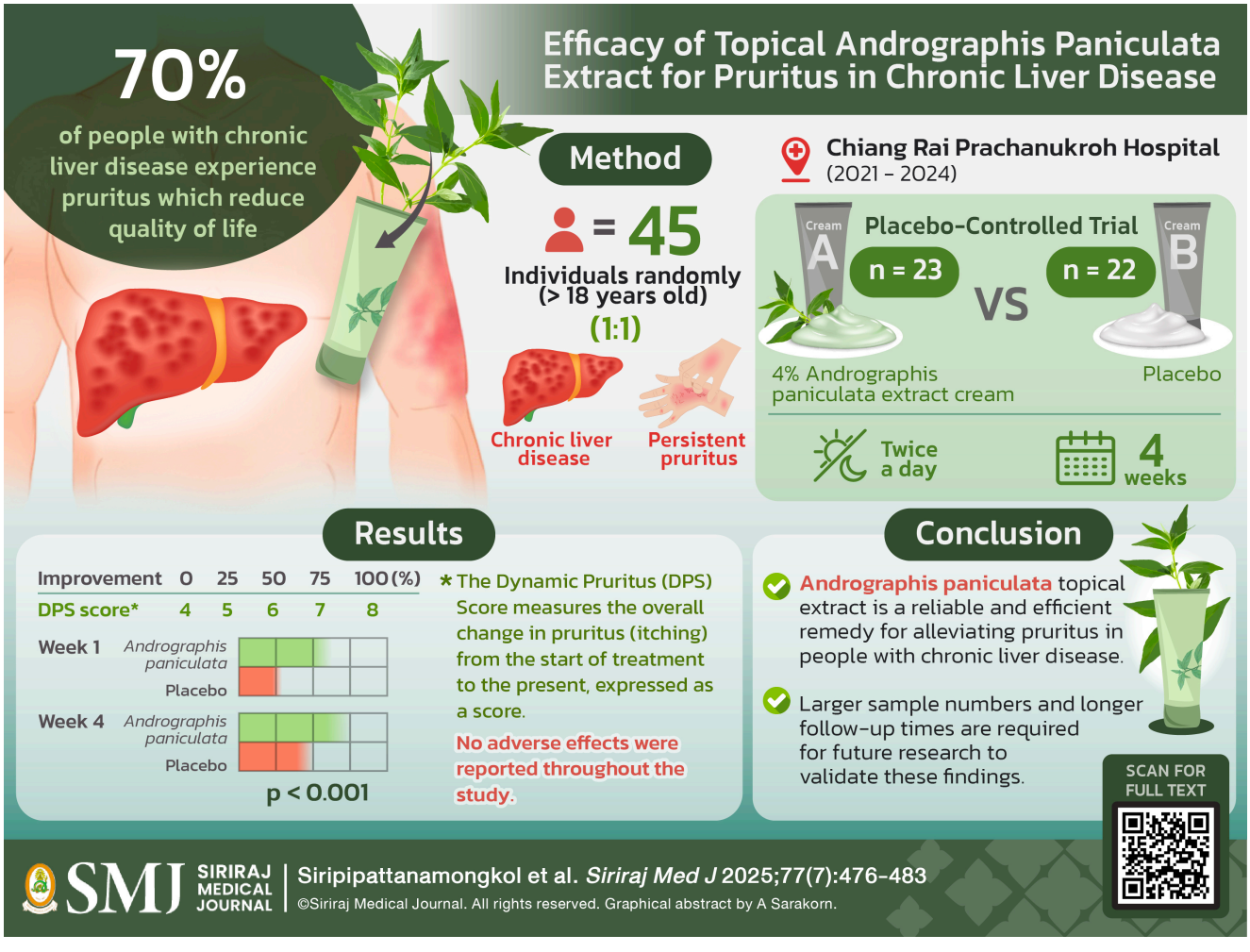
## REFERENCES


1. Methawasin K. The basic knowledge about abnormal movement. North-Eastern Thai Journal of Neuroscience. 2012;8(1):60-61.
2. Pelonero AL, Levenson JL, Pandurangi AK. Neuroleptic malignant syndrome: a review. Psychiatr Serv. 1998; 49(9):1163-72.
3. Tural U, Onder E. Clinical and pharmacologic risk factors for neuroleptic malignant syndrome and their association with death. Psychiatry Clin Neurosci. 2010;64(1):79-87.
4. Addonizio G, Susman VL, Roth SD. Symptoms of Neuroleptic malignant syndrome in 82 consecutive inpatients. Am J Psychiatry. 1986;143(12):1587-90.
5. Keck PE, Pope HG, JR, Cohen BM, McElroy SL, Nierenberg AA. Risk-factors for Neuroleptic malignant syndrome- a Case-Control study. Arch Gen Psychiatry. 1989;46(6):914-8.
6. Sachdev P, Mason C, Hadzi-Pavloric D. Case-Control study of Neuroleptic malignant syndrome. Am J Psychiatry. 1997;154(8): 1156-8.
7. Chen Y, Guo JJ, Steinbuch M, Buckley PF, Patel NC. Risk of Neuroleptic malignant syndrome in patients with bipolar disorder: a prospective, population-based case-control study. J Psychiatry Med. 2009;39(4):439-50.
8. Berardi D, Amore M, Keck PE, Troia M, Dell'Atti M. Clinical and pharmacologic risk factor for Neuroleptic malignant syndrome: A Case-Control study. Biol Psychiatry. 1998;44(8):748-54.
9. Viejo LF, Morales V, Punal P, Perez JL, Sancho RA. Risk factors in Neuroleptic malignant syndrome: A Case-Control study. Acta Psychiatr Scand. 2003;107(1):45-49.
10. Su YP, Chang CK, Hayes RD, Harrison S, Lee W, Broadbent M, et al. Retrospective chart review on exposure to psychotropic medications associated with neuroleptic malignant syndrome. Acta Psychiatr Scand. 2014;130(1):52-60.
11. Guinart D, Taipale H, Rubio JM, Tanskanen A, Correll CU, Tiihonen J, et al. Risk Factors, Incidence, and Outcomes of Neuroleptic Malignant Syndrome on Long-Acting Injectable vs Oral Antipsychotics in a Nationwide Schizophrenia Cohort. Schizophr Bull. 2021;47(6):1621-30.
12. Lao KSJ, Zhao J, Blais JE, Lam L, Wong ICK, Besag FMC, et al. Antipsychotics and Risk of Neuroleptic Malignant Syndrome: A Population-Based Cohort and Case-Crossover Study. CNS Drugs. 2020;34(11):1165-75.
13. Nielsen RE, Jensen SOW, Nielsen J. Neuroleptic malignant syndrome-an 11-year longitudinal case-control study. Can J Psychiatry. 2012;57(8):512-8.
14. Misawa F, Okumura Y, Takeuchi Y, Fujii Y, Takeuchi H. Neuroleptic malignant syndrome associated with long-acting injectable versus oral second-generation antipsychotics: Analyses based on a spontaneous reporting system database in Japan. Schizophr Res. 2021;231:42-46.
15. Langan J, Martin D, Shajahan P, Smith DJ. Antipsychotic dose escalation as a trigger for Neuroleptic malignant syndrome (NMS) : literature review and case series report. BMC Psychiatry. 2012;12:214.
16. Taemeeyapradit U, Tanchaiswad W. Neuroleptic Malignant Syndrome: A case report and literature review. J Psychiatr Ass Thailand. 1989; 34(3):205-16.
17. Kooptiwoot S, Vuthiganond S. Neuroleptic malignant syndrome in a diabetic patient with eye complication and acute psychosis. Siriraj Med J. 1999;51(6):367-72.
18. Wae-Alee D, Tanchaiswad W. Neuroleptic Malignant Syndrome: A review of 12 cases. J Psychiatr Assoc Thailand. 1996;41(2): 89-109.
19. Tantiphlachiva K. Neuroleptic Malignant Syndrome: A five-year review. J Psychiatr Ass Thailand. 1999; 44(3):189-200.
20. Kasantikul D, Kanchanatawan B. Neuroleptic malignant syndrome: A review and report of six cases. J Med Assoc Thai. 2006;89(12):2155-60.
21. Sirisuwannarat S. Neuroleptic malignant syndrome: A review and report of six cases. Journal of Charoenkrung Pracharak Hospital. 2020;16(1):1-24.
22. Shalev A, Munitz H. The neuroleptic malignant syndrome: agent and host interaction. Acta Psychiatr Scand. 1986;73(4):337-47.
23. Addonizio G, Susman VL, Roth SD. Neuroleptic malignant syndrome: review and analysis of 115 cases. Biol Psychiatry. 1987;22(8):1004-20.
24. Caroff SN. The neuroleptic malignant syndrome. J Clin Psychiatry. 1980;41(3):79-83.
25. Levenson JL. Neuroleptic malignant syndrome. Am J Psychiatry. 1985;142(10):1137-45.
26. Pearlman CA. Neuroleptic malignant syndrome: a review of the literature. J Clin Psychopharmacol. 1986; 6(5):257-73.
27. Wedzicha JA, Hoffbrand BI. Neuroleptic malignant syndrome and hyponatremia. Lancet. 1984;1:963.
28. Kato D, Kawanishi C, Kishida I, Furuno T, Suzuki K, Onishi H. Effects of CYP2D6 polymorphisms on neuroleptic malignant syndrome. Eur J Clin Pharmacol. 2007;63(11):991-6.
29. Bradford LD. CYP22D6 allele frequency in European Caucasians, Asians, Africans and their descendants. Pharmacogenomics. 2002;3(2):229-43.
30. Neuroleptic malignant syndrome. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Text revision. Washington, DC: American Psychiatric Association 2022.p.810-11.
31. Caroff SN, Mann SC. Neuroleptic malignant syndrome. Psychopharmacol Bull. 1988;24(1):25-29.
32. Davis JM. Dose equivalence of the antipsychotic drugs. J Psychiatr Res. 1974;11:65-69.
33. Lotrakul M, Ittasakul P. Antipsychotic drug. In: Lotrakul M, editor. Clinical use of Psychotropic drugs. 1st ed. Beyond Enterprise; 2017.p.80.

# Efficacy of Topical *Andrographis Paniculata* Extract for Pruritus in Chronic Liver Disease: A Randomized, Placebo-Controlled Trial

**Chotipong Siripipattanamongkol, M.D.[1,*], Vich Thampanya, M.D.[1], Piyaporn Sirijanchune, M.D.[1], Nonlawan Chueamuangphan, M.D.[1], Watcharee Khunkitti, Ph.D.[2]**

*[1]Department of Internal Medicine, Chiang Rai Prachanukroh Hospital, Chiang Rai 57000, Thailand, [2]Faculty of Pharmaceutical Sciences, Khon Kaen University, Khon Kaen 40002, Thailand.*




**Corresponding author: Chotipong Siripipattanamongkol*
*E-mail: dr.chotipong@gmail.com*
*Received 24 February 2025 Revised 14 March 2025 Accepted 14 March 2025*
*ORCID ID:http://orcid.org/0009-0001-0836-4752*
*https://doi.org/10.33192/smj.v77i7.273826*

## ABSTRACT

**Objective:** Assessing the effectiveness of topical extract from *Andrographis paniculata*, which has anti-inflammatory properties, in reducing pruritus in individuals with chronic liver disease.

**Materials and Methods:** At Chiang Rai Prachanukroh Hospital, a randomized, placebo-controlled trial was carried out from 2021 to 2024. Forty-five individuals with chronic liver disease and persistent pruritus who were at least eighteen years old were randomly assigned (1:1) to receive 4% extract cream from Andrographis paniculata or a placebo, which was applied twice a day for four weeks. The dynamic pruritus score (DPS), which was used to measure pruritus severity at baseline, week 1, and week 4, was the main outcome. The dermatology life quality index (DLQI) was used to measure the quality of life as the secondary outcome.

**Results:** The *Andrographis paniculata* extract significantly improved the pruritus severity compared to placebo, with DPS scores of 6.31±0.41 at week 1 and 6.82±0.40 at week 4 versus 5.01±0.24 and 5.59±0.35, respectively ($p<0.001$). However, there was no significant difference in DLQI scores between treatments. There were no adverse effects reported, and one participant in the treatment arm withdrew due to logistical issues.

**Conclusion:** *Andrographis paniculata* topical extract is a reliable and efficient remedy for alleviating pruritus in individuals with chronic liver disease. However, larger sample numbers and longer follow-up times are required for future research to validate these findings and evaluate the impact on quality of life.


**Keywords:** *Andrographis paniculata*; pruritus; quality of life; chronic liver disease; topical treatment (Siriraj Med J 2025; 77: 476-483)

## INTRODUCTION

Up to 70% of people with chronic liver disease, especially those with cholestatic liver disease, experience pruritus, a common and upsetting symptom.[1] It may significantly lower the patient's standard of living, with some patients experiencing such intense itching that it leads to depression or even suicidal ideation.[2,3] Despite its prevalence, little is known about how it develops in chronic liver disease. While bile acids were historically implicated as the primary culprits due to their toxic effects on skin and nerve endings[3,4], recent studies have found lysophosphatidic acid and its precursor enzyme, autotaxin, as key mediators of itching. Elevated levels of these substances correlate with pruritus severity and decrease with symptom improvement.[5-8]

Current treatments for pruritus in chronic liver disease are limited and often inadequate. While addressing the underlying cause, such as removing bile duct obstructions, can alleviate symptoms, numerous cholestatic conditions, including progressive familial intrahepatic cholestasis, primary biliary cirrhosis, and primary sclerosing cholangitis, are incurable.[9,10] Available pharmacological options, such as cholestyramine, rifampicin, naloxone, and sertraline, are taken orally and associated with significant side effects, some of which can be severe or life-threatening.[11-15] These drawbacks highlight how urgently safer and more efficient treatments are needed to control itching in this patient group.

*Andrographis paniculata*, a traditional medicinal herb, has garnered attention for its potent anti-inflammatory properties.[16-23] The cytokines TNF-α, IL-1β, and IL-6 that are linked to itching and inflammation are inhibited by its active ingredients, which include andrographolide and related diterpenes.[16-19] Preclinical studies have proved the herb's efficacy in reducing inflammation in animal models[20,21], and studies on humans have demonstrated potential in reducing the symptoms of diseases such as upper respiratory infections and rheumatoid arthritis.[22,23] Although evidence specific to dermatological applications is limited, *Andrographis paniculata*'s immunomodulatory and anti-inflammatory properties point to its potential as a pruritus treatment.

The purpose of this study was to assess Andrographis paniculata's effectiveness as a topical remedy for pruritus in individuals suffering from chronic liver disease. It is anticipated that this novel, complementary approach will improve the quality of life for individuals suffering from this debilitating symptom due to its anti-inflammatory properties and potential to alleviate itching.

## MATERIALS AND METHODS

### Study population and design

This randomised, placebo-controlled study was conducted at Chiang Rai Prachanukroh Hospital, Thailand, from 2021 to 2024. All participants provided written informed consent before enrolling in the study, and the

research protocol was registered with the Thai Clinical Trial Registry (TCTR20250206011) after being approved by the Chiang Rai Prachanukroh Hospital's Research Ethics Committee (approval number: EC CRH 125/63In).

The following were the requirements for inclusion: 18 years of age or older, with a diagnosis of cholestatic or chronic liver disease, and a history of widespread itching that lasts longer than two weeks. The exclusion criteria included: a known allergy to *Andrographis paniculata;* the use of topical steroids within the past 2 weeks; impaired consciousness due to hepatic encephalopathy or other conditions; life expectancy of fewer than 4 weeks; pregnancy or breastfeeding; blood pressure below 90/60 mmHg; low platelet count (<50,000/mm$^3$) or abnormal blood clotting (INR >1.5); or an underlying skin diseases (e.g., atopic dermatitis, psoriasis, dermatophytosis, candidiasis, or chronic skin infections).

All participants continued to receive standard treatments for pruritus, including antihistamines, ursodeoxycholic acid, and moisturising cream throughout the study period.

Participants were randomly assigned in equal proportions using a computer-generated block randomization method to receive either a 4% *Andrographis paniculata* extract cream or a placebo moisturizing cream. The creams were packaged in identical 30 g white tubes labelled as "Drug A" or "Drug B" with the active cream being light green while the placebo was white, ensuring visual similarity. Researchers, outcome assessors, and participants were blinded to treatment assignments throughout the study.

**Active Treatment:** The 4% *Andrographis paniculata* extract cream was prepared by SBU Corporation and contained standardised concentrations of active compounds including andrographolide and related diterpenes.

**Placebo:** The placebo cream was a moisturising cream with identical appearance, texture, and packaging to the active treatment but lacked the active ingredients.

Participants were instructed to apply the assigned cream to the affected area twice daily for four weeks (total dose: 360 g of cream). Compliance was checked through self-reporting and tube weight measurements at follow-up visits.

### Pruritus Severity

The primary outcome was the change in pruritus severity assessed using the dynamic pruritus score (DPS) at baseline, 1 week, and 4 weeks after treatment initiation. The DPS is a validated tool that captures fluctuations in pruritus intensity over time by incorporating both peak and average itch severity, providing a comprehensive evaluation of symptom changes.[25]

### The DLQI, or Dermatology Life Quality Index

The variance in quality of life, as measured by the DLQI in Thai at the same intervals, was the secondary result. Higher scores indicate a greater loss in quality of life. The DLQI is a well-known 10-item questionnaire that assesses how skin disorders affect patients' everyday activities, emotional health, and social interactions.[26-29]

### Analysis of statistics

Using a 5% alpha error, we determined that each group should initially consist of 19 participants to ensure there is an 80% chance of detecting a 50% difference in pruritus severity between the groups. Considering an expected dropout rate of 15%, the final sample size for each group will be adjusted to 22 individuals. Descriptive statistics were used to summarize the baseline clinical and demographic features. Depending on the data distribution, continuous variables are either provided as mean ± standard deviation (SD) or median (interquartile range, IQR). The Student's t-test or the Wilcoxon rank-sum test are used for analysis. Frequencies and percentages are used to compare categorical variables using the chi-square or Fisher's exact test. Repeated-measures ANOVA or appropriate non-parametric techniques were used to evaluate changes in DPS and DLQI scores over time. Stata Version 14 and the intention-to-treat (ITT) method were used for all statistical computations, and a $p < 0.05$ was deemed statistically significant.

## RESULTS

### Study population

Initially, 63 patients were assessed for eligibility, of which, 18 patients were excluded, with 12 not meeting the inclusion criteria, 5 were unable to adhere to the follow-up protocol, and 1 declined to take part (Fig 1). Consequently, 45 patients were enrolled and randomised into two groups: the *Andrographis paniculata* topical treatment (n=23) or placebo (n=22). One participant in the treatment group dropped out due to logistical challenges, resulting in a final cohort of 44 patients (n=22 in each group) completing the study.

### Baseline characteristics

Table 1 summarizes the baseline characteristics of the patients and reveals no significant differences between the groups. With comparable Child-Turcotte-Pugh scores in both groups, the causes of chronic liver disease (such as viral or alcoholic hepatitis) and the presence of

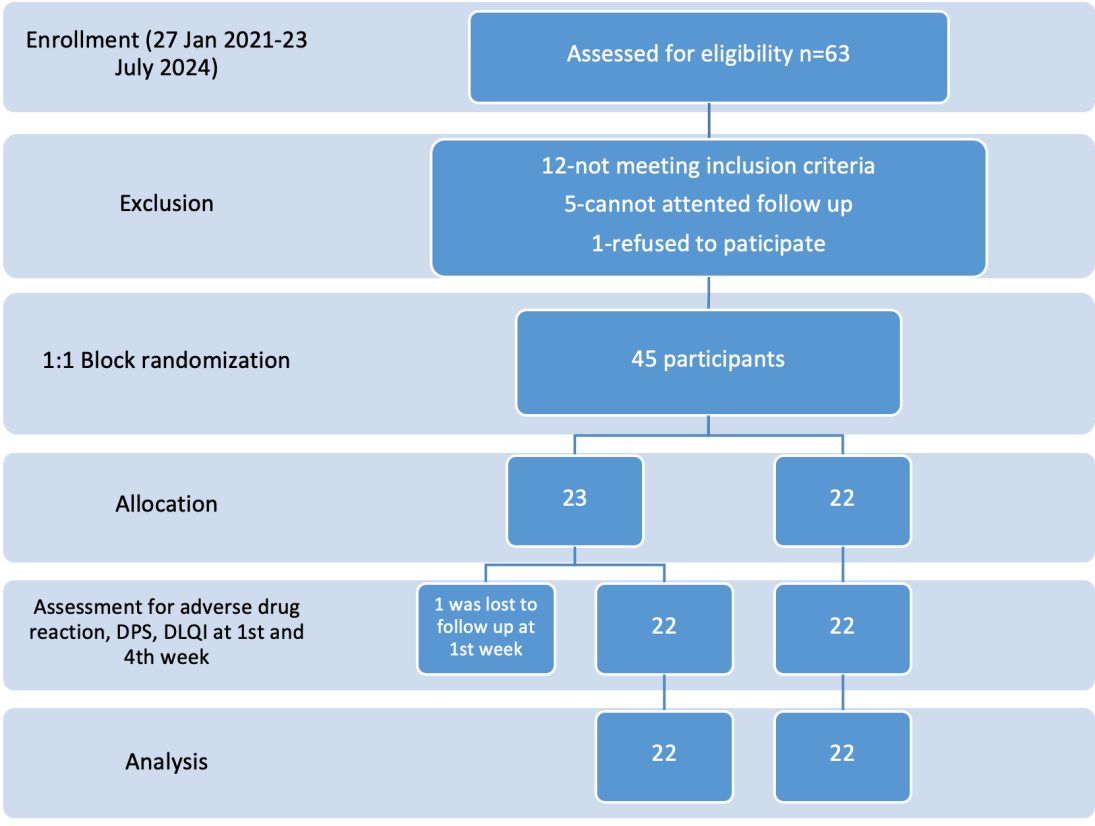



**Fig 1.** Study flowchart
**Abbreviations:** DPS: Dynamic pruritus score, DLQI: Dermatology Life Quality Index

comorbidities (such as diabetes, hypertension, or chronic kidney disease) were likewise equally distributed.

Overall, the mean duration of pruritus was 5.12 ± 7.60 weeks with a shorter duration in the treatment group (3.39 ± 2.10 weeks) compared to the placebo (7.00 ± 10.46 weeks), although this difference was not statistically significant (p = 0.112). At the outset, both groups experienced moderate quality of life impairment, with DLQI scores of 9.74 in the treatment group and 7.45 in the placebo group (p = 0.289).

### Effects of *Andrographis paniculata* compared to placebo

Participants treated with Andrographis paniculata showed significant improvement in pruritus severity compared to the placebo group (Table 2). At week 1, the mean DPS was 6.31 for the treatment group versus 5.04 for the placebo group. By week 4, the mean DPS increased to 6.82 in the treatment group compared to 5.59 in the placebo group (p < 0.001).

Over the course of the trial, both groups demonstrated improvements in quality of life as measured by the DLQI, with the treatment group showing a trend of more noticeable improvement, however there was no statistical significance (Table 2). The treatment group's baseline DLQI scores were 9.74 ± 1.64, while the placebo group's were 7.45 ± 1.33. Despite ongoing progress, the treatment group did not vary significantly from the placebo group (p = 0.096).

### Side effects

Both creams were well-tolerated by all participants with no adverse effects reported throughout the study period.

## DISCUSSION

This randomised, placebo-controlled study is the first to evaluate the efficacy of topical *Andrographis paniculata* extract cream for relieving itching in patients with chronic liver disease, showing that *Andrographis paniculata* significantly reduced pruritus severity, as measured by the DPS, compared to placebo. However, this was not accompanied by any considerable improvement in their quality of life. These findings suggest that *Andrographis paniculata* is a safe and effective adjunctive treatment for pruritus in this patient population.

The anti-pruritic effects of *Andrographis paniculata* observed align with its known anti-inflammatory and immunomodulatory properties. The herb contains

**TABLE 1.** Participant's baseline characteristics.

| Characteristic | *Andrographis paniculata* extract (N=23) | Placebo (N=22) | *p*-value |
|---|---|---|---|
| Age | 58.95 ± 11.77 | 57.40 ± 10.45 | 0.644 |
| Male, (n, %) | 14 (60.87) | 9(40.91) | 0.149 |
| BMI | 23.22 ± 3.36 | 24.51 ± 4.00 | 0.249 |
| Cause of chronic liver disease/cholestasis, n (%) | | | |
| Alcoholic hepatitis | 7 (30.43) | 9(40.91) | 0.779 |
| PBC/autoimmune hepatitis | 4 (17.39) | 3(13.64) | |
| Cholangiocarcinoma | 3 (13.04) | 1(4.55) | |
| Bile duct stone | 1 (4.35) | 0 (0) | |
| Viral hepatitis | 7 (30.43) | 6 (27.27) | |
| Drug-induced liver injury | 0 (0) | 1 (4.55) | |
| MASH with cirrhosis | 0 (0) | 1 (4.55) | |
| Idiopathic | 1(4.35) | 0 (0) | |
| Hypertension, n (%) | 4 (17.39) | 8 (36.36) | 0.135 |
| Diabetes, n (%) | 3 (13.04) | 5 (22.73) | 0.324 |
| Hyperlipidemia, n (%) | 1 (4.35) | 1 (4.55) | 0.744 |
| Chronic kidney disease, n (%) | 3 (13.04) | 4 (18.18) | 0.474 |
| HIV infection, n (%) | 1 (4.35) | 0 (0) | 0.511 |
| Atopic dermatitis, n (%) | 1 (4.35) | 0 (0) | 0.511 |
| Duration of pruritus (week) | 3.39 ± 2.10 | 7.00 ± 10.46 | 0.112 |
| Total bilirubin (mg/dL) | 8.04 ± 8.57 | 5.07 ± 6.21 | 0.192 |
| Albumin (g/dL) | 3.41 ± 0.62 | 3.19 ± 0.61 | 0.214 |
| Alkaline phosphatase (U/L) | 199.37 ± 128.28 | 239.23 ± 163.61 | 0.367 |
| Absolute eosinophil count (cell/mm$^3$) | 238.09 ± 193.23 | 640.87 ± 1353.26 | 0.164 |
| Child Turcotte Pugh score status, n (%) | | | 0.930 |
| Class A | 10 (43.48) | 8 (36.36) | |
| Class B | 5 (21.74) | 6 (27.27) | |
| Class C | 8 (34.78) | 8 (36.36) | |
| Dermatology Life Quality Index at baseline | 9.74 ± 7.88 | 7.45 ± 6.25 | 0.289 |

Data is presented as mean ± standard deviation

**Abbreviations:** BMI, body mass index; HIV, human immunodeficiency virus; MASH, metabolic associated steatohepatitis; PBC, primary biliary cirrhosis; SD, standard deviation

**TABLE 2.** A comparison of the clinical outcomes of *Andrographis paniculata* extract vs placebo using the dynamic pruritus score and dermatology life quality index.

| Outcome assessment | Andrographis paniculata extract (N=23) | Placebo (N=22) | Change in mean (95%CI) | *p*-value |
|---|---|---|---|---|
| Dynamic pruritus score (DPS)** | | | | <0.001* |
| At baseline | reference | reference | | |
| At first week | 6.31 ± 0.41 | 5.04 ± 0.24 | **-1.27(-2.28,-0.32)** | **0.010*** |
| At fourth week | 6.82 ± 0.40 | 5.59 ± 0.35 | **-1.23(-2.31,0.15)** | **0.027*** |
| Within-group p | <0.001* | <0.001* | | |
| Dermatology Life Quality Index (DLQI)*** | | | | 0.096 |
| At baseline | 9.74±1.64 | 7.45±1.33 | **-2.28(-6.57,2.00)** | **0.289** |
| At first week | 4.30 ± 1.23 | 5.32 ± 1.18 | **1.01(-2.50,4.46)** | **0.555** |
| At fourth week | 3.52 ± 1.27 | 4.5 ± 1.32 | **0.98(-2.71,4.67)** | **0.596** |
| Within-group p | <0.001* | <0.001* | | |

Data is presented as mean ± standard error
*Shows statistically significant difference
** The Dynamic Pruritus Score measures the overall change in pruritus (itching) from the start of treatment to the present, expressed as a score. Scoring 4 indicates no change from the baseline, while scores of 5, 6, 7, and 8 represent 25%, 50%, 75%, and 100% improvement, respectively. Since the scoring is based on changes relative to the baseline, the baseline scores themselves are not displayed.
*** Dermatology Life Quality Index, a higher score means more worsening impact on quality of life

active compounds such as andrographolide and dehydroandrographolide, which inhibit pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6, and deactivate histamine release and mast cell activity[16-20,24] to alleviate itching. Previous study had also proved the effectiveness of *Andrographis paniculata* in reducing itching in end-stage renal disease patients[30], supporting its potential utility across different conditions characterised by chronic itching.

The significant reduction in itching severity due to the topical administration of the *Andrographis paniculata* extract underscores its therapeutic potential. However, the DLQI scores did not significantly improve, possibly due to the small sample size restricting the statistical power needed to find a subtle improvement in the quality of life. Furthermore, patients with advanced liver disease often experience systemic symptoms such as fatigue and malnutrition, which may overshadow dermatologic improvements and limit the sensitivity of the DLQI in capturing the full impact of pruritus relief.[26-29]

This study's rigorous randomised, placebo-controlled design and use of standardised end measures (DLQI and DPS) are two of its many strong points. Additionally, the absence of adverse effects highlights the safety of topical *Andrographis paniculata* extract in this population. However, it is not without its limitations, such as the small sample size and the short treatment duration (four weeks) may have constrained the ability to detect significant changes in quality of life or long-term effects. Furthermore, the use of pre-existing anti-pruritic medications may have introduced confounding variables. To verify these results and investigate the processes behind *Andrographis paniculata*'s anti-pruritic benefits, future studies should employ larger sample numbers, longer follow-up times, and stricter management of concurrent drugs.

This study has important clinical implications. Pruritus in chronic liver disease is difficult to manage with few safe and effective treatments. Topical *Andrographis paniculata* extract shows promise as a well-tolerated, effective therapy for reducing itching. Moreover, its inclusion in Thailand's National List of Essential Medicines for other uses supports its safety and potential for wider clinical application.[31]

In conclusion, the *Andrographis paniculata* topical extract is a reliable and efficient remedy for alleviating pruritus in individuals with chronic liver disease. Its strong anti-pruritic effects and favourable safety profile make it a valuable adjunctive therapy. More research is necessary to confirm these findings, investigate the mechanisms of action, and assess the long-term benefits in larger and more diverse patient groups.

### Data Availability Statement

Due to the regulations of the Institutional Review Board (IRB) at Chiang Rai Prachaukroh Hospital and the privacy rules of the Health Insurance Portability and Accountability Act (HIPAA), the datasets generated and analysed during the current study are not publicly available. However, they can be accessed upon reasonable request from the corresponding author.

## ACKNOWLEDGEMENTS

Thanks to Dr Panjapone Kobpungton for his invaluable guidance throughout this research and Dr Vich Thampanya for highlighting the significance of *Andrographis paniculata* and facilitating collaboration with the pharmaceutical company involved in this project.

## DECLARATIONS

### Grants and Funding Information

This research received no grant or funding support.

### Conflict of Interest

The *Andrographis paniculata* extract cream and placebo used in this study were manufactured by SBU Corporation Co., Ltd. The researcher has no financial interest or shareholder position in the company.

### Registration Number of Clinical Trial

Thai Clinical Trial Registry (TCTR20250206011).

### Author Contributions

Conceptualization and methodology, CS, VT, PS, NC; Formal analysis, PS, NC; Investigation; CS, VT; Visualization and writing – original draft, CS; Writing – review and editing, CS; All authors have read and agreed to the final version of the manuscript.

### Use of Artificial Intelligence

AI was used to assist with translation from Thai to English in some parts of introduction and first part of discussion.

## REFERENCES

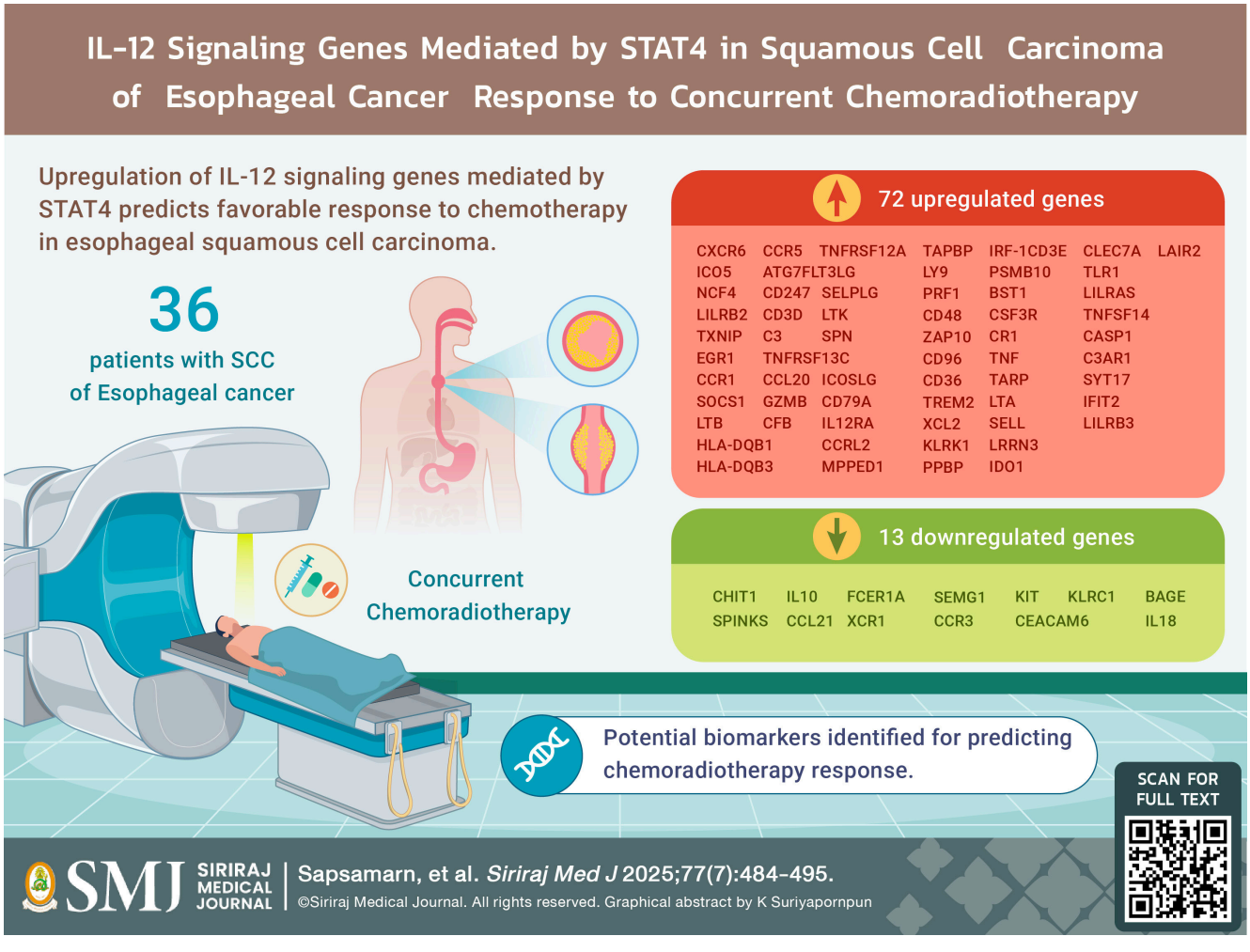
1. Mela M, Mancuso A, Burroughs AK. Review article: pruritus in cholestatic and other liver diseases. Aliment Pharmacol Ther. 2003;17(7):857–70.
2. Rishe E, Azarm A, Bergasa NV. Itch in primary biliary cirrhosis: a patient 'perspective. Acta Derm Venereol. 2008;88(1):34–37.
3. Bergasa NV. The pruritus of cholestasis. J Hepatol. 2005;43(6): 1078–88.
4. Ghent CN, Bloomer JR, Klatskin G. Elevations in skin tissue levels of bile acids in human cholestasis: relation to serum levels and pruritus. Gastroenterology. 1977;73(5):1125–30.
5. Kremer AE, Martens JJ, Kulik W, Rueff F, Kuiper EMM, van Buuren HR, et al. Lysophosphatidic acid is a potential mediator of cholestatic pruritus. Gastroenterology. 2010;139(3):1008–18.
6. Hagermark O. Peripheral and central mediators of itch. Skin Pharmacol. 1992;5(1):1–8.
7. van Meeteren LA, Moolenaar WH. Regulation and biological activities of the autotaxin-LPA axis. Prog Lipid Res. 2007;46(2): 145–60.
8. Hashimoto T, Ohata H, Momose K. Itch-scratch responses induced by lysophosphatidic acid in mice. Pharmacology. 2004;72(1): 51–56.
9. Bunchorntavakul C, Reddy KR. Pruritus in chronic cholestatic liver disease. Clin Liver Dis. 2012;16:331–46.
10. Kremer AE, Oude Elferink RP, Beuers U. Pathophysiology, and current management of pruritus in liver disease. Clin Res Hepatol Gastroenterol. 2011;35(2):89–97.
11. Ikoma A, Steinhoff M, Ständer S, Yosipovitch G, Schmelz M. The neurobiology of itch. Nat Rev Neurosci. 2006;7(7):535–47.
12. Kremer AE, Beuers U, Oude-Elferink RP, Pusl T. Pathogenesis and treatment of pruritus in cholestasis. Drugs. 2008;68:2163–82.
13. Vinod S Hegade, Stuart FW Kendrick, David EJ Jones. Drug treatment of pruritus in liver diseases. Clin Med. 2015;15(4):351–7.
14. Javitt NB. Letter: timing of cholestyramine doses in cholestatic liver disease. N Engl J Med. 1974;290(23):1328–9.
15. George RB, Allen TK, Habib AS. Serotonin receptor antagonists for the prevention and treatment of pruritus, nausea, and vomiting in women undergoing cesarean delivery with intrathecal morphine: a systematic review and meta-analysis. Anesth Analg. 2009;109(1): 174–82.
16. Thisoda P, Rangkadilok N, Pholphana N, Worasuttayangkurn L, Ruchirawat S, Satayavivad J. Inhibitory effect of Andrographis paniculata extract and its active diterpenoids on platelet aggregation. Eur J Pharmacol. 2006;553:39-45.
17. Sheeja K, Shihab PK, Kuttan G. Antioxidant and anti-inflammatory activities of the plant *Andrographis Paniculata Nees*. Immunopharmacol Immunotoxicol. 2006;28(1):129-40.
18. Song CH, Lee MH, No KO, Kang SJ. Andrographolide or Andrographis paniculata extract containing composition having an anti-itching effect (Patent No. KR20030098988). Republic of Korea. Issued December 29, 2003.
19. Kumar, RA, Sridevi K, Kumar NV, Nanduri S, Rajagopal S. Anticancer and immunostimulatory compounds from *Andrographis paniculata*. J Ethnopharmacol. 2004;92(2–3):291–5.
20. Zou W, Xiao Z, Wen X, Luo J, Chen S, Cheng Z, et al. The anti-inflammatory effect of *Andrographis paniculata* (Burm. f.) Nees on pelvic inflammatory disease in rats through down-regulation of the NF-κB pathway. BMC Complement Altern Med. 2016;16(1):483.

21. Li Z, Tan J, Wang L, Li Q. Andrographolide benefits rheumatoid arthritis via inhibiting MAPK pathways. Inflammation. 2017;40(5):1599–605.
22. Cáceres DD, Hancke JL, Burgos RA, Sandberg F, Wikman GK. Use of visual analogue scale measurements (VAS) to assess the effectiveness of standardized *Andrographis paniculata* extract SHA-10 in reducing the symptoms of the common cold. A randomized double-blind placebo study. Phytomedicine. 1999;6(4):217–23.
23. Burgos RA, Hancke JL, Bertoglio JC, Aguirre V, Arriagada S, Calvo M, et al. Efficacy of an *Andrographis paniculata* composition for the relief of rheumatoid arthritis symptoms: A prospective randomized placebo-controlled trial. Clin Rheumatol. 2009; 28(8):931–46.
24. Che D, Hou Y, Zeng Y, Li C, Zhang Y, Wei D, et al. Dehydroandrographolide inhibits IgE-mediated anaphylactic reactions via the calcium signaling pathway. Toxicol Appl Pharmacol. 2019;366:46–53.
25. Ständer S, Blome C, Anastasiadou Z, Zeidler C, Jung KA, Tsianakas A, et al. Dynamic pruritus score: evaluation of the validity and reliability of a new instrument to assess the course of pruritus. Acta Derm Venereol. 2017;97(2):230–4.
26. Zou Q, Luo Y, Hao D, Li M, Jihui C. Validation and application of the Dermatology Life Quality Index score, a modification of the DLQI score, in psoriasis patients. J Health Popul Nutr. 2024; 43(1):92.
27. Kulthana K, Jiamton S, Wanitphakdeedecha R, Chantharujikaphong S. The validity and reliability of the dermatology life quality index (DLQI) in Thais. Thai J Dermatol. 2004; 20:113–23.
28. Kulthanan K, Jiamton S, Kittisarapong R. Dermatology life quality index in Thai Patients with Acne. Siriraj Med J. 2007;59:3–7.
29. Noppakun N, Rajatanavin N, Suthipinittharm P, Puvabanditsin P, Akaraphanth R, Tuchinda C, et al. Clinical Practice Guideline for Psoriasis. Dermatological Society of Thailand 2011. Retrieved May 30, 2020, Available from: http://www.dst.or.th/files_news/007-Guideline_Psoriasis_2011.pdf.
30. Kobpungton P, Thampanya V, Sirijanchune P, Chueamuangphan N. Effect of topical *Andrographis paniculata* extract on end-stage renal disease pruritus: A randomized controlled trial. The Clinical Academia. 2021;45(2):53–59.
31. Rangkadilok N, Pholphana N, Suriyo T, Satayavivad J. ฟ้าทะลายโจร (Andrographis paniculata) – ข้อมูลวิชาการที่น่ารู้ [Andrographis paniculata – Academic information worth knowing]. Chulabhorn Research Institute 2016. Retrieved May 30, 2020, from http://www.eht.sc.mahidol.ac.th/-article/1818.

# Upregulation of IL-12 Signaling Genes Mediated by STAT4 Predicts a Complete Clinical Response in Patients with Squamous Cell Carcinoma of the Esophagus who Undergo Concurrent Chemoradiotherapy

Nattida Sapsamarn, M.D.[1,2], Asada Methasate, M.D., Ph.D.[3], Pariyada Tanjak, Ph.D.[3,4], Thanawat Suwatthanarak, Ph.D.[3,4], Onchira Acharayothin[3], Kullanist Thanormjit[3,4], Jirawat Swangsri, M.D., Ph.D.[3], Thammawat Parakonthun, M.D.[3], Tetsuya Taga, M.D., Ph.D.[5], Vitoon Chinswangwatanakul, M.D., Ph.D.[3,4,*]

*[1]Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok. [2]Institute of Science Tokyo, Japan. [3]Division of General Surgery, Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [4]Siriraj Cancer Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. [5]Department of Stem Cell Regulation, Medical Research Institute, Institute of Science of Tokyo, Tokyo, Japan.*




*Corresponding author: Vitoon Chinswangwatanakul
E-mail: vchinswa@gmail.com
Received 20 December 2024 Revised 5 Jaunary 2025 Accepted 9 Jaunary 2025
ORCID ID:http://orcid.org/0000-0001-9662-1669
https://doi.org/10.33192/smj.v77i7.272789

## ABSTRACT

**Objective:** This study compared immune transcriptome profiles between SCC patients achieving cCR and those with incomplete responses to CRT using the NanoString Ncounter® Pancancer Immune Profiling Panel.

**Materials and Methods:** A retrospective cohort of 36 SCC patients treated with CRT was analyzed. Clinical data and blood samples were collected, and immune transcriptome profiling was conducted. Differential gene expression analysis identified distinctions between cCR and non-clinical complete response (non-cCR) groups.

**Results:** Of the 36 patients, 20 achieved cCR, while 16 had non-cCR. Significant differences in immune transcriptomes were observed, particularly in IL-12 signaling mediated by STAT4. In the cCR group, 72 genes were upregulated, and 13 were downregulated, suggesting their role as predictive biomarkers for treatment response.

**Conclusions:** This study identified distinct immune transcriptome profiles in SCC patients with cCR versus non-cCR, highlighting genes related to IL-12 signaling as potential biomarkers. These findings emphasize the role of immune-related gene expression in determining patient outcomes and may support personalized therapeutic strategies to enhance CRT efficacy in SCC patients.


**Keywords:** Squamous cell carcinoma of the esophagus; Chemoradiotherapy; Interleukin 12 (Siriraj Med J 2025; 77: 484-495)

## INTRODUCTION

Esophageal cancer is the seventh most common cancer worldwide. It ranks as the seventh most frequently occurring cancer in men and the thirteenth most common cancer in women.[1] In Thailand, esophageal cancer is the thirteenth leading cause of cancer-related death. There are two main pathological types of esophageal cancer: squamous cell carcinoma (SCC), which can occur in the cervical, thoracic, or abdominal regions of the esophagus, and adenocarcinoma, which can develop at the junction of the esophagus and stomach (esophagogastric junction). There are various multimodal treatment options available for esophageal SCC, including endoscopic resection, primary esophagectomy, neoadjuvant chemotherapy or chemoradiotherapy (CRT) followed by surgery, and definitive CRT. The treatment of choice depends on the location and stage of the tumor.

There are two main types of immune resistance mechanisms: innate immunity and adaptive immunity. The mechanism that primarily targets tumors is adaptive immunity. Tumors can trigger specific adaptive immune resistance, which may prevent host immune response to limit their growth and spread. This immune response largely involves T cells, particularly CD8+ cytotoxic T lymphocytes (CTLs). Many tumors are surrounded by infiltrates of mononuclear cells, including T lymphocytes and macrophages. Activated lymphocytes and macrophages can also be found in the lymph nodes that drain the areas where tumors are located. The key cells involved in important cancer prevention mechanisms are T cells, macrophages, and natural killer cells.[2]

T cells, specifically T helper cells, are categorized into two types: Th1 and Th2 cells. Th1 cells secrete several cytokines, including IFN-gamma, IL-2, IL-3, TNF-alpha, TNF-beta, and the adjustment factor IFN-gamma. In contrast, Th2 cells produce cytokines such as IL-4, IL-5, IL-9, IL-10, and IL-13, especially those that regulate IL-4, IL-5, and IL-13.[2]

Interleukin-12 (IL-12) is a heterodimeric cytokine composed of two subunits: p40 and p35.[3] It is considered a proinflammatory cytokine and is produced by antigen-presenting cells such as dendritic cells and macrophages. IL-12 plays a significant role in facilitating effective antitumor immune responses. IL-12 signals through its receptors, IL-12R1 and IL-12R2, which are expressed on target cells. This signaling pathway activates downstream pathways involving Jak2 and Tyk2, leading to the phosphorylation and homodimerization of STAT4.[3] These processes are essential for the recruitment and effector functions of CD8+ T cells and natural killer cells. Given its capabilities, IL-12 is a strong candidate for immunotherapy-based interventions, as it enhances the activity of tumor-specific cytotoxic natural killer and CD8+ T cells, both of which are critical for killing tumor cells. However, systemic administration of IL-12 can be quite toxic. Therefore, alternative administration methods are being explored.[3-9] IL-12 signaling, which is mediated by STAT4, is a crucial pathway in the immune system that regulates the response to infections. IL-12 is produced by dendritic cells and macrophages and then binds to its receptor on T cells and natural killer cells before activating JAK kinases. This activation leads to the

phosphorylation of STAT4, which then dimerizes and translocates into the nucleus to promote gene expression. This pathway enhances the differentiation of Th1 cells and stimulates the release of IFN-γ, which is crucial for eliminating pathogens. Dysregulation of this pathway is associated with autoimmune diseases but can also enhance anticancer responses, indicating potential implications for cancer therapy in the future.[10,11]

A clinical complete response (cCR) occurs when no tumor remnants are detectable through nonsurgical methods, as proven by imaging studies such as esophagography, computed tomography (CT) scans, endoscopy, or positron emission tomography-CT following CRT. Neoadjuvant chemoradiotherapy increase survival rates in patients with locally advanced esophageal carcinoma. According to Monjazeb et al., achieving a cCR after concurrent CRT (CCRT) can lead to a more favorable prognosis.[12] A pathological response study in another type of cancer, specifically CA rectum, reported that downstaging of the T-category was observed in 10 patients (55.6%), while downstaging of the N-category was noted in 14 patients (77.8%) following neoadjuvant chemoradiation.

## MATERIALS AND METHODS

### 1. Study cohort

A total of 800 patients with SCC of the esophagus were enrolled in this study. All patients were over 18 years of age at the time of diagnosis. Among these patients, 36 met the study criteria. The remaining patients were excluded from our analyses. The datasets collected included various aspects of patient information and clinical data, which were part of the inclusion criteria. These datasets comprised demographic details such as age at disease onset, sex, and underlying conditions; treatments received following CCRT, categorized as surgical or nonsurgical; and laboratory results obtained prior to CCRT, including complete blood count, absolute neutrophil count, absolute lymphocyte count, neutrophil-lymphocyte ratio, absolute eosinophil count, platelet count, and platelet-lymphocyte ratio. Additional information included tumor location within the thoracic region (classified as upper, middle, or lower), tumor differentiation based on biopsy results (poorly differentiated, moderately differentiated, or well-differentiated), tumor staging, and lymph node status as per the National Comprehensive Cancer Network (NCCN) guidelines, details on preoperative radiation (dose and fractionation), preoperative chemoradiotherapy, and recurrence following CCRT. All patients included in the study were diagnosed with Stage 3 disease, meeting the inclusion criteria as defined by the NCCN guidelines.

#### *Sample selection and specimen retrieval*

The protocol used for the selection and retrieval of specimens was as follows:

- All samples must come from patients who meet the inclusion criteria.
- Clinical data for potential samples will be collected through reviews of outpatient department and inpatient department documents.
- Patient clinical information from operative records, as well as outpatient department and inpatient department documents, will be reviewed. Slides from formalin-fixed, paraffin-embedded (FFPE) blocks of tissue biopsies obtained from esophagogastroduodenoscopy will be obtained from the pathology department.

### 2. Treatment

All patients underwent upper gastrointestinal endoscopy along with a biopsy of the lesion to confirm a diagnosis of SCC. Patients received a combination of chemotherapy and radiation therapy. Imaging studies were subsequently conducted via posttreatment sequential CT scanning at approximately 4- to 6-week intervals. After being diagnosed with esophageal SCC, patients were treated with a combination of chemotherapy and radiation therapy. The chemotherapy regimens were divided into three different formulas: Formula 1: paclitaxel and carboplatin; Formula 2: cisplatin and 5-FU; and Formula 3: carboplatin and 5-FU. Radiation therapy was administered at doses ranging from 50–60 Gy. Following the chemotherapy and radiation treatments, patients who underwent CT scans were categorized into two groups on the basis of their response: (1) clinical complete response (cCR) and (2) non-clinical complete response (non-cCR). These categories were further classified according to specific criteria.

### 3. Laboratory methods for RNA extraction from FFPE tissue

#### *Deparaffinization*

FFPE slides were cut into 5-micron-thick sections, resulting in a total of 5 cuts. Deparaffinization was performed by adding 1000 µL of xylene, followed by vortexing the mixture. Next, 400 µL of absolute ethanol was added, and the mixture was vortexed again. The sample was then centrifuged in RCF mode at 16000 × g for 2 minutes. The supernatant was discarded. Subsequently, 1000 µL of absolute ethanol was added to the pellet, which was vortexed. The sample was centrifuged a second time in RCF mode at 16,000 × g for 2 minutes, and the

supernatant was discarded. Finally, the samples were dried in a heating box at 55 °C for 50 minutes.

***RNA extraction and qualification***

The process of extracting RNA from FFPE tissue via a Veracyte RNA Extraction Kit (Veracyte, San Francisco, CA, USA) began with the preparation of essential solutions. This involved adding ethanol to the wash solution and reconstituting DNase I with nuclease-free water, which was kept on ice. During the isolation phase, lysis buffer and proteinase K were added to the tissue sample. The mixture was then incubated at 56 °C and 80 °C to facilitate tissue breakdown and RNA release. Afterward, the sample was cooled on ice until it reached room temperature, followed by brief centrifugation to prepare for DNase treatment. Finally, DNase I was added to eliminate any DNA contaminants, ensuring high RNA purity for subsequent analyses. The RNA concentration and purity were measured via a Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA purity was assessed via the A260/A280 ratio, which was required to be not less than 2.0.

***Gene expression and experiments***

This study utilized an nCounter Pan-Cancer Immune Profiling Panel (NanoString Technologies, Seattle, WA, USA). This panel includes 40 reference genes as well as 730 targets related to immuno-oncology that feature 109 cell surface markers and represent 14 immune cell types.

After RNA extraction, immune transcriptome profiling was performed via the nCounter PanCancer Immune Panel. NanoString assays were performed with 100 ng of RNA, 8 µL of reporter probe, and 2 µL of capture probe per sample. The hybridization reaction was performed for 24 h at 65 °C in a thermocycler.

***NanoString data analysis***

The nSolver Analysis Software (version 4, NanoString Technologies) was utilized to assess quality control parameters, including binding density, limit of detection, and positive controls. No quality control issues were identified, allowing all 36 samples to proceed for further analysis. The advanced analysis module of the nSolver software was employed for housekeeping gene selection, data normalization, differential expression analysis, and calculation of immune-oncology-related scores. Differential gene expression analysis was conducted to identify the transcriptomic differences between patients with cCR and those with incomplete clinical responses. Rosalind software (NanoString Technologies) was subsequently used to compare the gene expression in the two patient groups. Differentially expressed genes were identified via the software. The threshold was set at a fold change greater than 1.5 combined with a $p$-value of $< 0.05$ to determine significant differentially expressed genes. To account for multiple t tests, the classical Benjamini–Hochberg method was applied to adjust $p$-values, controlling the false discovery rate and producing adjusted p-values.

***Gene set, pathway, and function analyses***

Gene set, pathway, and functional analyses were conducted via gene set analysis with a directed global significance score based on NanoString annotations version 46. Gene set analysis was used to evaluate the overall significance of functionally related gene sets, with the global significance score reflecting cumulative evidence for differential gene expression in a pathway through the square root of the mean squared t statistic. The directed global significance score considered the direction of over- or underexpression on the basis of t statistics. Biological networks were examined via Rosalind and Qiagen Ingenuity Pathway Analysis (Qiagen, Hilden, Germany). Fisher's exact test was used to calculate enrichment $p$ values, whereas Z scores were used to predict the activation or inhibition of molecular functions.

**4. Statistical analysis**

Non-normally distributed data were presented using the median and interquartile range (IQR) and analyzed with the Mann-Whitney test. Categorical data were analyzed using the Chi-square test and reported as proportions (percentages). All statistical analyses were performed using Stata version 19.1.

## RESULTS

***Patient characteristics***

The characteristics of the eligible patients are summarized in Table 1. The median age of patients in the cCR group and the non-cCR group was 64 years. Many of these patients had underlying conditions, including hypertension and diabetes. In terms of treatment, the cCR group consisted of 10 patients, some of whom underwent surgery. Similarly, the non-cCR group consisted of 8 patients, with a comparable distribution of surgical and nonsurgical treatments.

Data collection from laboratory results revealed no statistically significant differences between the two groups. The laboratory findings included the complete blood count, absolute neutrophil count, absolute lymphocyte count, neutrophil-lymphocyte ratio, absolute eosinophil count, platelet count, and platelet-lymphocyte ratio.

**TABLE 1.** Characteristics of patients with a cCR and those with a non-cCR.

| | Clinical complete response (n=20) | Non-clinical complete response (n=16) | *p* value |
|---|---|---|---|
| Sex | | | 0.106 |
| Male | 17 (85%) | 16 (100%) | |
| Female | 3 (15%) | 0 (0%) | |
| Age, mean ± SD. | 63.9 ± 9.04 | 64.13 ± 8.69 | 0.94 |
| U/D | 13 (65%) | 12 (75%) | 0.517 |
| HT | 8 (40%) | 8 (50%) | 0.549 |
| DM | 3 (15%) | 1 (6.3%) | 0.406 |
| Other U/D | 9 (45%) | 9 (56.3%) | 0.502 |
| Treatment | | | 1.000 |
| Surgical | 10 (50%) | 8 (50%) | |
| Nonsurgical | 10 (50%) | 8 (50%) | |
| CBC, median (IQR) | 7900 (6430, 10280) | 8040 (6380, 12100) | 0.386 |
| Absolute neutrophil (10*3/ul), median (IQR) | 5.19 (3.82, 6.01) | 4.7 (4.05, 7.71) | 0.800 |
| Absolute lymphocyte (10*3/ul), median (IQR) | 1.71 (1.24, 2.28) | 2.03 (1.64, 2.43) | 0.278 |
| NLR, median (IQR) | 3 (2, 4) | 3 (2, 4) | 0.955 |
| Absolute eosinophile, median (IQR) | 0.17 (0.11, 0.28) | 0.22 (0.12, 0.61) | 0.262 |
| Platelet (10*3/ul), median (IQR) | 280 (230, 330) | 317 (271, 408) | 0.148 |
| PLR, median (IQR) | 195.87 (104.79, 229.6) | 171.82 (120.35, 220.85) | 0.942 |
| Tumor location | | | 0.782 |
| Upper thoracic | 5 (25%) | 4 (25%) | |
| Middle thoracic | 7 (35%) | 4 (25%) | |
| Lower thoracic | 8 (40%) | 8 (50%) | |
| Tumor differentiation | | | 0.351 |
| Poorly differentiated | 2 (10%) | 0 (0%) | |
| Moderately differentiated | 16 (80%) | 13 (81.3%) | |
| Well differentiated | 2 (10%) | 3 (18.8%) | |
| Clinical stage | | | N/A |
| 3 | 20 (100%) | 16 (100%) | |
| Clinical tumor stage | | | 0.657 |
| 2 | 1 (5%) | 1 (6.3%) | |
| 3 | 18 (90%) | 15 (93.8%) | |
| 4 | 1 (5%) | 0 (0%) | |

**TABLE 1.** Characteristics of patients with a cCR and those with a non-cCR. (Continue)

| | Clinical complete response (n=20) | Non-clinical complete response (n=16) | *p* value |
|---|---|---|---|
| LN stage | | | 0.439 |
| Negative | 4 (20%) | 5 (31.3%) | |
| Positive | 16 (80%) | 11 (68.8%) | |
| cM stage | | | N/A |
| 0 | 20 (100%) | 16 (100%) | |
| Preop RT dose Gy, mean ± SD | 52.3 ± 8.31 | 51.53 ± 3.5 | 0.73 |
| Preop RT Fr, mean ± SD. | 27.95 ± 1.88 | 26.81 ± 1.94 | 0.084 |
| Preop CMT | | | 0.271 |
| Paclitaxel and carboplatin | 9 (45%) | 4 (25%) | |
| Cisplatin and 5FU | 6 (30%) | 9 (56.3%) | |
| Carboplatin and 5FU | 5 (25%) | 3 (18.8%) | |
| Recurrence | | | |
| Local recurrence | 2 (10%) | 6 (37.5%) | 0.049* |
| Single metastasis | 1 (5%) | 6 (37.5%) | 0.014* |
| Multiple metastases | 2 (10%) | 3 (18.8%) | 0.451 |

**Abbreviations:** 5FU, 5-Fluorouracil; CBC, Comeplete Blood Count; cM, clinical Metastasis; CMT, Chemotherapy; DM, Diabetes Mellitus; Fr, Fraction; Gr, Gray; HT, Hypertension; IQR, Interquatile Rang; LN, Lymph node; NLR, NeutrophiL-Lymphocyte Ratio; PLR, Platele-Lymphocyte Ratio; RT, Radiotherpy; SD, Standard Deviation; U/D, Undelying Disease

In the cCR group, the majority of tumors were located in the upper and middle thoracic areas. In terms of tumor differentiation, moderate differentiation was most prevalent in the cCR group, with 16 patients, compared with 13 patients in the non-cCR group.

All patients in this study were at clinical stage 3. The most common tumor stage in both groups was T3. The lymph nodes were categorized into clinical lymph node–positive and lymph node–negative subgroups. There were 20 patients (55%) in the cCR group and 16 patients (44%) in the non-cCR group. None of the patients had metastasis before treatment.

The radiation dose for patients in the cCR group was 52 Gy. For those in the non-cCR group, the dose was 51 Gy. Three chemotherapy regimens were used: (1) paclitaxel and carboplatin were administered to 9 patients (25%) in the cCR group and 4 patients (11%) in the non cCR group; (2) cisplatin and 5-FU were given to 6 patients (16.6%) in the cCR group and 9 patients (25%) in the non cCR group; and (3) carboplatin and 5-FU were provided to 5 patients (13.89%) in the cCR group and 3 patients (8.33%) in the non-cCR group.

In terms of follow-up after treatment, there was a statistically significant difference in local recurrence ($p = 0.049$) and single recurrence ($p = 0.014$). The incidence of recurrence was greater in the non-cCR group than in the cCR group.

Table 2 presents a list of differentially expressed genes organized into upregulated and downregulated groups. On the left side, the "Up" section features 72 genes whose expression increased (highlighted by the red background). Some of these genes are Interleukin 12 Receptor Subunit Alpha *(IL12RA)*, C-X-C Motif Chemokine Receptor 6*(CXCR6)*, C-C Motif Chemokine Receptor 5*(CCR5)*, Tumor Necrosis Factor Receptor Superfamily Member 12A *(TNFRSF12A)*, and Inducible T Cell Costimulator*(ICOS)*. On the right side, the "Down" section highlights 13 genes with decreased expression (indicated by a green background). Examples of these genes are Chitinase 1*(CHIT1)*, Interleukin 10*(IL10)*, Fc Fragment of IgE

Receptor IA *(FCER1A)*, and Semenogelin1*(SEMG1)*. This table helps identify key genes that may contribute to different biological responses, highlighting them as potential targets for further research

The heatmap in Fig 1 displays the expression levels of various genes across patient groups. The horizontal axis represents the patients, with the blue axis indicating those who achieved a cCR and the orange axis denoting the non-cCR group. The vertical axis represents the individual genes analyzed in the study. The blue shading

| UP=72 | | | | | | | DOWN=13 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CXCR6 | CCR5 | TNFRSF12A | TAPBP | IRF-1CD3E | CLEC7A | LAIR2 | CHIT1 | IL10 | FCER1A | SEMG1 | KIT | KLRC1 | BAGE |
| ICOS | ATG7FLT3LG | | LY9 | PSMB10 | TLR1 | | SPINK5 | CCL21 | XCR1 | CCR3 | CEACAM6 | | IL18 |
| NCF4 | CD247 | SELPLG | PRF1 | BST1 | LILRA5 | | | | | | | | |
| LILRB2 | CD3D | LTK | CD48 | CSF3R | TNFSF14 | | | | | | | | |
| TXNIP | C3 | SPN | ZAP10 | CR1 | CASP1 | | | | | | | | |
| EGR1 | TNFRSF13C | | CD96 | TNF | C3AR1 | | | | | | | | |
| CCR1 | CCL20 | ICOSLG | CD36 | TARP | SYT17 | | | | | | | | |
| SOCS1 | GZMB | CD79A | TREM2 | LTA | IFIT2 | | | | | | | | |
| LTB | CFB | IL12RA | XCL2 | SELL | LILRB3 | | | | | | | | |
| HLA-DQB1 | | CCRL2 | KLRK1 | LRRN3 | | | | | | | | | |
| HLA-DRB3 | | MPPED1 | PPBP | IDO1 | | | | | | | | | |

**TABLE 2.** In patients with a cCR, the differentially expressed genes were categorized into an upregulated group (72 genes, shown in red) and a downregulated group (13 genes, shown in green).

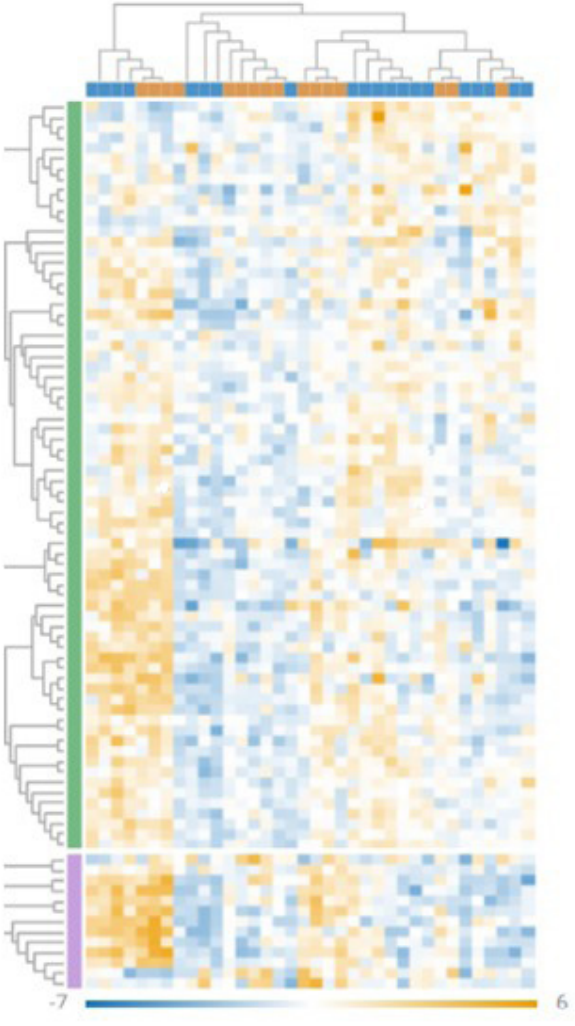



**Fig 1.** A heatmap displaying gene expression across patients, where blue represents clinical complete response (cCR) and orange indicates non-clinical complete response (non-cCR). The blue shading indicates downregulated genes, the orange shading represents upregulated genes, and the white shading signifies no change in expression.

(cool) indicates downregulated genes, whereas the orange shading (warm) indicates upregulated genes. The white cells presented no significant change in gene expression. The dendrograms on the top and left show clustering, with clustering for the top group based on the patient response profiles and clustering on the left based on the gene expression. The black markers highlight specific genes with notable expression changes and high statistical significance, suggesting potential key indicators for treatment response.

The volcano plot in Fig 2 compares gene expression between patients who achieved a cCR and those with a non-cCR. The x-axis represents the log2-fold change, indicating the ratio of gene expression between the cCR and non-cCR groups. Positive values on the right side of the x-axis represent genes with increased expression in the cCR group, whereas negative values on the left side denote genes with decreased expression. The y-axis displays the -log10 *p*-value, which represents the probability that the observed changes occurred by chance. Higher points on this axis signify changes that are statistically more significant.

The pathway interaction database in Fig 3 illustrates the signaling pathways in immune cells, with bars highlighting changes in pathway activity that are either upregulated (shown in red) or downregulated (shown in green) on the basis of adjusted p-values. The analyzed pathways included "IL12 signaling mediated by STAT4," "downstream signaling in naive CD8+ T cells," "TCR signaling in naive CD8+ T cells," "IL12-mediated signaling events," "caspase cascade in apoptosis," and "TCR signaling in naive CD4+ T cells." The length of each bar represents the adjusted *p* value, indicating the significance level of each pathway. This visualization emphasized the activity

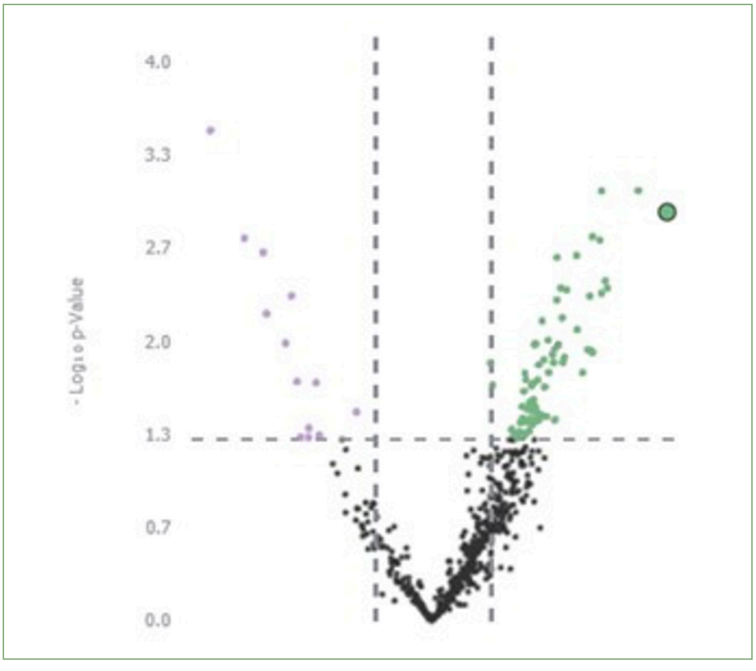



**Fig 2.** A volcano plot comparing gene expression between patients who achieved a clinical complete response (cCR) and those who did not achieve a complete response (non-cCR). Genes with increased expression in the cCR group are represented as green points on the right side of the plot, whereas genes with decreased expression are shown as purple points on the left side. The height of each point indicates the level of statistical significance, with taller points representing greater significance.

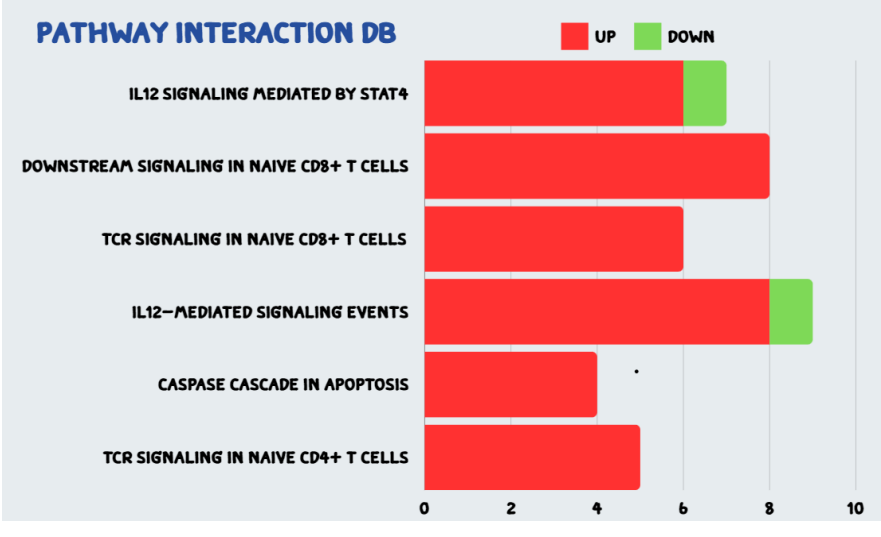



**Fig 3.** Pathway interaction database illustrating immune cell signaling pathways. The figure highlights upregulated activities in red and downregulated activities in green on the basis of adjusted *p*-values. The pathways that are represented include T-cell signaling and apoptosis, with the lengths of the bars indicating the significance levels of each pathway.

and suppression of specific pathways, particularly those involved in T-cell signaling and apoptosis.

The Pathways/BioPlanet chart in Fig 4 illustrates immune signaling pathways associated with the body's immune response, which are categorized into upregulated pathways (in red) and downregulated pathways (in green). All the signals presented a *p*-value of less than 0.05, indicating statistical significance. The analyzed pathways included "T-cell activation co-stimulatory signal," "leptin influence on the immune response," "interleukin-12/STAT4 pathway," "T-cell receptor and CD3 complex," "CTL-mediated immune response against target cells," and "PD-1 signaling". The length of each bar represents the significance level of either upregulation or downregulation in each pathway. The chart reveals that the "leptin influence on immune response" pathway is the most upregulated pathway, whereas the "interleukin-12/STAT4 pathway" is partially downregulated. This provides a comprehensive overview of variations in the immune response across different pathways, highlighting the importance of T-cell signaling and immune regulation.

### *Genes, biological processes, and signaling pathways*

To further investigate the different molecular functions at each resection margin, we conducted a functional analysis of differentially expressed genes with *p*- values < 0.05 via IPA. Our analysis revealed a total of 85 significantly 72 upregulated genes (positive fold change values) and 13 downregulated genes (negative fold change values) compared with cCR and non-cCR.

The analysis highlights key genes and pathways involved in immune regulation and inflammation. Central genes such as *TNF, IL6, IL1A,* and *IL1B* play pivotal roles in cytokine-mediated signaling, particularly in pathways such as cytokine storm signaling and Th1/Th2 activation. These processes drive immune responses, including the activation of cytotoxic T cells, monocyte adhesion, and the release of reactive oxygen species, which are essential for inflammation and defense mechanisms. Additionally, genes such as *RELA* (part of the NF-kappa B pathway) and *IFNG/IL17A* underscore the importance of adaptive immunity. These are depicted in Fig 5.

## DISCUSSION

This study demonstrated that patients who achieved a cCR had higher levels of IL-12 than those with a non-cCR. These findings suggest potential applications for these two groups of patients. The first group included patients who were not suitable for surgery due to their physical condition. For these patients, achieving a cCR could help reduce complications associated with the tumor mass, such as bleeding or difficulty swallowing.

The second group consisted of patients eligible for surgery, where IL-12 could be used as an adjuvant therapy. As a result, the present research could improve treatment options for both groups of patients. Further studies are necessary because this research was retrospective and limited by the availability of biopsy tissue, which must be collected before CCRT. This limitation may reduce the number of patients who meet the criteria. Future research could expand the patient sample size and use a prospective design, enabling the collection of both fresh and paraffin-embedded biopsy tissues. These samples could then be utilized for RNA extraction and further analysis via Geomax technology, allowing for more accurate calculations.

The current review highlights that IL-12 has been utilized in the treatment of several cancer types,[13] including melanoma,[4,14-22] based on certain studies, as well as in gynecologic cancer[10,12,22-29] and breast cancer,[8,14,25,30-40] as

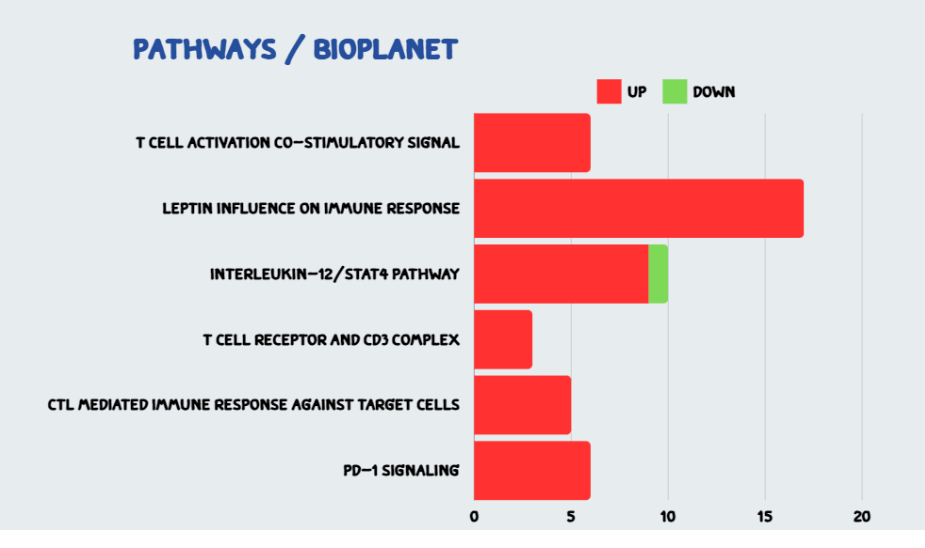



**Fig 4.** The Pathways/BioPlanet chart illustrates the immune signaling pathways involved in the body's immune response. Pathways that are upregulated are indicated in red, whereas those that are downregulated are shown in green. All signaling pathways had *p*-values less than 0.05. Notably, the "leptin influence on immune response" pathway was the most significantly upregulated pathway, whereas the "interleukin-12/STAT4 pathway" was partially downregulated.

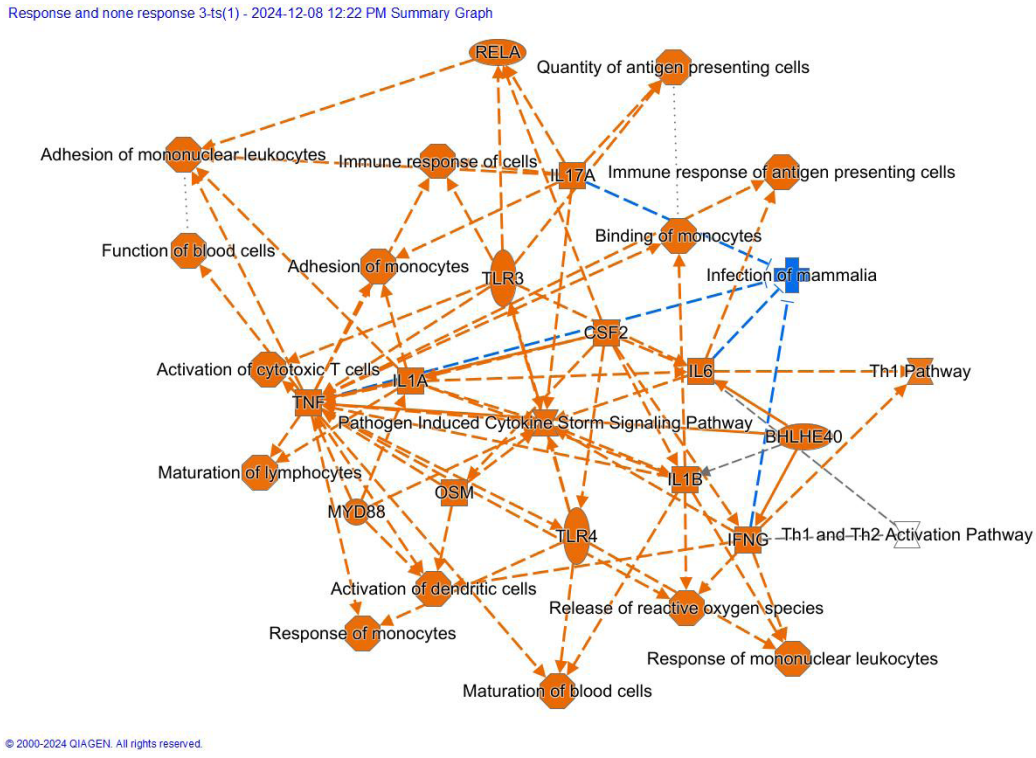



**Fig 5.** The diagram highlights critical immune pathways and key genes, such as *TNF, IL6, IL1A,* and *IL1B*, which regulate cytokine storms, Th1/Th2 pathways, and inflammation. These processes activate T cells, monocyte adhesion, and reactive oxygen species essential for immune responses.

indicated by other research. However, the use of IL-12 should be approached with caution due to potential side effects, such as flu-like symptoms, which some patients may find intolerable.[14] Various methods of administering IL-12 have been explored to mitigate these side effects.[36] Due to the severe side effects associated with systemic treatment via intravenous injections, alternative methods for administering IL-12, including intravenous injection, subcutaneous injection, intramuscular injection, and intratumoral injection, have been developed.[36] Future studies may focus on the intratumoral injection of IL-12, an adjuvant for cancer vaccines,[37] which could increase the effectiveness of treatment and potentially improve outcomes in patients with SCC of the esophagus. This recent discovery could prove beneficial in the future by improving treatment effectiveness for patients, particularly by increasing the rate of cCR.

## CONCLUSION

This study highlights significant differences in immune transcriptome profiles between patients with a cCR and those with a non-cCR to CRT in patients with SCC of the esophagus. The findings indicate that the upregulation of genes involved in IL-12 signaling, particularly those mediated by STAT4, may serve as potential biomarkers for predicting favorable treatment responses. Patients who achieved cCR presented a lower incidence of distant recurrence than non- cCR patients did, indicating that immune transcriptomic characteristics may influence patient prognosis. The results provide a strong foundation for future research that leverages transcriptomic data to develop personalized treatment strategies, potentially improving patient outcomes in patients with esophageal cancer.

### Data Availability Statement

The authors affirm that the data supporting the findings of this study are included within the article and its supplementary materials.

## ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to all individuals and organizations who contributed to the successful completion and publication of this research.

## DECLARATION

### Grants and Funding Information

No grants or funding applied.

## Conflict of Interest

All authors declare no personal or professional conflicts of interest relating to any aspect of this study.

## Registration Number of Clinical Trial

COA no. Si 203/2020

## Author Contributions

Conceptualization and methodology, N.S, V.T., A.M., P.T., T.T.; Specimen collection, A.M., J.S. and T.P.; Investigation, N.S., P.T., T.S., O.A., K.T.; Formal analysis, N.S. ,P.T. and T.S. ; Visualization and writing – original draft, N.S.; Writing – review and editing, N.S.,V.T. ; Funding acquisition, none ; Supervision, V.T. All authors have read and agreed to the final version of the manuscript.

## Use of Artificial Intelligence

No artificial intelligence tools or technologies were utilized in the writing, analysis, or development of this research.

## REFERENCES

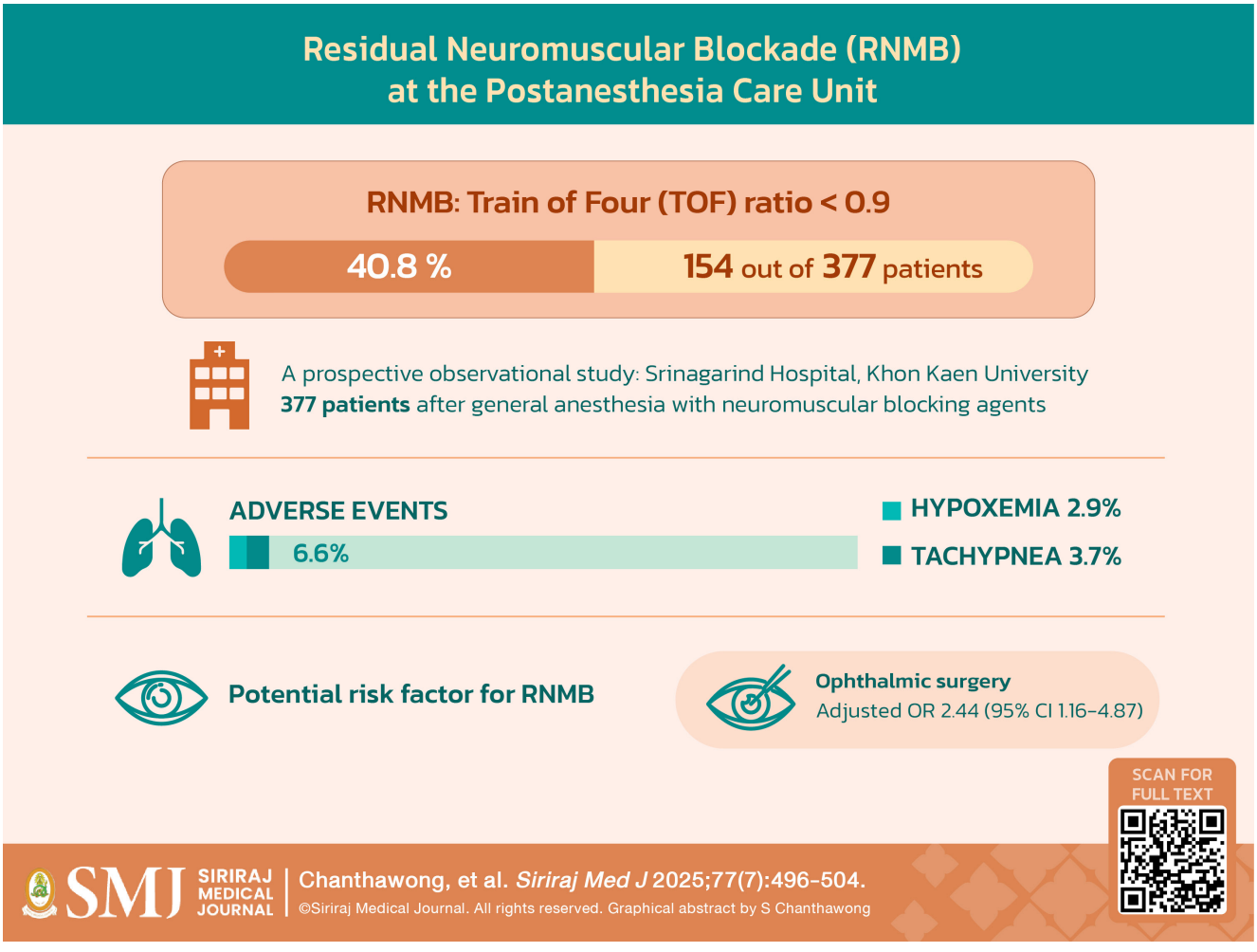

1. The Globocan Cancer Observatory - Globocan 2020. The Globocan Cancer Observatory - Globocan 2020.
2. Abul K. Abbas M, Andrew H. Lichtman, MD, PhD, Shiv Pillai, MBBS, PhD. Cellular and molecular immunology. Edition, editor, 2015.
3. Stamm LM, Satoskar AA, Ghosh SK, David JR, Satoskar AR. STAT-4 mediated IL-12 signaling pathway is critical for the development of protective immunity in cutaneous leishmaniasis. Eur J Immunol. 1999;29(8):2524-9.
4. Byrne-Hoffman CN, Deng W, McGrath O, Wang P, Rojanasakul Y, Klinke DJ. Interleukin-12 elicits a non-canonical response in B16 melanoma cells to enhance survival. Cell Commun Signal. 2020;18(1):78.
5. Jia Z, Ragoonanan D, Mahadeo KM, Gill J, Gorlick R, Shpal E, et al. IL12 immune therapy clinical trial review: Novel strategies for avoiding CRS-associated cytokines. Front Immunol. 2022; 13:952231.
6. Lasek W, Zagożdżon R, Jakobisiak M. Interleukin 12: still a promising candidate for tumor immunotherapy? Cancer Immunol Immunother. 2014;63(5):419-35.
7. Liu J, Cao S, Kim S, Chung EY, Homma Y, Guan X, et al. Interleukin-12: an update on its immunological activities, signaling and regulation of gene expression. Curr Immunol Rev. 2005;1(2):119-37.
8. Mirlekar B, Pylayeva-Gupta Y. IL-12 family cytokines in cancer and immunotherapy. Cancers. 2021;13(2):167.
9. Tugues S, Burkhard S, Ohs Ia, Vrohlings M, Nussbaum K, Vom Berg J, et al. New insights into IL-12-mediated tumor suppression. Cell Death Differ. 2015;22(2):237-46.
10. Su W, Ito T, Oyama T, Kitagawa T, Yamori T, Fujiwara H, et al. The direct effect of IL-12 on tumor cells: IL-12 acts directly on tumor cells to activate NF-κB and enhance IFN-γ-mediated STAT1 phosphorylation. Biochem Biophys Res Commun. 2001; 280(2):503-12.
11. Watford WT, Hissong BD, Bream JH, Kanno Y, Muul L, O'Shea JJ. Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4. Immunol Rev. 2004;202:139-56.
12. Wang J, Qin J, Jing S, Liu Q, Cheng Y, Wang Y, et al. Clinical complete response after chemoradiotherapy for carcinoma of thoracic esophagus: Is esophagectomy always necessary? A systematic review and meta-analysis. Thorac Cancer. 2018;9(12): 1638-47.
13. Lu X. Impact of IL-12 in Cancer. Curr Cancer Drug Targets. 2017;17(8):682-97.
14. Agliardi G, Liuzzi AR, Hotblack A, De Feo D, Núñez N, Stowe CL, et al. Intratumoral IL-12 delivery empowers CAR-T cell immunotherapy in a pre-clinical model of glioblastoma. Nat Commun. 2021;12(1):444.
15. Algazi AP, Twitty CG, Tsai KK, Le M, Pierce R, Browning E, et al. Phase II trial of IL-12 plasmid transfection and PD-1 blockade in immunologically quiescent melanoma. Clin Cancer Res. 2020; 26(12):2827-37.
16. Gao W, Pan J, Pan J. Antitumor activities of interleukin-12 in melanoma. Cancers. 2022;14(22):5592.
17. Gollob JA, Mier JW, Veenstra K, McDermott DF, Clancy D, Clancy M, et al. Phase I trial of twice-weekly intravenous interleukin 12 in patients with metastatic renal cell cancer or malignant melanoma: ability to maintain IFN-γ induction is associated with clinical response. Clin Cancer Res. 2000;6(5): 1678-92.
18. Greaney SK, Algazi AP, Tsai KK, Takamura KT, Chen L, Twitty CG, et al. Intratumoral plasmid IL12 electroporation therapy in patients with advanced melanoma induces systemic and intratumoral T-cell responses. Cancer Immunol Res. 2020;8(2): 246-54.
19. Komel T, Bosnjak M, Brezar SK, De Robertis M, Mastrodonato M, Scillitani G, et al. Gene electrotransfer of IL-2 and IL-12 plasmids effectively eradicated murine B16. F10 melanoma. Bioelectrochemistry. 2021;141:107843.
20. Moretti S, Chiarugi A, Semplici F, Salvi A, De Giorgi V, Fabbri P, et al. Serum imbalance of cytokines in melanoma patients. Melanoma Res. 2001;11(4):395-9.
21. Sun Y, Jurgovsky K, Möller P, Alijagic S, Dorbic T, Georgieva J, et al. Vaccination with IL-12 gene-modified autologous melanoma cells: preclinical results and a first clinical phase I study. Gene Ther. 1998;5(4):481-90.
22. Chen X, Han S, Wang S, Zhou X, Zhang M, Dong J, et al. Interactions of IL-12A and IL-12B polymorphisms on the risk of cervical cancer in Chinese women. Clin Cancer Res. 2009;15(1): 400-5.
23. Kumar N, Vyas A, Agnihotri SK, Chattopadhyay N, Sachdev M, editors. Small secretory proteins of immune cells can modulate gynecological cancers. Seminars in Cancer Biology; 2022: Elsevier.
24. Marana HRC, da Silva JS, de Andrade JM. NK cell activity in the presence of IL-12 is a prognostic assay to neoadjuvant chemotherapy in cervical cancer. Gynecol Oncol. 2000;78(3):318-23.
25. Ruffell B, Chang-Strachan D, Chan V, Rosenbusch A, Ho CM, Pryer N, et al. Macrophage IL-10 blocks CD8+ T cell-dependent responses to chemotherapy by suppressing IL-12 expression in intratumoral dendritic cells. Cancer Cell. 2014;26(5):623-37.
26. Whitworth JM, Alvarez RD. Evaluating the role of IL-12 based therapies in ovarian cancer: a review of the literature. Expert

Opin Biol Ther. 2011;11(6):751-62.
27. Xiong H-Y, Ma T-T, Wu B-T, Lin Y, Tu Z-G. IL-12 regulates B7-H1 expression in ovarian cancer-associated macrophages by effects on NF-κB signalling. Asian Pac J Cancer Prev. 2014; 15(14):5767-72.
28. Zeimet AG, Widschwendter M, Knabbe C, Fuchs D, Herold M, Müller-Holzner E, et al. Ascitic interleukin-12 is an independent prognostic factor in ovarian cancer. J Clin Oncol. 1998;16(5): 1861-8.
29. Zijlmans HJ, Fleuren GJ, Baelde HJ, Eilers PH, Kenter GG, Gorter A. Role of tumor-derived proinflammatory cytokines GM-CSF, TNF-α, and IL-12 in the migration and differentiation of antigen-presenting cells in cervical carcinoma. Cancer. 2007; 109(3):556-65.
30. Bekaii-Saab TS, Roda JM, Guenterberg KD, Ramaswamy B, Young DC, Ferketich AK, et al. A phase I trial of paclitaxel and trastuzumab in combination with interleukin-12 in patients with HER2/neu-expressing malignancies. Mol Cancer Ther. 2009; 8(11):2983-91.
31. Eliopoulos N, Francois M, Boivin M-N, Martineau D, Galipeau J. Neo-organoid of marrow mesenchymal stromal cells secreting interleukin-12 for breast cancer therapy. Cancer Res. 2008; 68(12):4810-8.
32. Gyorffy S, Palmer K, Podor TJ, Hitt M, Gauldie J. Combined treatment of a murine breast cancer model with type 5 adenovirus vectors expressing murine angiostatin and IL-12: a role for combined anti-angiogenesis and immunotherapy. J Immunol. 2001;166(10):6212-7.
33. Kaarvatn MH, Vrbanec J, Kulic A, Knezevic J, Petricevic B, Balen S, et al. Single nucleotide polymorphism in the interleukin 12B gene is associated with risk for breast cancer development. Scand J Immunol. 2012;76(3):329-35.
34. Kovacs E. The serum levels of IL-12 and IL-16 in cancer patients. Relation to the tumour stage and previous therapy. Biomed Pharmacother. 2001;55(2):111-6.
35. Mohamed Amin Z, Che Ani MA, Tan SW, Yeap SK, Alitheen NB, Syed Najmuddin SUF, et al. Evaluation of a recombinant Newcastle disease virus expressing human IL12 against human breast cancer. Sci Rep. 2019;9(1):13999.
36. Nguyen KG, Vrabel MR, Mantooth SM, Hopkins JJ, Wagner ES, Gabaldon TA, et al. Localized interleukin-12 for cancer immunotherapy. Front Immunol. 2020;11:575597.
37. Portielje JE, Gratama J, van Ojik HH, Stoter G, Kruit WH. IL-12: a promising adjuvant for cancer vaccination. Cancer Immunol Immunother. 2003;52(3):133-44.
38. Roszak A, Mostowska A, Sowińska A, Lianeri M, Jagodziński PP. Contribution of IL12A and IL12B polymorphisms to the risk of cervical cancer. Pathol Oncol Res. 2012;18(4):997-1002.
39. Sabel MS, Skitzki J, Stoolman L, Egilmez NK, Mathiowitz E, Bailey N, et al. Intratumoral IL-12 and TNF-α–loaded microspheres lead to regression of breast cancer and systemic antitumor immunity. Ann Surg Oncol. 2004;11(2):147-56.
40. Telli ML, Nagata H, Wapnir I, Acharya CR, Zablotsky K, Fox BA, et al. Intratumoral plasmid IL12 expands CD8+ T cells and induces a CXCR3 gene signature in triple-negative breast tumors that sensitizes patients to anti–PD-1 therapy. Clin Cancer Res. 2021;27(9):2481-93.

# Incidence of Postoperative Residual Neuromuscular Blockade at the Postanesthesia Care Unit Following General Anesthesia

Sarinya Chanthawong, M.D., Thidamas Apithambundit, M.D., Netinai Chaimala, B.N.S., Sangmukda Saras, B.N.S., Thanat Yotepanya, M.D.*, Saranyoo Nonphiaraj, M.D.

*Department of Anesthesiology, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand.*




*Corresponding author: Thanat Yotepanya
E-mail: thanyo@kku.ac.th
Received 27 December 2024 Revised 27 February 2025 Accepted 27 February 2025
ORCID ID:http://orcid.org/0009-0006-2973-4368
https://doi.org/10.33192/smj.v77i7.272875

## ABSTRACT

**Objective:** This study aimed to investigate residual neuromuscular blockade (RNMB), respiratory adverse events, and identify risk factors for RNMB.

**Material and Methods:** A prospective observational study enrolled 377 elective adult patients with American Society of Anesthesiologists (ASA) classifications 1–3 who underwent GA with NMBAs. At the conclusion of surgery, endotracheal tubes were removed once clinical extubation criteria were met. The Train-of-Four (TOF) ratio was immediately measured upon the patients' arrival in the postanesthesia care unit (PACU), with RNMB defined as a TOF ratio < 0.9. Respiratory adverse events were documented.

**Results:** The incidence of RNMB was 40.8% upon PACU arrival, with a median TOF ratio of 0.92 (interquartile range [IQR], 0.85–0.98). Cisatracurium was the primary NMBA used (98.1%). The incidence of respiratory adverse events was 6.6%, including hypoxemia (2.9%) and tachypnea (3.7%). No cases of reintubation or unplanned ICU admission occurred. The ophthalmologic surgery emerged as the only significant risk factor for RNMB, adjusted OR 2.44 (95% CI, 1.16-4.38, p = 0.02).

**Conclusion:** The incidence of RNMB after GA in the PACU was common, though no serious adverse events were observed. The type of surgery was identified as the sole significant risk factor for RNMB.


**Keywords:** Residual neuromuscular blockade; NMBAs; muscle relaxant; postanesthesia care unit; train-of-four (Siriraj Med J 2025; 77: 496-504)

## INTRODUCTION

Neuromuscular blocking agents (NMBAs) are essential components of most anesthetic techniques. They are used to facilitate tracheal intubation, reduce airway injury, and maintain muscle relaxation, thereby promoting the success of surgical procedures, particularly in thoracic and abdominal surgeries.[1,2] However, several studies have reported that the effects of NMBAs often persist in the post-anesthesia care unit (PACU), even after the administration of acetylcholinesterase inhibitors,[3] indicating that these patients may experience residual neuromuscular blockade (RNMB).

RNMB is linked to postoperative pulmonary complications such as hypoxemia, hypercapnia, upper airway obstruction, respiratory muscle weakness, dyspnea, the need for airway equipment, reintubation, pneumonia, pharyngeal muscle dysfunction, aspiration, and atelectasis.[4,5] These complications increase morbidity and mortality within the first 24 hours postoperatively[6] and may delay discharge from the PACU.[7]

According to the 2020 guidelines on muscle relaxants and reversal in anesthesia published by the French Society of Anaesthesia & Intensive Care Medicine (SFAR), quantitative monitoring of the adductor pollicis muscle of the ulnar nerve is likely recommended for diagnosing RNMB.[8] Achieving a Train-of-Four (TOF) ratio of ≥ 0.9 can eliminate the possibility of RNMB. The guidelines also recommend administering an acetylcholinesterase inhibitor once a measured TOF ratio at the adductor pollicis muscle is ≥ 0.2.

Several studies[7,9–13] emphasize the importance of appropriate administration of neuromuscular blocking agents (NMBAs) and their reversal agents. Specifically, the use of neuromuscular monitoring is crucial to prevent the incidence and adverse effects of RNMB. In current practice at our center, clinical criteria are used at the end of surgery before extubation. If these clinical extubation criteria are not met or RNMB is suspected, a TOF ratio is measured. An additional reversal agent is considered if the TOF ratio is consistent with RNMB.

Due to the lack of data on the incidence of RNMB and the appropriate administration of NMBAs and their reversal agents in our tertiary health care center, we aimed to investigate the incidence of RNMB upon PACU arrival and to identify associated risk factors. The results of this study could contribute to a new database and be analyzed to improve our institution's patient management guidelines.

## MATERIAL AND METHODS

This was a prospective observational study approved by the Ethics Committee in Human Research, Khon Kaen University (HE651283). Data collection occurred from August 2022 to July 2023. Patients aged over 18 years with American Society of Anesthesiologists (ASA) classification I-III who underwent elective surgery with general anesthesia using NMBAs were enrolled. Exclusion criteria included patients with pre-existing neuromuscular diseases such as myasthenia gravis, Guillain-Barre syndrome which could interfere with TOF responses[14,15], undergoing

airway surgery, sensitive population such as pregnant patients, and patients with preoperative respiratory conditions such as hypoxemia ($SpO_2$ < 90%), tachypnea (respiratory rate > 20 breaths/min), use of oropharyngeal or nasopharyngeal airways, retention of an endotracheal tube, use of accessory respiratory muscles, or pneumonia. Additionally, patients who declined participation were excluded.

Patients who met the inclusion criteria were selected and informed about the details of the study. On the day of surgery, general anesthesia with NMBAs was conducted according to routine practice determined by the attending anesthesiologist. At the end of the procedure, reversal agents (either neostigmine at 0.05 mg/kg with atropine at 0.02 mg/kg or sugammadex at 2 mg/kg) were administered. Tracheal extubation was performed once the patient met all clinical extubation criteria which included hemodynamic stability, spontaneous breathing with tidal volume > 5 mL/kg, $SpO_2$ > 94%, $EtCO_2$ < 50 mmHg, intact gag reflex, sustained head lift for 5 seconds, hand grip, and following commands. During transfer to the PACU following extubation, patients received supplemental oxygen at 6–8 L/min via an oxygen mask with a reservoir bag.

Upon arrival at the PACU, the acceleromyography TOF-Watch® SX (Organon) was applied by the PACU nurse anesthetist who was not involved in this study, within the first 5 minutes to measure the TOF ratio. TOF stimulations at 30 mA were applied to the ulnar nerve twice, with a 15-second interval, and TOF responses were measured at the adductor pollicis muscle. The final TOF ratio was calculated as the average of the two values obtained. If the discrepancy between the two readings was ≥ 0.2, an additional TOF stimulation was performed, and the average of the two closest TOF ratios was used as the final TOF ratio. All study procedures were conducted without interfering with standard PACU care.

Patients' demographic data, including sex, age, weight, height, ASA physical status, and underlying diseases, were collected. Perioperative data such as type of surgery, anesthesia duration, type, dose, and interval of NMBAs, type and dose of reversal agents, and body temperature in the operating room and PACU were recorded. RNMB was diagnosed if the average TOF ratio was < 0.9 and defined as severe RNMB if the average TOF ratio was < 0.7. Respiratory adverse events, including hypoxemia, tachypnea, the need for airway support, use of accessory respiratory muscles, reintubation, and unplanned ICU admission, were documented within 1 hour postoperatively.

### *Sample size calculation*

The sample size was calculated using a formula for estimating the proportion in an infinite population, based on the study by Ismail Aytac et al.[14], which reported an RNMB incidence of 43%. A 95% confidence level ($\alpha = 0.05$) and a 5% margin of error were used. The final calculated sample size was 377 patients.

### *Statistical analysis*

Data analysis was conducted using STATA version 18. Categorical data were analyzed using Fisher's exact test or the Chi-squared test and presented as numbers and percentages. Continuous data which normally distributed were presented as mean with standard deviation (SD) and analyzed using Student's t-test while non-normally distributed data were presented as median with interquartile range (IQR) and analyzed using Wilcoxon rank-sum test. A p-value of < 0.05 was considered statistically significant. Univariate and multivariate logistic regression analyses were performed to identify risk factors for RNMB.

## RESULTS

During the study period, 377 patients were enrolled, the majority of whom were female (66%). Participants' demographic data are shown in Table 1. The most common ASA classification was II (49.1%), and the most frequently performed procedure was gynecologic surgery (24.1%). The average patient age was 52 years. The mean anesthesia duration was 154.1 ± 64.7 minutes. Cisatracurium was the most frequently used NMBA (98.1%). There were 3 patients received a combination of NMBAs: 2 patients received rocuronium and cisatracurium, and 1 patient received cisatracurium and atracurium. The average time between the last NMBA dose and administration of a reversal agent was 52.4 ± 21.5 minutes. The median TOF ratio was 0.92 (IQR: 0.85, 0.98). All patients received a reversal agent prior to extubation, with neostigmine being the most commonly used (99.7%).

The incidence of RNMB upon arrival at the PACU was 40.8% (154 out of 377 patients), as illustrated in Fig 1. Two of these patients received a combination of NBMAs: one patient received rocuronium and cisatracurium, and another patient received cisatracurium and atracurium. The incidence rates of RNMB were 40.8% for cisatracurium, 40% for atracurium, and 60% for rocuronium. Severe RNMB was observed in 5.6% of patients (21 out of 377), with 20 patients having received cisatracurium and 1 patient having received a combination of rocuronium and cisatracurium. (Fig 2)

**TABLE 1.** Clinical features of patients who underwent elective surgery with general anesthesia using neuromuscular blocking agents (n=377).

| Characteristics | Study population (n=377) |
|---|---|
| Female* | 249 (66%) |
| Age (years)** | 52.3 ± 15.3 |
| Weight (kg)** | 62.6 ± 12.4 |
| Height (cm)** | 160.0 ± 10.0 |
| BMI (kg/m$^2$)** | 24.3 ± 4.4 |
| ASA classification* | |
| I | 165 (43.7%) |
| II | 185 (49.1%) |
| III | 27 (7.2%) |
| Surgery type* | |
| Gynecologic | 91 (24.1%) |
| Otolaryngologic | 74 (19.6%) |
| Ophthalmologic | 72 (19.1%) |
| General | 72 (19.1%) |
| Orthopedic | 51 (13.5%) |
| Plastic | 13 (3.5%) |
| Neurologic | 3 (0.8%) |
| Maxillofacial | 1 (0.3%) |
| Anesthetic time (min)** | 154.1 ± 64.7 |
| Temperature in OR (˚C)*** | 36.0 (35.7, 36.3) |
| Temperature in PACU (˚C)*** | 36.4 (36.2, 36.6) |
| Neuromuscular blocker used* | |
| Cisatracurium | 370 (98.1%) |
| Atracurium | 5 (1.3%) |
| Rocuronium | 5 (1.3%) |
| Reversal agent used* | |
| Neostigmine | 376 (99.7%) |
| Sugammadex | 1 (0.3%) |
| Dosage of neuromuscular blocker (mg/kg)** | |
| Cisatracurium | 0.20 ± 0.08 |
| Atracurium | 0.70 ± 0.43 |
| Rocuronium | 0.91 ± 0.14 |
| Dosage of reversal agent (mg/kg)** | |
| Neostigmine | 0.04 ± 0.01 |
| Interval from last dose NMB administration (mins)** | 52.4 ± 21.5 |
| TOF ratio*** | 0.92 (0.85, 0.98) |

* = number (%), **= mean ± SD, ***= median (IQR)

**Abbreviations:** BMI = body mass index; ASA = American society of Anesthesiologists; OR = operating rooms, PACU = post anesthesia care unit; TOF = train of four

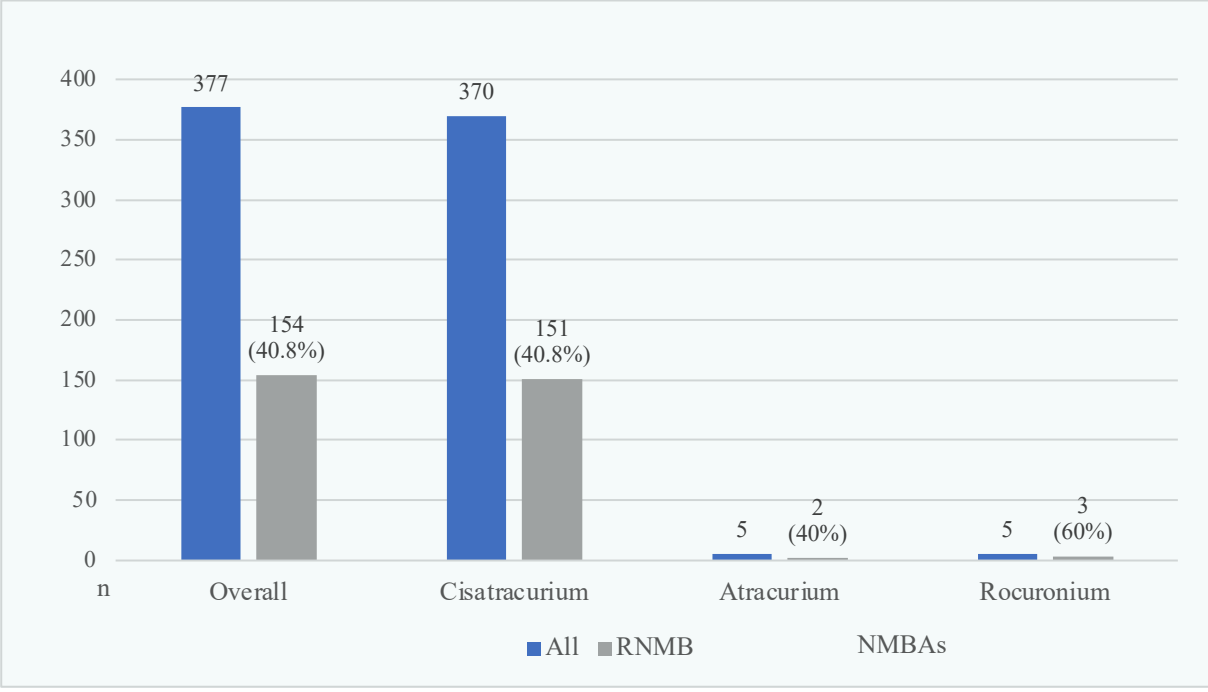



**Fig 1.** Incidence of residual neuromuscular blockade in patients who underwent elective surgery with general anesthesia using neuromuscular blocking agents.

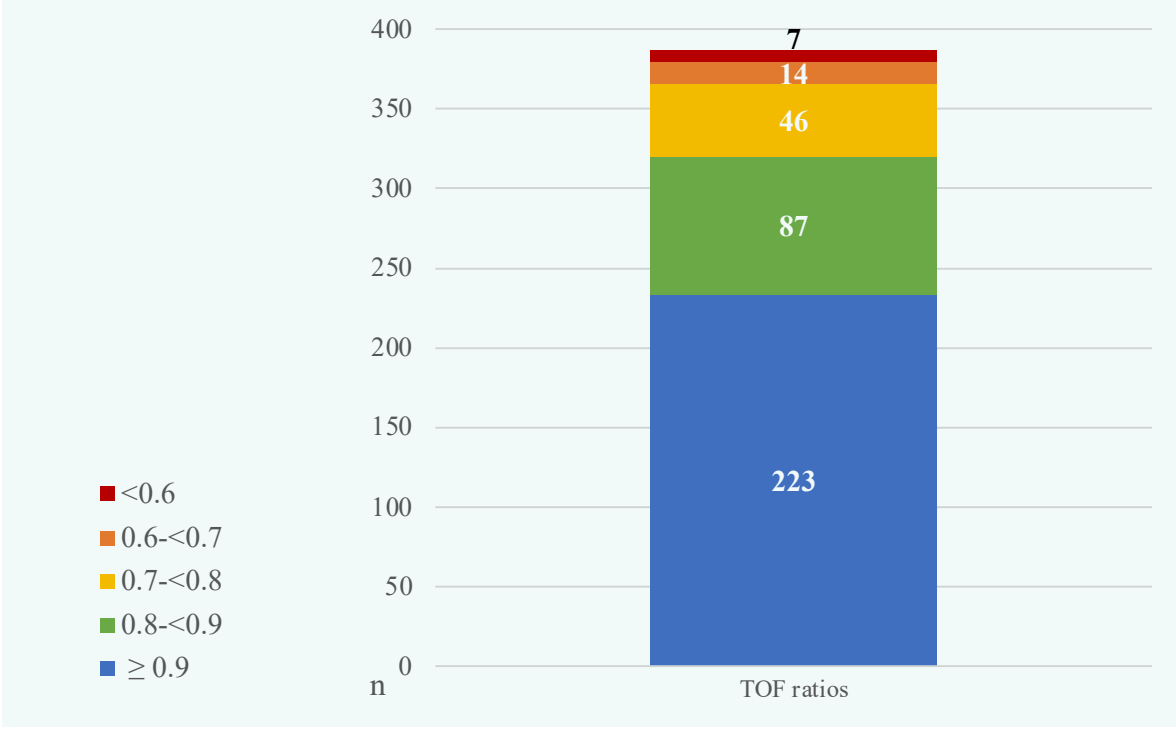



**Fig 2.** Number of patients in each TOF ratio range.

The incidence of adverse events within the first hour postoperatively was 6.6% (25 patients). Adverse respiratory events included hypoxemia in 11 patients (2.9%) and tachypnea in 14 patients (3.7%). All adverse events occurred in patients who had received cisatracurium. Hypoxemia was managed with low-flow oxygen delivered via nasal cannula or mask with a reservoir bag. No cases of reintubation or unplanned ICU admission were reported.

Univariable and multivariable analyses were conducted to identify potential risk factors for residual neuromuscular blockade, including sex, age, BMI > 30 kg/m², ASA classification I-III, chronic kidney disease (CKD) stage 4-5, type of surgery, time between the last

NMBA dose and reversal agent administration, anesthesia duration, and hypothermia upon arrival at the PACU. The analysis revealed that ophthalmologic surgery was the only significant risk factor for RNMB, adjusted OR of 2.44 (95% CI: 1.16 to 4.87, p=0.02).

## DISCUSSION

### *Incidence of residual neuromuscular blockade*

This study found an incidence of RNMB of 40.8%, despite all patients receiving a reversal agent. This is comparable to an observational study by Aytac et al.,[16] which included 415 patients undergoing anesthesia with intermediate-acting NMBAs and reported an overall RNMB incidence of 43%. In that study, TOF ratios < 0.9 were observed in 42% of patients receiving vecuronium, 52% of those receiving rocuronium, and 30% of those receiving atracurium. Additionally, the incidence of RNMB was 46.3% among patients who received neostigmine at a mean dose of 20 ± 10 mcg/kg. We hypothesize that the lower incidence of RNMB in our study may be attributed to the higher mean dose of neostigmine used (40 ± 10 mcg/kg). Furthermore, there was a difference in the type of NMBAs used, with 98.1% of patients in our study receiving cisatracurium.

A meta-analysis by Carvalho et al.[17] in 2020, which included 53 studies and 12,664 patients, reported the incidence of RNMB based on TOF ratios measured at or post-extubation of <0.7, <0.9, or <1.0. Specifically

**TABLE 2.** Correlation between patient/anesthesia/surgery-related factors and residual neuromuscular blockade.

| Variables | Crude OR 95% CI | *p*-value | Adjusted OR 95% CI | *p*-value |
|---|---|---|---|---|
| Male | 0.73 (0.47, 1.14) | 0.16 | 0.66 (0.38, 1.14) | 0.13 |
| Advanced age | 1.26 (0.77, 2.06) | 0.35 | 1.16 (0.68, 1.20) | 0.58 |
| BMI ≥ 30 kg/m² | 0.75 (0.33, 1.73) | 0.50 | 0.65 (0.25, 1.68) | 0.37 |
| Surgery type | | | | |
| Ophthalmologic | 2.36 (1.20, 4.64) | 0.01* | 2.44 (1.16, 4.87) | 0.02* |
| Gynecologic | 1.79 (0.94, 3.40) | 0.22 | 1.42 (0.68, 2.96) | 0.35 |
| Otolaryngologic | 1.52 (0.78, 2.98) | 0.07 | 1.32 (0.65, 2.67) | 0.44 |
| Neurologic | 1.00 (0.09, 11.59) | > 0.99 | 1.15 (0.10, 13.74) | 0.91 |
| Orthopedic | 0.68 (0.31, 1.52) | 0.35 | 0.62 (0.27, 1.40) | 0.25 |
| Plastic | 0.36 (0.07, 1.77) | 0.47 | 0.36 (0.07, 1.87) | 0.22 |
| CKD stage 4-5 | 0.58 (0.09, 3.72) | 0.65 | 0.87(0.25, 2.95) | 0.82 |
| Hypothermia (BT < 36˚C) | 1.48 (0.92, 2.40) | 0.11 | 1.46 (0.86, 2.46) | 0.16 |
| Type of muscle relaxant | | | | |
| Cisatracurium | 0.92 (0.20, 4.17) | > 0.99 | 0.75 (0.03, 19.67) | 0.71 |
| Atracurium | 0.96 (0.16, 5.84) | > 0.99 | 0.88 (1.22, 6.35) | 0.90 |
| Rocuronium | 2.20 (0.36, 13.3) | > 0.40 | 2.67 (0.25, 27.65) | 0.41 |
| Reversal dose (mcg/kg) | 0.15 (<0.01, >99) | 0.62 | 0.08 (0.00, 532) | 0.75 |
| Time between the last dose of NMBAs and reversal agents < 30 min | 0.76 (0.31, 1.84) | 0.54 | 0.91 (0.35, 2.34) | 0.84 |
| Anesthetic time > 240 min | 0.47 (0.20, 1.09) | 0.16 | 0.48 (0.19, 1.25) | 0.22 |

* p-value < 0.05 is considered statistically significant

**Abbreviations:** BMI = body mass index; CKD = chronic kidney disease; BT = body temperature; NMBAs = neuromuscular blocking agents

focusing on a TOF ratio of <0.9, the study found that the RNMB incidence in adult patients undergoing elective surgery under general anesthesia with intermediate-acting NMBAs was 33.1% in the no-neuromuscular-monitoring (NMM) group, 30.6% in the qualitative NMM group (peripheral nerve stimulation), and 11.5% in the quantitative NMM group (TOF ratio monitoring). The study concluded that patients monitored with TOF had a significantly lower incidence of RNMB compared to those monitored qualitatively or not monitored at all. Our findings are comparable to those of the no-NMM group.

An observational study conducted in Thailand by Khamtuikrua et al.[18] in 2017, which included 209 patients in a similar setting to our study, reported an incidence of RNMB upon PACU arrival of 53.1%, higher than the incidence observed in our study. Despite the similarity in patient populations, a possible explanation for this difference is that 3.3% of patients in their study did not receive a reversal agent. In contrast, all patients in our study received reversal agents, with 99.7% receiving neostigmine and 0.3% receiving sugammadex. Since our center does not routinely employ TOF monitoring, we administer a reversal agent to most patients to reduce the incidence of RNMB.

### *Adverse respiratory events*

In this study, we found that 45.5% of the 11 patients with hypoxemia had RNMB. This result contrasts with the findings of Aytac et al.[16], where 82.4% of patients with hypoxemia had RNMB. Respiratory complications were more frequent among patients with a TOF ratio < 0.7, with 42.6% requiring airway support. Additionally, Murphy et al.19 reported that patients experiencing critical respiratory events after general anesthesia had a higher incidence of RNMB.

Interestingly, no major adverse events, such as reintubation or unplanned ICU admission, were observed. The adverse events identified were hypoxemia and tachypnea, which were successfully managed with conventional low-flow oxygen supplementation. We recorded only the incidence of these adverse respiratory events and did not explore their underlying causes, which may have been multifactorial.

Therefore, despite the incidence of RNMB being 40.8%, our current practice—which includes the use of clinical extubation criteria, administration of reversal agents to most patients, and selective use of neuromuscular monitoring (e.g., in patients with neuromuscular diseases or delayed awakening)—appears to be relatively safe for our general patient population, as no major adverse events were recorded.

### *Risk factors for residual neuromuscular blockade*

We found that only the type of surgery was significantly associated with RNMB, with ophthalmologic surgery showing the highest odds ratio. A possible explanation for this finding is that 40% (29 out of 72) of patients undergoing ophthalmologic surgery were of advanced age, a group known to be at increased risk for RNMB.[20] Additionally, we observed that 65% of older adults who underwent ophthalmologic surgery experienced RNMB. This is consistent with a recent study by Murphy et al.[21], which suggested that RNMB may be more common in older adults (57%) compared to younger patients (30%). Therefore, it is recommended that a combination of TOF monitoring and clinical assessments be used for these patients.[22]

Aytac et al.[16] identified female sex, ASA physical status class III, and short anesthesia duration as risk factors for RNMB, which differed from our findings. Additionally, Carvalho et al.[17] reported that the use of sugammadex as a reversal agent was associated with a lower rate of RNMB compared to neostigmine. However, we were unable to investigate differences between neostigmine and sugammadex in our study, as it was only administered to a single patient.

Khamtuikrua et al.[18] reported that an increase in age by 10 years, a time interval of less than 30 minutes between the last dose of NMBA and administration of a reversal agent, and hypothermia (body temperature < 36˚C) were significant risk factors for RNMB.

### *Limitations*

Since our primary outcome was to assess the incidence of RNMB, we calculated the sample size based on the proportion reported in the previous study by Aytac et al.[16] However, our sample size may have been insufficient to identify all possible risk factors for RNMB. Additionally, in order to identify RNMB, we did not use clinical assessments such as hand grip strength combined with the TOF ratio, which would have been helpful in determining early neuromuscular impairment.[23] To properly identify risk factors, further studies with an appropriately powered sample size, a multi-center study, or a combination of clinical assessment and TOF for RNMB diagnosis may be required.

## CONCLUSION

The incidence of RNMB in the PACU following general anesthesia was high, even with the use of reversal agents. However, no serious respiratory adverse events were observed. Among the potential risk factors, ophthalmic surgery was the only factor significantly associated with RNMB.

***Clinical implications***

1. What is already known on this topic

Residual neuromuscular blockade, defined as a TOF ratio < 0.9, can lead to postoperative pulmonary complications and increase both morbidity and mortality. Although neuromuscular monitoring is recommended to exclude residual neuromuscular blockade, it is not routinely used in our center.

2. What this study adds

The incidence of RNMB upon arrival at the PACU was 40.8%, despite the administration of a reversal agent. No major respiratory adverse events were observed. Ophthalmic surgery was identified as a risk factor for RNMB. Therefore, routine TOF monitoring is recommended for patients at risk of RNMB, specifically those undergoing ophthalmic surgery, to enhance safety.

### Data Availability Statement

The data of this study are available upon reasonable request.

## ACKNOWLEDGMENTS

We would like to thank Dr. Dylan Southard for editing this manuscript via the KKU Publication Clinic (Thailand).

## DECLARATION

### Grants and Funding Information

This work is not funded by any external resources.

### Conflict of Interest

The authors declare no conflict of interest.

### Registration Number of Clinical Trial

This was a prospective observational study approved by the Ethics Committee in Human Research, Khon Kaen University (HE651283)

### Author Contributions

Conceptualization and methodology, S.C., T.A., and T.Y.; data collection, T.A., N.C., and S.S.; formal analysis, S.C., T.A., and T.Y.; writing – original draft, S.C., T.A.; writing – review and editing, S.C., T.Y., and S.N. All authors have read and agreed to the published version of the manuscript.

### Use of Artificial Intelligence

None

## REFERENCES

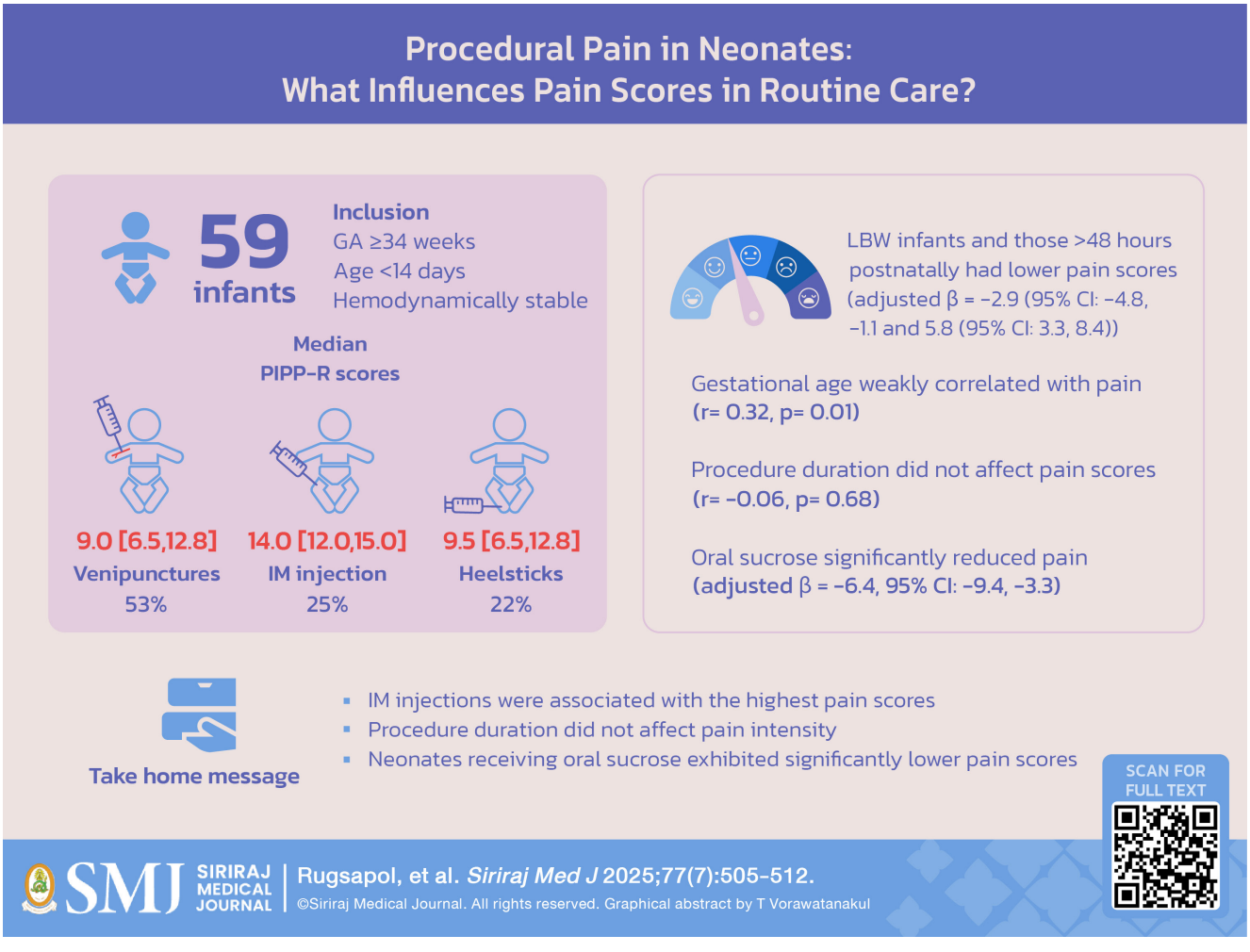
1. Baillard C, Clec'h C, Catineau J, Salhi F, Gehan G, Cupa M, et al. Postoperative residual neuromuscular block: a survey of management. Br J Anaesth. 2005;95(5):622–6.
2. Yu B, Ouyang B, Ge S, Luo Y, Li J, Ni D, et al. Incidence of postoperative residual neuromuscular blockade after general anesthesia: a prospective, multicenter, anesthetist-blind, observational study. Curr Med Res Opin. 2016;32(1):1–9.
3. Plaud B, Debaene B, Donati F, Marty J. Residual paralysis after emergence from anesthesia. Anesthesiology. 2010;112(4):1013–22.
4. Murphy GS, Szokol JW, Marymont JH, Greenberg SB, Avram MJ, Vender JS. Residual neuromuscular blockade and critical respiratory events in the postanesthesia care unit. Anesth Analg. 2008;107(1):130–7.
5. Grosse-Sundrup M, Henneman JP, Sandberg WS, Bateman BT, Uribe JV, Nguyen NT, et al. Intermediate acting non-depolarizing neuromuscular blocking agents and risk of postoperative respiratory complications: prospective propensity score matched cohort study. BMJ. 2012;345:e6329.
6. Arbous MS, Meursing AEE, van Kleef JW, de Lange JJ, Spoormans HHAJM, Touw P, et al. Impact of anesthesia management characteristics on severe morbidity and mortality. Anesthesiology. 2005;102(2):257–68; quiz 491–2.
7. Butterly A, Bittner EA, George E, Sandberg WS, Eikermann M, Schmidt U. Postoperative residual curarization from intermediate-acting neuromuscular blocking agents delays recovery room discharge. Br J Anaesth. 2010;105(3):304–9.
8. Plaud B, Baillard C, Bourgain JL, Bouroche G, Desplanque L, Devys JM, et al. Guidelines on muscle relaxants and reversal in anaesthesia. Anaesth Crit Care Pain Med. 2020;39(1):125–42.
9. Murphy GS, Brull SJ. Residual neuromuscular block: lessons unlearned. Part I: definitions, incidence, and adverse physiologic effects of residual neuromuscular block. Anesth Analg. 2010;111(1):120–8.
10. Brull SJ, Murphy GS. Residual neuromuscular block: lessons unlearned. Part II: methods to reduce the risk of residual weakness. Anesth Analg. 2010;111(1):129–40.
11. Murphy GS. Residual neuromuscular blockade: incidence, assessment, and relevance in the postoperative period. Minerva Anestesiol. 2006;72(3):97–109.
12. Berg H, Roed J, Viby-Mogensen J, Mortensen CR, Engbaek J, Skovgaard LT, et al. Residual neuromuscular block is a risk factor for postoperative pulmonary complications. A prospective, randomised, and blinded study of postoperative pulmonary complications after atracurium, vecuronium and pancuronium. Acta Anaesthesiol Scand. 1997;41(9):1095–103.
13. Esteves S, Martins M, Barros F, Barros F, Canas M, Vitor P, et al. Incidence of postoperative residual neuromuscular blockade in the postanaesthesia care unit: an observational multicentre study in Portugal. Eur J Anaesthesiol. 2013;30(5):243–9.
14. Moningi S, Durga P, Mantha S, Ramachandra G. Train of Four Responses in Paretic Limbs. Journal of Neurosurgical Anesthesiology [Internet]. 2009;21(4). Available from: https://journals.lww.com/jnsa/fulltext/2009/10000/train_of_four_responses_in_paretic_limbs.9.aspx
15. Thongsing A, Likasitwattanakula S, Netsuwan T, Sanmaneechai O. Pediatric Neuromuscular Diseases Prevalence in Siriraj Hospital, Thailand's Largest Tertiary Referral Hospital. Siriraj Med J. 2020;72(2):125–31.

16. Aytac I, Postaci A, Aytac B, Sacan O, Alay GH, Celik B, et al. Survey of postoperative residual curarization, acute respiratory events and approach of anesthesiologists. Braz J Anesthesiol. 2016;66(1):55–62.
17. Carvalho H, Verdonck M, Cools W, Geerts L, Forget P, Poelaert J. Forty years of neuromuscular monitoring and postoperative residual curarisation: a meta-analysis and evaluation of confidence in network meta-analysis. Br J Anaesth. 2020;125(4):466–82.
18. Khamtuikrua C, Suksompong S, Rhoopanwong S, Sangsab P, Chaikittisilpa N, Bormann B von. Risk Factors for Residual Neuromuscular Blockade after General Anesthesia. J Med Assoc Thai. 2017;100(Suppl 7):S75-S84.
19. Murphy GS, Szokol JW, Marymont JH, Greenberg SB, Avram MJ, Vender JS, et al. Intraoperative acceleromyographic monitoring reduces the risk of residual neuromuscular blockade and adverse respiratory events in the postanesthesia care unit. Anesthesiology. 2008;109(3):389–98.
20. Lee LA, Athanassoglou V, Pandit JJ. Neuromuscular blockade in the elderly patient. J Pain Res. 2016;9:437–44.
21. Murphy GS, Szokol JW, Avram MJ, Greenberg SB, Shear TD, Vender JS, et al. Residual Neuromuscular Block in the Elderly: Incidence and Clinical Implications. Anesthesiology. 2015; 123(6):1322–36.
22. Sagir O, Yucesoy Noyan F, Koroglu A, Cicek M, Ilksen Toprak H. Comparison between the Effects of Rocuronium, Vecuronium, and Cisatracurium Using Train-of-Four and Clinical Tests in Elderly Patients. Anesth Pain Med. 2013;2(4):142–8.
23. Tantibhaedhyangkul P, Kuptniratsaikul V, Tosayanonda O. Grip and Quadriceps Strength: Normative Values in the Thai Population. Siriraj Med J. 2020;53(4):224–30.

# Factors Associated with Pain Scores in Late Preterm and Term Infants Undergoing Routine Procedures

**Supranee Rugsapol, B.NS.[1], Ittichote Amornjiraporn, M.D.[2], Peerawit Thanasarnpaiboon, M.D.[2], Ratchada Kitsommart, M.D.[2,*]**

*[1]Division of Nursing, Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*




*Corresponding author: Ratchada Kitsommart
E-mail: rkitsommart@hotmail.com, ratchada.kit@mahidol.ac.th
Received 2 April 2025 Revised 30 April 2025 Accepted 30 April 2025
ORCID ID:http://orcid.org/0000-0002-7592-9899
https://doi.org/10.33192/smj.v77i7.274583

## ABSTRACT

**Objective:** To identify factors associated with Premature Infant Pain Profile-Revised (PIPP-R) scores in late preterm and term infants undergoing routine skin-puncture procedures.

**Materials and Methods:** A secondary analysis of a prospective cohort study was conducted in hemodynamically stable late preterm and term infants (gestational age [GA] 34–41 weeks) undergoing routine procedures. PIPP-R scores were evaluated through video recordings by a blinded assessor.

**Results:** Among 59 infants, 31 (52.5%) underwent venipuncture, 15 (25.4%) received intramuscular (IM) injections, and 13 (22%) underwent heelsticks. The median PIPP-R score was 11.0 [7.0, 14.0]. IM injections and heelsticks resulted in significantly higher pain scores compared to venipunctures (regression coefficient: 4.6, 95% CI: 2.3–6.8, and 2.8, 95% CI: 0.6–5.0, respectively). No correlation was observed between procedure duration and pain scores ($r = -0.06$, $p = 0.68$), but GA showed a weak positive correlation ($r = 0.32$, $p = 0.01$). After adjustment, low birthweight (<2500 g) was independently associated with lower PIPP-R scores (-2.9, 95% CI: -4.8, -1.1), while postnatal age <48 hours was linked to higher scores (5.8, 95% CI: 3.3–8.4). Oral sucrose solution significantly reduced pain scores (-6.4, 95% CI: -9.4, -3.3).

**Conclusion:** IM injections were associated with the highest pain scores, while procedure duration did not affect pain intensity. LBW infants and those >48 hours old exhibited lower pain scores, suggesting developmental factors. The demonstrated efficacy of oral sucrose underscores its importance in neonatal pain management, emphasizing the need for routine implementation of analgesic strategies.


**Keywords:** Heelstick; intramuscular; late-preterm; PIPP-R scores; venipuncture (Siriraj Med J 2025; 77: 505-512)

## INTRODUCTION

Newborn infants are routinely subjected to painful procedures, including essential newborn screenings and vaccination injections. Higher-risk infants may require additional procedures. Research suggests that early pain exposure can influence both short-term and long-term nervous system development. Studies indicate a correlation between pain and abnormal brain development, as evidenced by dysmaturation in the brains of newborns.[1] Autopsy findings in infants with abnormal brain tissues reveal impaired preoligodendrocyte development, which is critical for myelin production and may result from brain injury repair processes during infancy.[2] Studies tracking early brain changes using MRI have shown an association between white matter abnormalities, altered diffusivity, and abnormal brain metabolites.[3] Additionally, research indicates that the number of painful procedures during infancy is linked to reduced intelligence and altered white matter development.[4,5] Furthermore, early pain experiences have been shown to influence pain perception later in life.[6-8] Consequently, caregivers should prioritize minimizing infant pain by reducing procedural interventions, mitigating factors that exacerbate pain, and assessing pain severity to determine the appropriate use of analgesic medications. Nevertheless, most studies focus on very preterm infants, a high-risk group undergoing multiple painful procedures due to severe illness.[3,9] While the impact of painful procedures in late preterm and term infants is acknowledged, it is less emphasized compared to very preterm infants. This discrepancy arises because procedures in late preterm and term infants are brief and less invasive, leading caregivers to underestimate the importance of pain assessment and management.

Due to infants' inability to directly communicate pain, assessments primarily rely on behavioral observations, such as changes in facial expressions, using tools like the neonatal facial coding system and the facial action coding system.[10] However, preterm infants may exhibit reduced facial pain expressions of pain.[11] Therefore, additional psychometric factors are incorporated, including the Premature Infant Pain Profile – revised (PIPP-R), Cry, Requires oxygen, Increased vital signs, Expression, Sleeplessness (CRIES), Neonatal Infant Pain Scale (NIPS), COMFORT, and COMFORTneo, as well as the Neonatal Pain, Agitation, and Sedation Scale (N-PASS).[12] Among these, NIPS and PIPP-R scores are widely used.[13] In our previous study, we found no significant changes in cerebral oxygenation in late preterm and term infants during painful procedures.[14] Consequently, pain score assessment remains the primary method for evaluating pain in this population. Since newborn infants may exhibit varying pain responses depending on gestational age, postnatal age, or other factors, as well as differing levels of pain depending on procedure type[15], this study hypothesizes that both infant characteristics and procedure

type influence pain scores. Therefore, the aim of this study was to identify factors associated with PIPP-R scores, one of the most widely used pain assessment tools, in late preterm and term infants undergoing blood sampling procedures.

## MATERIALS AND METHODS

A secondary analysis was conducted using data from a previous prospective cohort study on pain response in late preterm and term infants.[14] The study protocol was approved by Siriraj Institutional Review Board (COA no. Si 435/2022), and written parental consent was obtained prior to infant recruitment. The study was carried out at the high-risk nursery at Siriraj Hospital, Mahidol University in Bangkok, Thailand. In accordance with institutional protocol, hemodynamically stable preterm infants (<37 weeks of gestation) who did not require positive-pressure ventilation for respiratory compromise were admitted to the high-risk nursery. Infants with very low birthweight, respiratory compromise or hemodynamic instability were admitted to either intermediate care or the neonatal intensive care unit. Gestational age (GA) was determined using maternal ultrasonographic data from the first trimester, or when available, through postnatal clinical examination.

The enrolled infants satisfied the following inclusion criteria: 1) born within the GA range of 34 to 41 weeks, 2) postnatal age less than 14 days, 3) exhibited respiratory stability without requiring positive-pressure ventilation, maintained hemodynamic stability without requiring vasoactive agents, and 4) had a physician's order for skin puncture procedures. The exclusion criteria included infants exhibiting signs of neonatal encephalopathy, such as abnormal movements, hypotonia, or apnea, as well as those with severe congenital anomalies. To ensure biological diversity and reduce potential biases, each infant was enrolled only once, and in cases of multiple pregnancies, only one twin was included to minimize the influence of genetic factors.

The care of neonates, including the choice of blood sampling, either heelstick or venipuncture, was determined at the clinical discretion of the attending physician. For procedures, a 25G needle was used for IM injections at the anterolateral thigh, a 24G needle for venipunctures or IV procedures, and a 28G needle for heelsticks. The institutional protocol for procedural pain control includes oral sucrose solution and non-pharmacologic measures such as swaddling and non-nutritive sucking. However, to allow for unobstructed assessment of facial expressions required for accurate pain scoring, non-nutritive sucking was not provided in this cohort. The infant was placed supine on a radiant warmer during the procedure. Cerebral oxygenation ($CrSO_2$) was monitored using the INVOS 5100C device as part of the primary study, while oxygen saturation ($SpO_2$) and heart rate were concurrently evaluated using the Nellcor™ Bedside $SpO_2$ Patient Monitoring System with a MAXN sensor affixed to the infant's right hand. The procedure adhered to standard protocols and included simultaneous video documentation for subsequent pain evaluation. The video recordings focused on capturing the infant's facial expressions and monitoring heart rate and $SpO_2$, providing a comprehensive evaluation of responses to the procedures. Pain assessment was conducted using the Premature Infant Pain Profile-Revised (PIPP-R) score.[16] A single assessor (PT), blinded to the type and duration of each procedure, analyzed the recordings to ensure unbiased evaluation. The intraclass correlation coefficients for PIPP-R scores was 1.00. Another co-author (SR) independently extracted procedure-related variables from the same set of video recordings to minimize ascertainment bias.

### Statistical analysis

This secondary analysis included all eligible infants from the parent cohort (n = 59) who underwent skin-puncture procedures including heelstick, venipuncture, or intramuscular injection during the study period.[14] Demographic characteristics were reported as numbers (percentages), mean ± standard deviation (SD), or median [$25^{th}$, $75^{th}$ percentile], depending on variable type and distribution. Correlations between gestational age or procedure duration and pain scores were assessed using Pearson correlation coefficients ($r$). Differences in PIPP-R scores across groups exposed to specific variables were analyzed using the Mann-Whitney U or Kruskal-Wallis tests. Potential factors associated with pain scores were identified through univariate linear regression, with variables showing $p$ <0.2 further evaluated via backward multiple linear regression analysis. Results were reported as regression coefficients with 95% confidence intervals (CI). Statistical analyses were performed using SPSS version 29.0, with $p$ <0.05 considered statistically significant.

## RESULTS

This secondary analysis included 59 procedures from 59 infants, of whom 29 (49.2%) were born at a gestational age (GA) ≥ 37 weeks, and 22 (37.3%) were classified as low birthweight (LBW; birthweight <2,500 g). Maternal and infant demographic characteristics are detailed in Table 1. The median postnatal age at the time of the procedures was 5 [2, 18] hours. Among the procedures,

**TABLE 1.** Maternal and infant demographic characteristics.

| | |
|---|---|
| **Mothers (n = 59)** | |
| Age (year) | 30.9 ± 5.7 |
| Gestational age (week) | 36 [35,38] |
| Twin pregnancy | 3 (5.1) |
| Hypertension related condition | 6 (10.2) |
| Diabetes | 18 (30.5) |
| Antenatal magnesium sulphate | 3 (5.1) |
| Antenatal dexamethasone | 24 (40.7) |
| Systemic intrapartum pain control | 13 (22) |
| Cesarean section | 30 (50.8) |
| **Infants (n = 59)** | |
| Birth weight (g) | 2,690 ± 520 |
| Male sex | 34 (57.6) |
| 1-minute Apgar score | 8 [8,9] |
| 5-minute Apgar score | 10 [9,10] |
| Birth resuscitation | 8 (13.6) |
| Small-for-gestational age | 9 (15.3) |
| Large-for-gestational age | 4 (6.8) |
| Postnatal age (hour) | 5 [2,18] |

Data are presented as mean ± SD, number (%), or median [P25, P75]. * $p$ <0.05 indicates statistical significance.

31 (52.5%) were venipunctures for intravenous (IV) fluid administration, 15 (25.4%) were intramuscular (IM) injections, and 13 (22%) were heelsticks. Five infants (8.5%) received oral glucose solution for pain relief. The median procedure durations were 34 [5, 51] seconds for venipunctures, 11 [6, 12] seconds for IM injections, and 1 [1, 1] second for heelsticks. The median PIPP-R score across all procedures was 11.0 [7.0, 14.0]. A weak positive correlation was observed between GA and PIPP-R scores ($r$= 0.32, $p$= 0.01) but no significant correlation was found between procedure duration and PIPP-R scores ($r$= -0.06, $p$= 0.68).

Table 2 summarizes the associations between maternal and infant factors and PIPP-R scores, using both univariate and multiple regression analyses. Infants who received oral sucrose solution had significantly lower PIPP-R scores compared to those who did not, with an unadjusted regression coefficient of -6.8 (95% CI: -10.9, -2.7; $p$ = 0.002). This association remained robust after adjustment, with an adjusted coefficient of -6.4 (95% CI: -9.4, -3.3; $p$ < 0.001). LBW was associated with lower PIPP-R scores. In the univariate analysis, the regression coefficient was -5.3 (95% CI: -7.5, -3.2; $p$ < 0.001), and this association persisted in the adjusted analysis, with a coefficient of -2.9 (95% CI: -4.8, -1.1; $p$ = 0.002). Postnatal age <48 hours was significantly associated with higher PIPP-R scores. The univariate regression coefficient was 3.2 (95% CI: 0.0, 6.5; $p$ = 0.05), and this association strengthened in the adjusted analysis, with a coefficient of 5.8 (95% CI: 3.3, 8.4; $p$ < 0.001).

The type of procedure also significantly influenced pain scores. IM injections were associated with the highest PIPP-R scores compared to venipunctures, with an adjusted regression coefficient of 4.6 (95% CI: 2.3, 6.8; $p$ < 0.001). Heelstick procedures had higher PIPP-R scores compared to venipunctures, with an adjusted coefficient of 2.8 (95% CI: 0.6, 5.0; $p$ = 0.02).

**TABLE 2.** Factors associated with PIPP-R scores.

| Factors | n | PIPP-R scores | *p** | Univariate linear regression analysis Coefficient (95% CI) | *p*† | Multiple linear regression analysis Coefficient (95% CI) | *p** |
|---|---|---|---|---|---|---|---|
| **Mothers (n = 59)** | | | | | | | |
| Systemic intrapartum pain control | | | 0.50 | -1.3 (-4.3,1.6) | 0.37 | | |
| No | 46 | 11.5 [7.0,14.0] | | | | | |
| Yes | 13 | 10.0 [3.0,14.5] | | | | | |
| Antenatal magnesium sulphate | | | 0.01* | -5.1 (-10.6,0.4) | 0.07† | -2.7 (-6.5,1.1) | 0.17 |
| No | 56 | 12.0 [7.0,14.0] | | | | | |
| Yes | 3 | 5.0 [4.0,-] | | | | | |
| Antenatal dexamethasone | | | 0.19 | -1.5 (-4.0,1.0) | 0.25 | | |
| No | 35 | 12.0 [7.0,15.0] | | | | | |
| Yes | 24 | 10.5 [4.3,13.0] | | | | | |
| Cesarean section | | | 0.86 | -0.3 (-2.8,2.2) | 0.82 | | |
| No | 29 | 12.0 [7.0,14.0] | | | | | |
| Yes | 30 | 11.0 [4.8,14.0] | | | | | |
| **Infants (n = 59)** | | | | | | | |
| Birthweight <2,500 g | | | <0.001* | -5.3 (-7.5,-3.2) | <0.001† | -2.9 (-4.8,-1.1) | 0.002* |
| No | 37 | 13.0 [11.0,15.0] | | | | | |
| Yes | 22 | 5.0 [3.0,10.3] | | | | | |
| Gestational age <37 weeks | | | 0.06 | -2.1 (-4.5,0.3) | 0.09† | | |
| No | 29 | 12.0 [7.5,15.0] | | | | | |
| Yes | 30 | 10.5 [4.0,13.0] | | | | | |
| Postnatal age <48h | | | 0.03* | 3.2 (0.0,6.5) | 0.05† | 5.8 (3.3,8.4) | <0.001* |
| No | 49 | 11.0 [5.0,13.0] | | | | | |
| Yes | 10 | 14.5 [9.5,15.3] | | | | | |
| Female sex | | | 0.87 | -0.1 (-2.6,2.4) | 0.95 | | |
| No | 34 | 12.0 [6.3,14.0] | | | | | |
| Yes | 25 | 11.0 [6.0,14.5] | | | | | |
| Resuscitation | | | 0.38 | 1.8 (-1.8,5.4) | 0.33 | | |
| No | 51 | 11.0 [5.0,14.0] | | | | | |
| Yes | 8 | 13.0 [10.3,14.0] | | | | | |
| Size-for-gestational age | | | 0.21 | | | | |
| Appropriate-for-gestational age | 46 | 11.0 [6.5,15.0] | | Reference | | | |
| Small-for-gestational age | 9 | 7.0 [3.5,12.5] | | -2.4 (-5.8,1.0) | 0.17† | | |
| Large-for-gestational age | 4 | 13.5 [12.3,14.0] | | 2.9 (-2.0,7.8) | 0.24 | | |
| Glucose oral solution | | | 0.004* | -6.8 (-10.9,-2.7) | 0.002† | -6.4 (-9.4,-3.3) | <0.001* |
| No | 54 | 12.0 [7.0,14.3] | | | | | |
| Yes | 5 | 4.0 [2.0,6.0] | | | | | |
| Procedure duration <1 minute | | | 0.54 | -1.0 (-4.3,2.4) | 0.56 | | |
| No | 10 | 12.5 [8.5,15.0] | | | | | |
| Yes | 49 | 11.0 [6.0,14.0] | | | | | |
| Procedure type | | | 0.007* | | | | |
| Intravenous | 31 | 9.0 [6.5,12.8] | | Reference | | | |
| Intramuscular | 15 | 14.0 [12.0,15.0] | | 4.4 (1.6,7.2) | 0.002† | 4.6 (2.3,6.8) | <0.001* |
| Heelstick | 13 | 9.5 [6.5,12.8] | | 0.5 (-2.4,3.4) | 0.74 | 2.8 (0.6,5.0) | 0.02* |

## DISCUSSION

This study investigated factors influencing pain scores in late preterm and term infants undergoing routine procedures. By focusing on hemodynamically stable infants, the findings are broadly applicable to a low-risk neonatal population. Using the PIPP-R, a validated pain assessment tool that incorporates gestational age, behavioral, and physiological indicators, we observed that even brief procedures resulted in PIPP-R scores within the moderate to severe range (interquartile range: 7–14). These findings highlight the importance of implementing pain mitigation strategies during routine neonatal care, regardless of procedure duration.[17]

Interestingly, LBW infants were associated with lower PIPP-R scores, with a regression coefficient of -2.9 (95% CI: -4.8, -1.1). LBW infants are often preterm or experience fetal growth restriction, which can result in developmental immaturity that impacts their behavioral and physiological pain responses.[17] These infants may exhibit less pronounced facial expressions or movements, partly due to underdeveloped motor control and limited energy reserves.[8] Additionally, postnatal illnesses such as hypoglycemia, commonly seen in LBW infants, may further blunt their responses to painful stimuli.

Although not statistically significant, our findings showed lower pain scores in infants whose mothers received MgSO4. This may reflect a potential dampening effect of MgSO4 on the infant's ability to respond to painful stimuli or a direct neuromodulatory action. Given the small number of exposed infants in our study, this observation should be interpreted with caution and warrants further investigation.

While this study included infants with GA starting at 34 weeks, where late preterm infants receive an additional point on the PIPP-R score due to their GA indicator, our findings revealed that term infants had higher PIPP-R scores compared to late preterm infants. Despite this adjustment, term infants exhibited higher PIPP-R scores, and a weak positive correlation between GA and PIPP-R scores was observed. This correlation suggests that advancing GA might contribute to slightly heightened pain perception or expression. However, these findings indicate that the additional scoring point for late preterm infants did not significantly impact the overall pain score. This aligns with our previous findings showing no significant differences in cerebral oxygenation responses between late preterm and term infants,[14] emphasizing that factors other than GA likely play a more critical role in shaping pain perception in this population. Unexpectedly, postnatal age <48 hours was associated with higher PIPP-R scores, with a regression coefficient of 5.8 (95% CI: 3.3, 8.4). The selection of 48 hours as a cutoff assumed that neonates beyond this age would have transitioned through the critical phase of postnatal adaptation. This finding may be explained by the reduced exposure to medical interventions in neonates under 48 hours old, as the intensity of acute postnatal care typically decreases after this period. The lower procedural frequency may result in reduced cumulative stress and less sensitization to painful stimuli, potentially explaining the lower pain scores. Additionally, neonates older than 48 hours may exhibit better behavioral regulation, further attenuating their pain responses.

The duration of the procedure did not significantly correlate with pain scores ($r$ = -0.06, $p$ = 0.68). However, the procedural route significantly influenced pain intensity, with IM punctures resulting in higher PIPP-R scores compared to IV and heelstick procedures. Factors such as the depth of needle insertion, muscle contractions, and localized tissue drug accumulation during IM injections likely contribute to more intense nociceptor activation compared to the more superficial nature of IV or heelstick procedures. In this study, the needle sizes used were 25G for IM injections, 24G for IV procedures, and 28G for heelsticks. Although the IM needle was slightly larger than the heelstick needle, it was not the primary factor contributing to the increased pain scores. Previous research in adults has shown no significant association between needle gauge and pain response.[18] Instead, the deeper tissue penetration and mechanical irritation of muscle fibers during IM injections may trigger stronger muscle contractions and nociceptor activation.[19] Additionally, the localized accumulation of injected substances in muscle tissue likely amplifies discomfort compared to the smaller, more superficial punctures of IV and heelstick procedures. Our findings indicate that PIPP-R scores for heelstick procedures were higher than those for IV procedures [regression coefficients 2.8 (95% CI: 0.6, 5.0), $p$ = 0.02]. This trend aligns with findings from a meta-analysis conducted in term infants[20], which suggests that IV routes should be prioritized when feasible to reduce pain intensity. Given the necessity of IM injections for routine neonatal care, such as vitamin K administration and vaccinations, it is crucial to implement effective analgesic measures to minimize discomfort, even for brief procedures. Strategies such as oral sucrose solution, swaddling, or facilitated tucking should be considered to improve procedural comfort and reduce pain.[5,17]

Oral sucrose solution demonstrated significant efficacy in reducing pain, with neonates receiving sucrose exhibiting lower PIPP-R scores compared to those who did not [regression coefficients: -6.4 (95% CI: -9.4, -3.3);

$p < 0.001$]. Its analgesic effect, mediated through activation of endogenous opioid pathways, aligns with previous studies documenting reduced pain and distress during procedures such as heelsticks, venipunctures, and IM injections.[21,22] However, only 8.5% of neonates in this study received sucrose, highlighting gaps in protocol adherence. To address this, a quality improvement project has been initiated to promote consistent and efficient use of oral sucrose for procedural pain management.

This study has several strengths, including its focus on routine neonatal procedures and the use of PIPP-R, a widely accepted pain assessment tool. Blinding assessors to procedure type and duration minimized bias, enhancing validity. Nevertheless, certain limitations exist. The analyzed procedures were brief and may not fully reflect pain responses to longer or more invasive interventions. Additionally, all included late preterm and term infants were hemodynamically stable, which likely reduced variability in physiological responses. Critically ill neonates, such as very preterm infants or those with hemodynamic instability, may exhibit more pronounced differences in pain responses,[23] highlighting the need for further investigation. PIPP-R scores may partially capture anxiety rather than pain alone, suggesting the need for exploring complementary non-pharmacological interventions such as swaddling or facilitated tucking. Finally, our findings must be considered in light of the relatively small sample inherent to this secondary analysis. Because our cohort size was determined by the parent study, no pre-specified sample size calculation was possible, and we may have been underpowered to detect associations for less common variables or modest effect sizes. Although this limitation cannot be remedied retrospectively, the trends and confidence intervals we observed offer valuable insights and generate hypotheses for further investigation. Larger, prospectively powered studies in more diverse neonatal populations will be essential to validate and extend these results and to deepen our understanding of how infant characteristics and procedural factors influence pain responses.

## CONCLUSION

IM injections were associated with higher pain scores compared to IV procedures or heelsticks. Gestational age demonstrated a weak correlation with pain scores, while procedure duration did not influence pain levels. Lower pain scores were observed in LBW infants, whereas higher scores were observed in those <48 hours postnatal. Oral sucrose solution was effective in reducing pain. These findings underscore the critical need for consistent pain mitigation strategies to enhance procedural comfort in neonates.

### Data Availability Statement

The datasets generated and/or analysed during this study are available from the corresponding author upon reasonable request.

## ACKNOWLEDGEMENT

We thank Miss Juthatip Suwannatrai for her assistance in facilitating the submission process.

## DECLARATIONS

### Grants and Funding Information

The authors received no financial support for this study.

### Conflict of Interest

All authors declare no conflicts of interest.

### Registration Number of Clinical Trial

This study is not a registered clinical trial. Ethical approval was obtained from the Siriraj Institutional Review Board, Faculty of Medicine Siriraj Hospital, Mahidol University (COA no. Si 435/2022).

### Use of Artificial Intelligence

Artificial intelligence was used solely to assist with English language editing during the preparation of the initial draft of this manuscript. All study concepts, analysis, interpretation, and writing were initiated and carried out by the authors. The final version of the manuscript, prior to submission, was further reviewed and edited by a professional English-language editor to ensure clarity and accuracy.

### Author Contributions

The authors confirm their contribution to the paper as follows: study conception and design: RK and SR; data collection: SR and PT; analysis and interpretation of results: RK, SR and IA; draft manuscript preparation and critical revision: RK, SR, and IA. All authors reviewed the results and approved the final version of the manuscript.

## REFERENCES

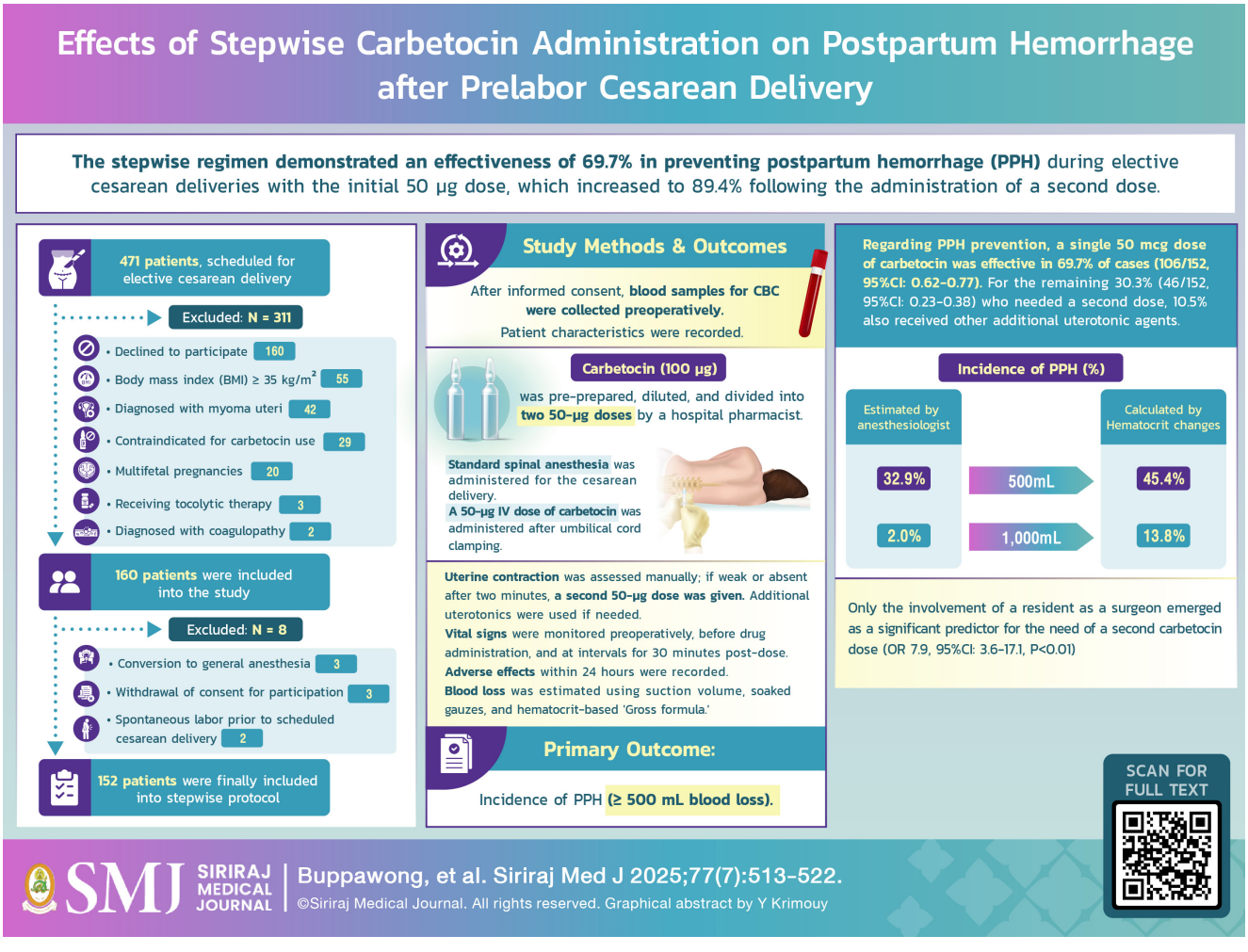
1. McPherson C, Miller SP, El-Dib M, Massaro AN, Inder TE. The influence of pain, agitation, and their management on the immature brain. Pediatr Res. 2020;88(2):168-75.
2. Buser JR, Maire J, Riddle A, Gong X, Nguyen T, Nelson K, et al. Arrested preoligodendrocyte maturation contributes to myelination failure in premature infants. Ann Neurol. 2012;71(1):93-109.
3. Brummelte S, Grunau RE, Chau V, Poskitt KJ, Brant R, Vinall J, et al. Procedural pain and brain development in premature newborns. Ann Neurol. 2012;71(3):385-96.
4. Vinall J, Miller SP, Bjornson BH, Fitzpatrick KP, Poskitt KJ, Brant R, et al. Invasive procedures in preterm children: brain

and cognitive development at school age. Pediatrics. 2014;133(3): 412-21.
5. Johnston CC, Fernandes AM, Campbell-Yeo M. Pain in neonates is different. Pain. 2011;152(3 Suppl):S65-s73.
6. Walker SM, Melbourne A, O'Reilly H, Beckmann J, Eaton-Rosen Z, Ourselin S, et al. Somatosensory function and pain in extremely preterm young adults from the UK EPICure cohort: sex-dependent differences and impact of neonatal surgery. Br J Anaesth. 2018;121(3):623-35.
7. Hermann C, Hohmeister J, Demirakça S, Zohsel K, Flor H. Long-term alteration of pain sensitivity in school-aged children with early pain experiences. Pain. 2006;125(3):278-85.
8. Williams MD, Lascelles BDX. Early neonatal pain—a review of clinical and experimental implications on painful conditions later in life. Front Pediatr. 2020;8:30.
9. Slater R, Fabrizi L, Worley A, Meek J, Boyd S, Fitzgerald M. Premature infants display increased noxious-evoked neuronal activity in the brain compared to healthy age-matched term-born infants. Neuroimage. 2010;52(2):583-9.
10. Craig KD, Hadjistavropoulos HD, Grunau RV, Whitfield MF. A comparison of two measures of facial activity during pain in the newborn child. J Pediatr Psychol. 1994;19(3):305-18.
11. Gibbins S, Stevens B, McGrath PJ, Yamada J, Beyene J, Breau L, et al. Comparison of pain responses in infants of different gestational ages. Neonatology. 2008;93(1):10-8.
12. Maxwell LG, Fraga MV, Malavolta CP. Assessment of pain in the newborn: an update. Clin Perinatol. 2019;46(4):693-707.
13. Olsson E, Ahl H, Bengtsson K, Vejayaram DN, Norman E, Bruschettini M, et al. The use and reporting of neonatal pain scales: a systematic review of randomized trials. Pain. 2021; 162(2):353-60.
14. Amornjiraporn I, Rugsapol S, Thanasarnpaiboon P, Paes B, Kitsommart R. A comparison of the effect of procedural pain on cerebral oxygen saturation between late preterm and term infants. J Perinatol. 2024; doi: 10.1038/s41372-024-01978-4.
15. Wilson-Smith EM. Procedural pain management in neonates, infants and children. Rev Pain. 2011;5(3):4-12.
16. Gibbins S, Stevens BJ, Yamada J, Dionne K, Campbell-Yeo M, Lee G, et al. Validation of the Premature Infant Pain Profile-Revised (PIPP-R). Early Hum Dev. 2014;90(4):189-93.
17. Good practice in postoperative and procedural pain management, 2nd edition. Paediatr Anaesth. 2012;22 Suppl 1:1-79.
18. Yoon SH, Rah UW, Sheen SS, Cho KH. Comparison of 3 needle sizes for trigger point injection in myofascial pain syndrome of upper- and middle-trapezius muscle: a randomized controlled trial. Arch Phys Med Rehabil. 2009;90(8):1332-9.
19. Tucker MH, Tiwari P, Carter BS. The physiology, assessment, and treatment of neonatal pain. Semin Fetal Neonatal Med. 2023; 28(4):101465.
20. Shah VS, Ohlsson A. Venepuncture versus heel lance for blood sampling in term neonates. Cochrane Database Syst Rev. 2011; doi: 10.1002/14651858.CD001452.pub4(10).
21. Stevens B, Yamada J, Ohlsson A, Haliburton S, Shorkey A. Sucrose for analgesia in newborn infants undergoing painful procedures. Cochrane Database Syst Rev. 2016;7(7):Cd001069.
22. Witt N, Coynor S, Edwards C, Bradshaw H. A guide to pain assessment and management in the neonate. Curr Emerg Hosp Med Rep. 2016;4(1):1-10.
23. Moolmai P, Rattanachamnongk P, Yangthara B, Wutthigate P. Factors influencing bronchopulmonary dysplasia: an eight-year study in a single tertiary care unit in Thailand. Siriraj Med J. 2025;77(2):158-67.

# Effects of Stepwise Carbetocin Administration on Postpartum Hemorrhage after Prelabor Cesarean Delivery

**Punyaniti Buppawong, M.D.[1], Tripop Lertbunnaphong, M.D.[1,*], Pattarawan Limsiri, M.D.[1], Anisa Phetses, R.N.[1], Patchareya Nivatpumin, M.D.[2], Namtip Triyasunant, M.D.[2]**

*[1]Department of Obstetrics and Gynecology, Faculty of Medicine Siriraj Hospital, Mahidol University. Bangkok, Thailand, [2]Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University. Bangkok, Thailand.*




*Corresponding author: Tripop Lertbunnaphong
E-mail: tripop.ler@mahidol.ac.th
Received 11 January 2025 Revised 25 February 2025 Accepted 26 February 2025
ORCID ID:http://orcid.org/0000-0002-1571-1679
https://doi.org/10.33192/smj.v77i7.273108

## ABSTRACT

**Objective:** To evaluate the effectiveness of stepwise administration of carbetocin in reducing postpartum hemorrhage (PPH) in term pregnancies undergoing elective cesarean delivery.

**Materials and Methods:** This study included term pregnancies scheduled for cesarean delivery. Exclusion criteria were a BMI ≥ 35 kg/m$^2$, any underlying medical conditions that contraindicated the use of carbetocin and conditions affecting uterine contraction. Following newborn delivery, an initial dose of 50 µg of carbetocin was administered, with a subsequent 50 µg given if uterine contraction was inadequate. The primary outcome was the incidence of PPH exceeding 500 mL, while secondary outcomes included total blood loss ≥ 1,000 mL, administration of additional uterotonic agents, adverse effects, other pregnancy-related and surgical outcomes.

**Results:** Of the 152 pregnant women analyzed, adequate uterine contraction was observed in 69.7% of cases after the first carbetocin dose, with an additional 19.7% achieving adequate contraction after the second dose. Estimated blood loss exceeding 500 mL was 32.9% and exceeding 1,000 mL was 2.0% of cases, respectively. Recalculations based on pre- and post-hematocrit levels indicated higher rates of blood loss: more than 500 mL in 45.4% and over 1,000 mL in 13.8%. No serious adverse events or complications were reported.

**Conclusion:** The stepwise regimen demonstrated an effectiveness of 69.7% in preventing postpartum hemorrhage (PPH) during elective cesarean deliveries with the initial 50 µg dose, which increased to 89.4% following the administration of a second dose.


**Keywords:** Carbetocin; stepwise administration; cesarean delivery; postpartum hemorrhage (Siriraj Med J 2025; 77: 513-522)

## INTRODUCTION

Postpartum hemorrhage is the leading cause of maternal mortality worldwide.[1] While women with high-risk factors—such as prolonged third stage of labor, pregnancy-induced hypertensive disorders, uterine atony, birth canal injury, or retained placental tissue—are particularly susceptible, postpartum hemorrhage can also occur in low-risk pregnancies, regardless of the mode of delivery.[1,2] To reduce its incidence, the administration of uterotonic agents in the third stage of labor is a critical intervention.[3] Carbetocin is a next-generation synthetic oxytocin that acts as a long-acting oxytocin agonist. It is used to prevent postpartum hemorrhage (PPH) by promoting rhythmic uterine contractions.[4] Recent meta-analyses have confirmed that carbetocin reduces the need for additional uterotonic agents compared with oxytocin in preventing PPH, particularly in elective cesarean delivery.[5,6]

Determining the optimal dosage of carbetocin for PPH prevention remains a question. Initially, the standard dose of carbetocin (100 µg) was proved to be effective for preventing PPH after cesarean delivery and was recommended to be the 1$^{st}$ line drug for PPH prevention in guidelines.[7,8] The standard dose of carbetocin, set at 100 µg, was based on the equivalent dose of oxytocin 5 IU.[4] However, in a sequential allocation trial, the effective dose for 90% of cases (ED90) was found to be 14.8 µg of carbetocin.[9] Among obese women (BMI ≥ 40 kg/m$^2$) the minimum effective dose of carbetocin was 62.9 µg.[10] In pregnant women with abnormal labor and a history of oxytocin usage curves who underwent cesarean delivery, the minimum effective dose was 121 µg of carbetocin.[11]

In two studies, a wide range of carbetocin doses (20-120 µg) was administered but, the results were inconclusive.[12,13] Similar to those of oxytocin, carbetocin has some adverse effects, including nausea, vomiting, headache, hypotension, and chills.[1] Therefore, reducing the dose may also reduce adverse effects. We hypothesized that a half dose of carbetocin (50 µg) would be both sufficient and practical for effectively reducing PPH in pre-labor cesarean delivery. The aim of this prospective study was to analyze the effects of stepwise administration of carbetocin (50 µg/step) in reducing PPH in term pregnancies undergoing cesarean delivery.

## MATERIALS AND METHODS

The study protocol was approved by the Institutional Review Board (IRB) on April 18, 2022 (COA No. Si 313/2022). This prospective study was conducted from January 1$^{st}$ to June 30$^{th}$, 2023. The study employed a stepwise approach using carbetocin in women undergoing a pre-labor cesarean section. The inclusion criteria were women with a term singleton pregnancy scheduled for elective cesarean delivery at Siriraj Hospital. The exclusion criteria were women with a history of carbetocin allergy, cardiovascular, liver, kidney diseases, epilepsy, migraine,

or asthma. Additionally, obese women (BMI ≥ 35 kg/m$^2$) and those with high-risk pregnancies affecting uterine contractility, such as preeclampsia, placenta previa, placenta accreta spectrum (PAS), polyhydramnios, uterine structural abnormalities, hematologic diseases, were excluded. Women who had received tocolytic drugs within 24 hours preoperatively were also excluded. The standard anesthetic approach for pregnant women undergoing cesarean section is regional anesthesia (spinal or epidural anesthesia). Therefore, cases requiring conversion to general anesthesia, which might impact uterine contraction, are excluded from this study.

After obtaining informed consent, blood samples for a complete blood count were collected the day prior to surgery. Characteristics of the women were recorded. Carbetocin was prepared in advance from a 100 µg ampule (Duratocin®, Ferring Pharmaceutical Co., Ltd, Thailand), divided into two syringes, and diluted with normal saline to 5 mL per syringe by a hospital pharmacist on the day of surgery. Women underwent cesarean delivery under spinal anesthesia, which was administered using local anesthetic combined with intrathecal morphine, with standard intraoperative monitoring. Following fetal delivery and clamping of the umbilical cord, a 50-µg dose of diluted carbetocin was administered intravenously. The assessment of uterine contraction was performed through manual palpation by the surgeon, and was categorized as strong, moderate, weak, or absent. If after two minutes of administration, uterine contraction was assessed weak or absent, a second dose of 50 µg was administered. Additional uterotonic agents were administered if uterine contractions remained inadequate. The anesthesia team monitored and recorded the pulse and blood pressure before surgery, before drug administration, and then two minutes after administration, and subsequently every five minutes for 30 minutes. Any adverse effects within the first 24 hours post-operatively were documented by interviewing the patient, with medications used for treatment recorded. To estimate blood loss during surgery, the anesthesiologist documented the volume of blood collected in the suction canister and from weighed soaked gauzes or swabs. Twenty-four hours postoperatively, a blood sample was obtained from the patient for a complete blood count (CBC), and the results were subsequently used to calculate the volume of blood loss utilizing the 'Gross formula', a method for assessing blood loss based on changes in hematocrit levels.[14]

The primary outcome of this study was the incidence of PPH (blood loss ≥ 500 mL).[1] Secondary outcomes included total blood loss, blood loss ≥ 1,000 mL, administration of additional uterotonic agents, adverse effects, other pregnancy-related and surgical outcomes including operative time, pre- and post-operative hemoglobin and hematocrit levels, adverse events after operation.

The sample size was calculated based on 2022 data from Siriraj Hospital. The retrospective data of 50 pregnant women undergoing elective cesarean delivery in early 2022 showed average rate of blood loss exceeding 500 mL following cesarean delivery was 40%. Considering a type I error rate of 0.05 and a power of 0.8, the calculated sample size was 145. To account for potential dropout, an additional 10% was added, bringing the total to 160 pregnant women.

### Statistical analyses

Data analysis was conducted using SPSS Statistics version 18.0 for Windows (IBM Corp., Armonk, NY, USA). Descriptive data were presented as numbers and percentages (n, %) for categorical variables and as mean ± standard deviation (SD) or median (IQR) for continuous variables, depending on data distribution.

Normality of continuous variables was assessed using the Kolmogorov-Smirnov or Shapiro-Wilk test. If normally distributed, variables were summarized as mean ± SD and compared using the independent t-test or one-way ANOVA. If non-normally distributed, data were reported as median (IQR) and compared using the Mann-Whitney U test or the Kruskal-Wallis test. Categorical variables were presented as n (%) and compared using the Chi-square test. Fisher's exact test was used when any expected count was <5. A p-value < 0.05 was considered statistically significant.

To identify factors associated with the need for an additional dose of carbetocin, univariate analysis was first performed for each predictor variable. Variables with p ≤ 0.2 in the univariate analysis were considered for further evaluation in multivariate analysis. Multivariate logistic regression was then conducted using SPSS stepwise regression (both forward and backward selection), which iteratively adds or removes variables based on statistical significance criteria. The final model retained only independent predictors with p-values < 0.05. Results were reported as odds ratios (ORs) with 95% confidence intervals (CIs). A p-value of < 0.05 was considered statistically significant.

## RESULTS

A total of 471 pregnant women participated, with 152 included (Fig 1). Scheduled cesarean deliveries occurred at approximately 38 weeks of gestation. The majority of pregnant women were overweight (55.3%) or obese (13.8%). Maternal anemia (Hb <11 g/dL) was observed in 19.7% of cases (Table 1).

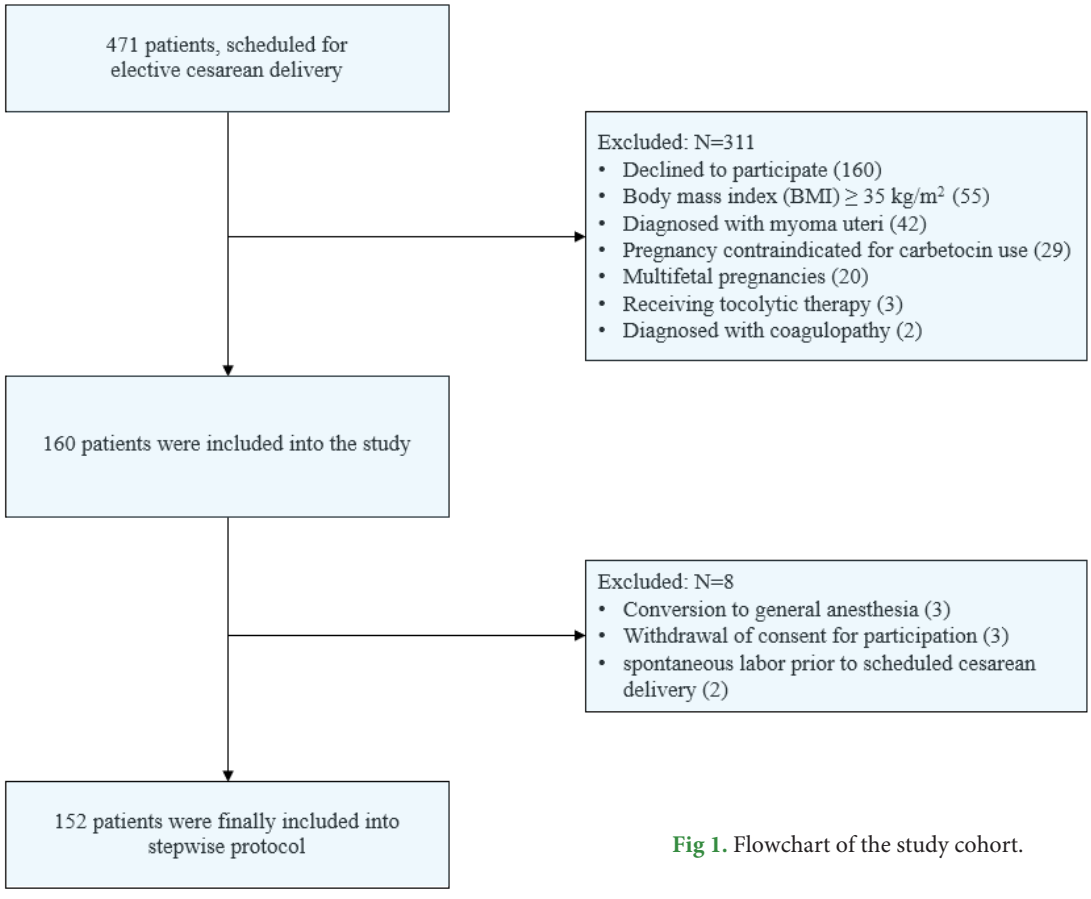



**Fig 1.** Flowchart of the study cohort.

**TABLE 1.** Characteristics of study population (N=152).

| Characteristic | Mean ± SD or N (%) (n=152) |
|---|---|
| Age (year) | 34.01 ± 4.78 |
| Body weight on admission (kg) | 69.10 ± 9.85 |
| Weight gain (kg) | 13.58 ± 5.03 |
| Gestational age (week) | 38.18 ± 0.55 |
| BMI (kg/m²) | 26.65 ± 3.35 |
| BMI (kg/m²) | |
| Underweight (<18.5) | 2 (1.3) |
| Normal weight (18.5-24.9) | 45 (29.6) |
| Overweight (25.0-29.9) | 84 (55.3) |
| Obesity (30.0-35.0) | 21 (13.8) |
| Baseline hemoglobin (g/dL) | 11.95 ± 1.03 |
| Baseline hematocrit (%) | 36.42 ± 2.94 |
| Maternal anemia | 30 (19.7) |
| Fetal birth weight (g) | 3122.76 ± 377.94 |

Carbetocin was effective as a single dose in 69.7% (106/152, 95% CI: 0.624–0.770), while 30.3% (46/152, 95% CI: 0.230–0.376) required an additional dose, with 10.5% also receiving other uterotonic agents. Based on the estimate of the anesthesiologist, 32.9% of pregnant women experienced blood loss exceeding 500 mL and 2.0% exceeded 1,000 mL, increasing to 45.4% and 13.8%, respectively, when assessed by hematocrit changes. Those requiring a second dose or additional uterotonic agents had significantly greater blood loss and hematologic changes ($P < 0.05$). Hypotension (≥20% decrease in SBP or DBP) occurred in 74% (113/152) of pregnant women, with 48% (54/113) requiring vasopressors. Vasopressor use did not differ significantly between groups (50 µg: 34% vs. 100 µg: 39%, $P \geq 0.05$). No severe adverse events occurred, and adverse effects did not differ between groups (Table 2).

Table 3 demonstrates median (IQR) values for pulse rate, SBP, and DBP post-carbetocin administration. Median vital sign changes were minimal, with no significant differences in pulse rate and SBP ($P \geq 0.05$). However, DBP was significantly lower in the 100 µg group at 2 and 5 minutes post-administration (49.0 (17) vs. 57.0 (14) mmHg, $P < 0.01$ at 2 min; 49.0 (14) vs. 56.0 (12) mmHg, $P < 0.01$ at 5 min).

**TABLE 2.** Outcomes and factors associated with the stepwise protocol (N=152).

| | N (%) (n=152) | Carbetocin dose N (%) 50 ug group (n=106) | 100 ug group (n=46) | p-value* |
|---|---|---|---|---|
| Parity | | | | |
| Nullipara | 72 (47.4) | 58 (54.7) | 14 (30.4) | 0.01 |
| Multipara | 80 (52.6) | 48 (45.3) | 32 (69.6) | |
| Previous cesarean section | | | | |
| Yes | 70 (46.1) | 43 (40.6) | 27 (58.7) | 0.04 |
| No | 82 (53.9) | 63 (59.4) | 19 (41.3) | |
| Surgeon | | | | |
| Attending physician | 99 (65.1) | 84 (79.2) | 15 (32.6) | <0.001 |
| Resident | 53 (34.9) | 22 (20.8) | 31 (67.4) | |
| Maternal anemia (Before cesarean section) | 30 (19.7) | 22 (20.8) | 8 (17.4) | 0.63 |
| Maternal anemia (After cesarean section) | 79 (52.0) | 53 (50.0) | 26 (56.5) | 0.46 |
| Need of additional uterotonic agents | 16 (10.5) | 0 (0.0) | 16 (34.8) | <0.001 |
| Need of vasopressor drugs | 54 (35.5) | 36 (34.0) | 18 (39.1) | 0.54 |
| Operative time | 55.7 (18.2) | 51.9 (17.0) | 64.4 (17.9) | <0.01 |
| Blood loss | | | | |
| Estimated blood loss | | | | |
| < 500 mL | 99 (65.1) | 82 (77.4) | 17 (37.0) | <0.001 |
| ≥ 500-999 mL | 50 (32.9) | 23 (21.7) | 27 (58.7) | <0.001 |
| ≥ 1,000 mL | 3 (2.0) | 1 (0.9) | 2 (4.3) | 0.17 |
| Calculated blood loss | | | | |
| < 500 mL | 62 (49.6) | 57 (51.0) | 8 (17.4) | 0.03 |
| ≥ 500-999 mL | 69 (45.4) | 42 (39.6) | 27 (58.7) | 0.03 |
| ≥ 1,000 mL | 21 (13.8) | 10 (9.4) | 11 (23.9) | 0.02 |
| Adverse effects | | | | |
| Headache | 3 (2.0) | 3 (2.8) | 0 (0.0) | 0.25 |
| Dizziness | 24 (15.8) | 16 (15.1) | 8 (17.4) | 0.72 |
| Nausea/vomiting | 63 (41.4) | 40 (37.7) | 23 (50.0) | 0.16 |
| Flushing | 4 (2.6) | 1 (0.9) | 3 (6.5) | 0.05 |
| Bitterness | 10 (6.6) | 6 (5.7) | 4 (8.7) | 0.49 |
| Chills | 28 (18.4) | 19 (17.9) | 9 (19.6) | 0.81 |
| Shivering | 15 (9.9) | 12 (11.3) | 3 (6.5) | 0.36 |
| Pruritus | 74 (48.7) | 55 (51.9) | 19 (41.3) | 0.23 |

**TABLE 2.** Outcomes and factors associated with the stepwise protocol (N=152). (Continue)

| | Mean ± SD (n=152) | Mean ± SD 50 ug (n=106) | 100 ug (n=46) | p-value* |
|---|---|---|---|---|
| Estimated blood loss (mL) | 416.1 ± 179.2 | 378.8 ± 161.5 | 502.2 ± 189.7 | <0.001 |
| Postoperative hemoglobin (g/dL) | 11.9 ±1.0 | 11.9 ±1.0 | 12.1 ± 1.0 | 0.16 |
| Postoperative hematocrit (%) | 36.4 ± 2.9 | 36.2 ± 3.0 | 36.9 ± 2.8 | 0.16 |
| Postoperative hemoglobin (g/dL) | 11.0 ± 1.2 | 11.0 ± 1.2 | 10.9 ± 1.3 | 0.57 |
| Postoperative hematocrit (%) | 33.4 ± 3.4 | 33.5 ± 3.3 | 33.2 ± 3.7 | 0.52 |
| | **Median (IQR) (n=152)** | **Median (IQR) 50 ug (n=106)** | **100 ug (n=46)** | **p-value*** |
| Calculated blood loss (mL) | 419.6 (625.0) | 364.5 (577.0) | 555.9 (820.0) | 0.02 |
| Change of Hemoglobin (g/dL) | 0.9 (1.4) | 0.8 (1.2) | 1.3 (1.6) | 0.03 |
| Change of Hematocrit (%) | 2.7 (4.1) | 2.4 (3.75) | 3.6 (4.7) | 0.03 |

* Statistical significance (P<0.05)

**Abbreviations:** SD: standard deviation, IQR: interquartile range

Multiparity, previous cesarean delivery, resident and operative time were identified as significant factors in the univariate analysis (Table 2). However, the only significant predictor of requiring an additional carbetocin dose was surgeries performed by residents (OR: 7.9, 95% CI: 3.6–17.1, P < 0.001), whereas multiparity and prior cesarean delivery were not retained in the final model (P > 0.05) (Table 4). Additionally, surgeries performed by residents were significantly longer than those conducted by attending physician (68.3 ± 17.8 vs. 49.0 ± 14.5 minutes, P < 0.01). However, the incidence of PPH did not significantly differ between groups. The proportion of cases with blood loss ≥ 500 mL was 44.4% in faculty-performed surgeries and 47.2% in resident-supervised surgeries (P > 0.05). For blood loss ≥ 1,000 mL, the rates were 11.1% and 18.9%, respectively (P > 0.05).

## DISCUSSION

This study demonstrated that a single 50 µg dose of carbetocin effectively prevented PPH after cesarean delivery in 69.7% of cases, while 30.3% required an additional dose. Pregnant women requiring additional uterotonic agents experienced greater blood loss and more pronounced hematologic changes. However, no severe adverse events were reported. These findings support the efficacy of carbetocin in reducing the need for additional uterotonic agents.

Postpartum hemorrhage is the leading cause of maternal mortality worldwide.[3] Current evidence unequivocally supports the pivotal role of medical interventions in reducing the incidence of PPH. According to a meta-analysis by *Gallos et al.*, carbetocin may be considered as the preferred agent for preventing PPH during elective cesarean deliveries.[6] While its efficacy in this context has been well established, the cost-effectiveness of carbetocin remains the subject of debate in clinical practice. Studies conducted in developed countries have demonstrated its cost-effectiveness and potential to reduce overall patient care costs; however, these findings have not been consistently replicated in developing countries.[15-17] Based on the findings of this study, reducing the carbetocin dosage to 50 mcg may be a viable and clinically appropriate strategy. However, cost-effectiveness analyses of the half-dose regimen are needed to support its wider clinical implementation.

The potential for reducing carbetocin dosage to use for PPH prevention among elective cesarean deliveries has been demonstrated in clinical research. Previous studies indicate the efficacy of various carbetocin regimens, especially in lower doses, in sustaining sufficient uterine

**TABLE 3.** Pulse and blood pressure after drug administration (N=152).

| | Median (IQR) (n=152) | Carbetocin dose Median (IQR) 50 ug (n=106) | 100 ug (n=46) | p-value* |
|---|---|---|---|---|
| **Pulse (bpm)** | | | | |
| 2 min | 91.0 (17) | 90.5 (16) | 93.5 (16) | 0.35 |
| 5 min | 89.0 (19) | 88.5 (22) | 90.0 (14) | 0.88 |
| 10 min | 88.5 (15) | 90 (15) | 87.5 (18) | 0.42 |
| 15 min | 88.0 (16) | 88.0 (16) | 87.5 (16) | 0.99 |
| 20 min | 87.0 (17) | 86.0 (17) | 88.0 (17) | 0.86 |
| 25 min | 84.0 (18) | 83.0 (16) | 85.0 (20) | 0.32 |
| 30 min | 82.0 (16) | 82.0 (15) | 82.0 (20) | 0.67 |
| **Systolic blood pressure (mmHg)** | | | | |
| 2 min | 110 (16) | 111 (16) | 105 (21) | 0.06 |
| 5 min | 111 (17) | 112 (16) | 110 (21) | 0.32 |
| 10 min | 111 (15) | 109.5 (16) | 111 (16) | 0.33 |
| 15 min | 113 (17) | 112.5 (17) | 116.0 (18) | 0.50 |
| 20 min | 112 (19) | 111.0 (17) | 112.0 (20) | 0.25 |
| 25 min | 115 (16) | 114.5 (17) | 115.0 (18) | 0.40 |
| 30 min | 115 (15) | 115.0 (15) | 115.0 (14) | 0.76 |
| **Diastolic blood pressure (mmHg)** | | | | |
| 2 min | 54 (15) | 57.0 (14) | 49.0 (17) | <0.01 |
| 5 min | 54 (14) | 56.0 (12) | 50.0 (13) | <0.01 |
| 10 min | 53.5 (11) | 54.0 (13) | 53.0 (9) | 0.61 |
| 15 min | 55 (11) | 54.0 (11) | 58.5 (9) | 0.55 |
| 20 min | 56 (13) | 55.0 (11) | 57.0 (13) | 0.34 |
| 25 min | 60 (12) | 60.0 (12) | 60.0 (11) | 0.80 |
| 30 min | 61 (13) | 61.0 (13) | 63.0 (11) | 0.29 |

*Statistical significance (P<0.05)
**Abbreviations:** bpm: beat per minute, IQR: interquartile range

**TABLE 4.** Multiple logistic regression analysis of factors associated with 100 ug carbetocin.

| Variable | Multivariate analysis | | | | |
|---|---|---|---|---|---|
| | beta | SE | $\chi^2$ | p-value* | OR (95%CI) |
| Multiparity | 1.24 | 0.79 | 2.48 | 0.12 | 3.4 (0.74-16.01) |
| Previous CS | -1.39 | 0.80 | 2.98 | 0.08 | 0.3 (0.05-1.21) |
| Operative time > 60 minutes | -1.32 | 0.50 | 0.07 | 0.79 | 0.9 (0.33-2.31) |
| Resident | 2.29 | 0.52 | 18.83 | <0.01 | 9.8 (3.50-27.61) |
| Constant | -1.78 | 0.34 | 28.10 | <0.01 | 0.2 |

*Statistical significance (P<0.05), SE: standard errors, $\chi^2$: Wald $\chi^2$, OR: odds ratios, and 95% CI: 95% confidence intervals, CS: cesarean section

contraction during cesarean delivery across different patient conditions.[9,10,13] Building on these findings, our study suggests that a reduced dose of 50 µg of carbetocin is adequate for promoting uterine contraction in elective cesarean deliveries. However, we excluded pregnant women with a BMI over 35 kg/m$^2$ for ethical reasons, as these cases required carbetocin doses greater than 50 µg.[10] It's important to note that despite excluding pregnancies preoperatively diagnosed with myomas, we encountered 8 cases with intraoperative diagnoses of myomas, highlighting the need for careful risk assessment when applying this stepwise protocol.

The stepwise administration of 50 µg of carbetocin demonstrated its efficacy in reducing PPH following elective cesarean delivery. In our study, two-third of pregnant women (69.7%) required only a single dose, while one-third (30.3%) required an additional dose. The need for additional uterotonic agents was low (10%), with no severe adverse events reported. These findings align with previous research highlighting carbetocin's role in minimizing PPH risk.[5,18,19]

Blood loss during cesarean delivery is frequently underestimated due to the inherent limitations of traditional measurement techniques, such as suction canister readings and visual estimation of gauze saturation, both of which are susceptible to significant error. Our analysis revealed that conventional estimates underreported blood loss by a factor of 0.7 for cases exceeding 500 mL (32.9% estimated vs. 45.4% recalculated) and by a factor of seven for losses greater than 1,000 mL (2.0% estimated vs. 13.8% recalculated). While hematocrit-based recalculations provide a more objective assessment, their accuracy remains influenced by intraoperative fluid administration and physiological variations. Despite these constraints, hematocrit-based estimation represents an improvement over the conventional method. To further enhance accuracy and optimize postpartum hemorrhage management in cesarean deliveries, the implementation of advanced quantitative blood loss measurement tools, similar to those used in vaginal deliveries, may improve accuracy and enhance PPH management in cesarean deliveries.[20,21] Multiple logistic regression analysis identified surgical inexperience as a significant factor associated with the need for additional doses of carbetocin. Conducted at a tertiary care center, our residency training program follows a standard structure comparable to other university hospitals worldwide. First-year residents assisted in cesarean delivery under faculty supervision, progressing to independent performance in later years, with oversight and consultation available as needed. In practice, the decision to administer an additional dose was guided by the study protocol rather than anesthesiologist warnings or concerns about liability for PPH. Nevertheless, the diverse levels of experience among these surgeons may have impacted their decisions regarding the administration of additional doses of carbetocin. While residents were more likely to administer an additional dose, the study found no significant difference in the incidence of PPH between surgeries performed by residents and those performed by faculty. This finding suggests that surgical experience affects clinical decision-making during cesarean delivery rather than directly influencing the occurrence of complications.

In line with prior research, the most frequently reported adverse effects included pruritus (48.7%), nausea/vomiting (41.4%), chills (18.4%), and dizziness (15.8%).[11,22,23] Hypotension was also prevalent, and observed in 74.3% of pregnant women, with a subset requiring vasopressor intervention. It's important to note that some of these adverse effects cannot be definitively attributed to carbetocin alone as they could be caused by other anesthetic agents used during surgery. For example, spinal block anesthesia can cause a drop in diastolic blood pressure (DBP) due to reduced parasympathetic tone and intrathecal morphine is known to cause pruritus, as well as nausea and vomiting. Nevertheless, it's worth emphasizing that all these clinical manifestations were not severe and were manageable.

A major strength of this study is its prospective experimental design, which provides a structured evaluation of a low-dose carbetocin regimen with a stepwise administration protocol in a real-world setting. The preparation of the diluted carbetocin dose was conducted under strict quality control by the Siriraj pharmacist unit, ensuring accuracy and consistency. Additionally, blood loss assessment incorporated both intraoperative estimates and an objective standardized equation based on pre- and postoperative hematocrit levels, offering a more comprehensive evaluation of actual blood loss.

However, the study has certain limitations. As a single-center study, the findings are influenced by institutional protocols and patient demographics, which may limit their direct applicability to other healthcare settings. Nevertheless, the documentation and analysis of clinical practice in a controlled manner provide valuable insights that may inform broader clinical implementation. Furthermore, the study employed a single-unblinded design to reflect real-world clinical decision-making. This approach may have introduced variability in the administration of additional uterotonic agents, as the determination of uterine tone relied on individual clinical judgment.

Given the widespread use of oxytocin as the first-line agent for PPH prevention, further research is needed to compare the efficacy and safety of this stepwise carbetocin regimen with standard oxytocin protocols. Multicenter randomized trials could clarify its benefits in uterine tone stabilization, blood loss reduction, and broader clinical applications.

## CONCLUSION

The stepwise regimen demonstrated an effectiveness of 69.7% in preventing postpartum hemorrhage (PPH) during elective cesarean deliveries with the initial 50 μg dose, which increased to 89.4% following the administration of a second dose.

### Data Availability Statement

The data supporting this study are not publicly available due to ethical concerns and patient confidentiality but may be obtained from the corresponding author upon reasonable request and with appropriate institutional approval.

## ACKNOWLEDGMENTS

The authors sincerely appreciate Mrs. Alisala Lertbannaphong and the Siriraj pharmacist team for their invaluable assistance in drug preparation, as well as the labor room and postpartum ward nursing team, led by Ms. Ratree Sirisomboon, R.N., Miss Metpapha Sudphet, and Mrs. Amporn Ownon, R.N. We also thank Dr. Saowalak Hunnangkul for her expertise in statistical analysis. Additionally, we acknowledge the financial support from the Siriraj Research Development Fund, administered by Routine to Research (R2R).

## DECLARATION

### Grants and Funding Information

This study was supported by the Siriraj Research Development Fund (Managed by Routine to Research; R2R), Grant Number (IO) R016535046.

### Conflict of Interest

Tripop Lertbunnaphong has received a speaker honorarium from Ferring Pharmaceutical Co., Ltd, Thailand. The other authors have no relevant financial or non-financial interests to disclose.

### Registration Number of Clinical Trial

TCTR20231009002 (retrospective registration on October 9, 2023)

### Author Contributions

P.B.: Protocol development, Material preparation, Data collection, Data analysis, Manuscript writing, Funding management. T.L.: Protocol development, Manuscript editing, Funding management. P.L.: Data collection, Data analysis. A.P.: Material preparation, Data collection. P.N.: Protocol development, Manuscript editing. N.T.: Protocol development, Manuscript editing. All authors reviewed and agreed to the final version of the manuscript.

### Use of Artificial Intelligence

During the preparation of this work, the authors used ChatGPT to refine the English language for publication. After utilizing this tool, the authors reviewed and edited the content as necessary and took full responsibility for the final version of the publication.

## REFERENCES

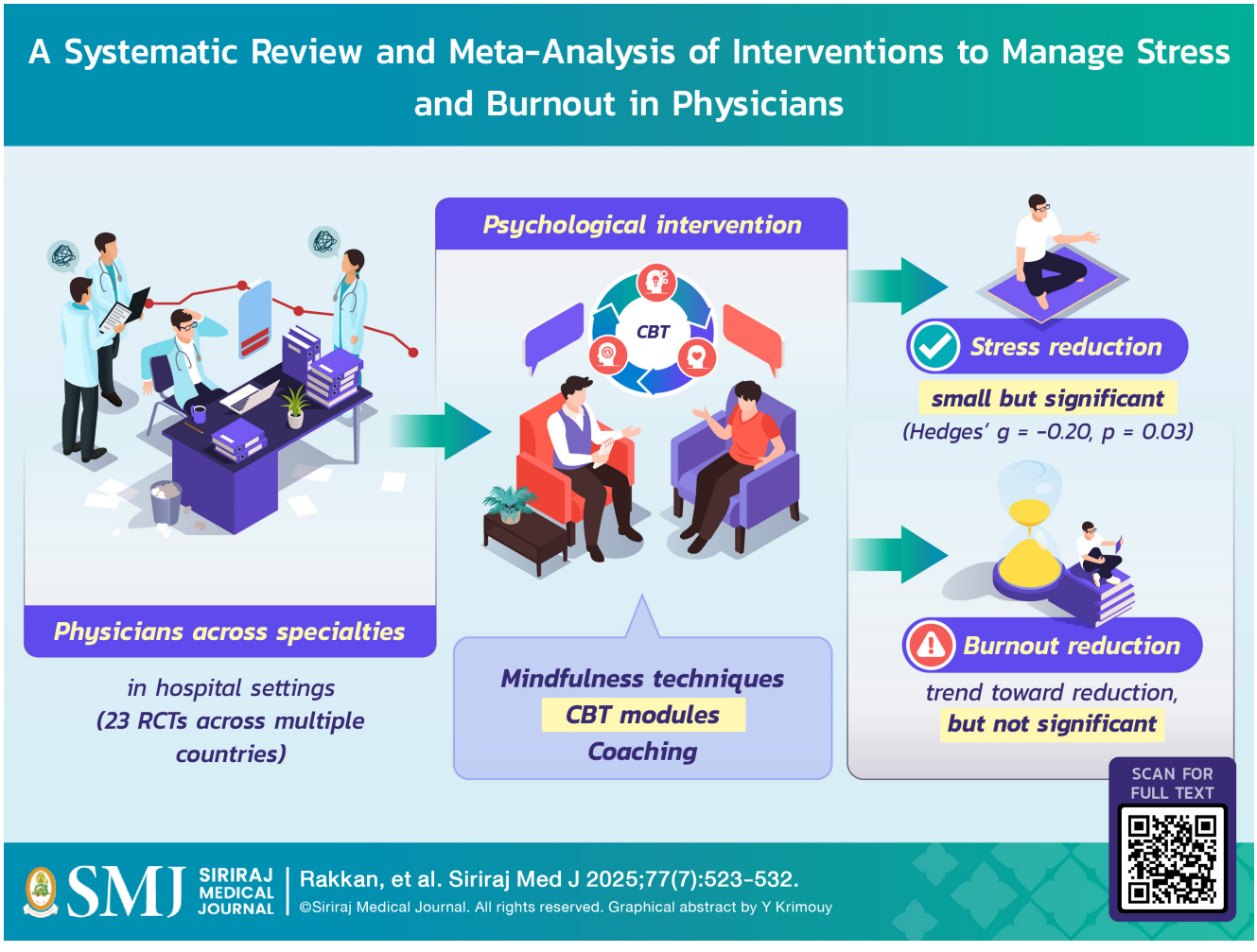
1. Vogel JP, Williams M, Gallos I, Althabe F, Oladapo OT. WHO recommendations on uterotonics for postpartum haemorrhage prevention: what works, and which one? BMJ Glob Health. 2019;4(2):e001466.
2. Lertbunnaphong T LJ, Thitadilok W. Risk Factors of Primary Postpartum Hemorrhage in Siriraj Hospital. Siriraj Med J. 2020;62(5):195-8.
3. Group FPTW, Begum F, Beyeza J, Burke T, Evans C, Hanson C, et al. FIGO and the International Confederation of Midwives endorse WHO guidelines on prevention and treatment of postpartum hemorrhage. Int J Gynaecol Obstet. 2022;158 Suppl 1(Suppl 1):6-10.
4. Hunter DJ, Schulz P, Wassenaar W. Effect of carbetocin, a long-acting oxytocin analog on the postpartum uterus. Clin Pharmacol Ther. 1992;52(1):60-7.
5. Onwochei DN, Van Ross J, Singh PM, Salter A, Monks DT. Carbetocin reduces the need for additional uterotonics in elective caesarean delivery: a systematic review, meta-analysis and trial sequential analysis of randomised controlled trials. Int J Obstet Anesth. 2019;40:14-23.
6. Gallos I, Williams H, Price M, Pickering K, Merriel A, Tobias A, et al. Uterotonic drugs to prevent postpartum haemorrhage: a network meta-analysis. Health Technol Assess. 2019;23(9): 1-356.
7. Leduc D, Senikas V, Lalonde AB, Clinical Practice Obstetrics C. Active management of the third stage of labor: prevention and treatment of postpartum hemorrhage. J Obstet Gynaecol Can. 2009;31(10):980-93.
8. Heesen M, Carvalho B, Carvalho JCA, Duvekot JJ, Dyer RA, Lucas DN, et al. International consensus statement on the use of uterotonic agents during caesarean section. Anaesthesia. 2019; 74(10):1305-19.
9. Khan M, Balki M, Ahmed I, Farine D, Seaward G, Carvalho JC. Carbetocin at elective Cesarean delivery: a sequential allocation trial to determine the minimum effective dose. Can J Anaesth. 2014;61(3):242-8.
10. Drew T, Balki M, Farine D, Ye XY, Downey K, Carvalho JCA. Carbetocin at elective caesarean section: a sequential allocation

trial to determine the minimum effective dose in obese women. Anaesthesia. 2020;75(3):331-7.
11. Nguyen-Lu N, Carvalho JC, Farine D, Seaward G, Ye XY, Balki M. Carbetocin at Cesarean delivery for labor arrest: a sequential allocation trial to determine the effective dose. Can J Anaesth. 2015;62(8):866-74.
12. Cordovani D, Balki M, Farine D, Seaward G, Carvalho JC. Carbetocin at elective Cesarean delivery: a randomized controlled trial to determine the effective dose. Can J Anaesth. 2012;59(8):751-7.
13. Anandakrishnan S, Balki M, Farine D, Seaward G, Carvalho JC. Carbetocin at elective Cesarean delivery: a randomized controlled trial to determine the effective dose, part 2. Can J Anaesth. 2013;60(11):1054-60.
14. Lin YM, Yu C, Xian GZ. Calculation methods for intraoperative blood loss: a literature review. BMC Surg. 2024;24(1):394.
15. Matthijsse S, Andersson FL, Gargano M, Yip Sonderegger YL. Cost-effectiveness analysis of carbetocin versus oxytocin for the prevention of postpartum hemorrhage following vaginal birth in the United Kingdom. J Med Econ. 2022;25(1):129-37.
16. You JHS, Leung TY. Cost-effectiveness analysis of carbetocin for prevention of postpartum hemorrhage in a low-burden high-resource city of China. PLoS One. 2022;17(12):e0279130.
17. Briones JR, Talungchit P, Thavorncharoensap M, Chaikledkaew U. Economic evaluation of carbetocin as prophylaxis for postpartum hemorrhage in the Philippines. BMC Health Serv Res. 2020;20(1):975.
18. Amornpetchakul P, Lertbunnaphong T, Boriboonhiransarn D, Leetheeragul J, Sirisomboon R, Jiraprasertwong R. Intravenous carbetocin versus intravenous oxytocin for preventing atonic postpartum hemorrhage after normal vaginal delivery in high-risk singleton pregnancies: a triple-blind randomized controlled trial. Arch Gynecol Obstet. 2018;298(2):319-27.
19. Kalafat E, Gokce A, O'Brien P, Benlioglu C, Koc A, Karaaslan O, et al. Efficacy of carbetocin in the prevention of postpartum hemorrhage: a systematic review and Bayesian meta-analysis of randomized trials. J Matern Fetal Neonatal Med. 2021;34(14):2303-16.
20. Bell SF, Watkins A, John M, Macgillivray E, Kitchen TL, James D, et al. Incidence of postpartum haemorrhage defined by quantitative blood loss measurement: a national cohort. BMC Pregnancy Childbirth. 2020;20(1):271.
21. Quantitative Blood Loss in Obstetric Hemorrhage: ACOG COMMITTEE OPINION, Number 794. Obstet Gynecol. 2019;134(6):e150-e6.
22. Widmer M, Piaggio G, Nguyen TMH, Osoti A, Owa OO, Misra S, et al. Heat-Stable Carbetocin versus Oxytocin to Prevent Hemorrhage after Vaginal Birth. N Engl J Med. 2018;379(8):743-52.
23. van der Nelson H, O'Brien S, Burnard S, Mayer M, Alvarez M, Knowlden J, et al. Intramuscular oxytocin versus Syntometrine((R)) versus carbetocin for prevention of primary postpartum haemorrhage after vaginal birth: a randomised double-blinded clinical trial of effectiveness, side effects and quality of life. BJOG. 2021;128(7):1236-46.

# A Systematic Review and Meta-Analysis of Interventions to Manage Stress and Burnout in Physicians

Nopparat Rakkan, M.D.[1,2], Punyisa Prachgosin, M.D., FRCPsychT,[1,*] Kamonporn Wannarit, M.D.[1]

[1]*Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand,* [2]*Udonthani Hospital, Udon Thani, Thailand.*




*Corresponding author: Punyisa Prachgosin
E-mail: punyisa.pra@mahidol.edu
Received 12 March 2025 Revised 17 April 2025 Accepted 20 April 2025
ORCID ID:http://orcid.org/0000-0001-6612-5891
https://doi.org/10.33192/smj.v77i7.273470

## ABSTRACT

**Objective:** To assess interventions aimed at reducing stress and burnout among physicians.

**Materials and Methods:** A systematic review and meta-analysis was performed in line with PRISMA guidelines. Randomized controlled trials published in English up to December 2022 were identified through searches in medical databases, including PubMed, EMBASE, and CINAHL. The primary outcomes were stress and burnout levels.

**Results:** A total of 13 studies investigated interventions to reduce stress and burnout, including mindfulness and non-mindfulness-based approaches. Non-mindfulness interventions included professional coaching, group-coaching programs, and training in humanism and professionalism. Outcomes were assessed at two time points: immediately post-intervention and at the final follow-up. Stress-reduction interventions significantly decreased stress levels (Hedges' g = -0.20, 95% CI: -0.37 to -0.02, p = 0.03; $I^2$ = 6.66%). In contrast, burnout-reduction interventions showed a non-significant trend toward improvement (Hedges' g = -0.08, 95% CI: -0.32 to 0.17, p = 0.53; $I^2$ = 43.89%).

**Conclusion:** Stress-reduction interventions are effective in lowering stress among physicians. However, improvements in burnout remain limited, possibly due to variations in intervention effectiveness across studies. Further research should focus on long-term follow-ups to address well-being and resilience of healthcare professionals.


**Keywords:** Burnout interventions; stress interventions; physicians (Siriraj Med J 2025; 77: 523-532)

## INTRODUCTION

Stress is a natural response to life pressures, particularly in demanding professions like medicine. The Health and Safety Executive (HSE) identifies six key factors contributing to workplace stress: demand, control, support, relationships, role and change.[1] Addressing these factors can help reduce stress, which impacts both individuals and organizations.[2,3] In individuals, stress is linked to anxiety and frustration, which can lead to mistakes, substance abuse, and reduced quality of life and self-confidence. At the organizational level, stress contributes to higher absenteeism and turnover rates, lower work quality and job satisfaction, and increased workplace conflicts.

Research shows that medical professionals, such as doctors and nurses, experience stress due to work-related pressures, patient expectations, and inadequate support. A recent study from Thailand also highlights the impact of occupational stress among physicians, including mental health professionals, indicating that this issue is a global concern.[41] Key factors include the lack of counseling services, ineffective workplace communication, and insufficient feedback.[4] *Weinberg and Creed* (2000) found a link between stress and an increased risk of depression and anxiety among healthcare professionals. Therefore, organizations must prioritize managing stress among medical personnel.[5]

*Giga et al.* (2003) reported that cognitive behavior therapy (CBT) and Employee Assistance Programs (EAP) are widely used stress management interventions in the United Kingdom.[6] Similarly, *Ruotsalainen et al.* (2015) found in a Cochrane review that stress prevention strategies, including relaxation training, moderately reduce stress levels. Additionally, organizational adjustments, such as revising work schedules, can also help.[7] A meta-analysis by *Regehr et al.* (2014) further highlighted the effectiveness of CBT and mindfulness-based techniques in reducing stress and burnout among physicians.[8] Emerging online stress management programs show promise. *Grime* (2004) demonstrated that an eight-week computerized CBT program effectively reduced depression and anxiety, with participants valuing its accessibility and flexibility.[9]

However, research on stress management for physicians, particularly through online methods, remains limited and often relies on non-randomized trials. This study focuses on randomized controlled trials to enhance stress management guidelines for this group.

## MATERIALS AND METHODS

This meta-analysis was approved by the Siriraj Institutional Review Board (Si 770/2561) and conducted in accordance with the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA).[34]

Two researchers (N.P. and P.P.) independently conducted a comprehensive search in Embase (from January 1988), and in PubMed and CINAHL (from January 1999) through to December 2022, on 17 June 2023. The search included the following keywords: "physicians", "stress management", "computerized stress management", "stress intervention", "stress prevention","burnout intervention", and "wellness intervention". We applied a language filter during the database search to include only English-

language articles. Therefore, only randomized controlled trials published in English were screened for eligibility. The primary outcomes of interest were measures of stress and burnout.

Two authors, (N.P. and P.P.) independently extracted data. Any disagreements were resolved through discussion with a third author (K.W.). Extracted data included first author, publication year, country, study type, sample size, participant characteristics, intervention type and duration, stress and burnout measures, and post-intervention/final follow-up scores (mean and standard deviation).

Risk of bias was assessed using the Cochrane Risk-of-Bias Tool for Randomized Trials (RoB 2).[42]

### Statistical analyses

The primary outcomes were stress and burnout levels at post-intervention and final follow-up. To ensure robustness, an intention-to-treat analysis was performed. Heterogeneity was assessed using the $I^2$ statistic, with thresholds interpreted as follows: low (<25%), moderate (25–50%), and high (>50%).[43] A fixed-effects model was used when heterogeneity was low ($I^2 \leq 25\%$), whereas a random-effects model was applied when moderate to high heterogeneity ($I^2 > 25\%$) was present. Publication bias was evaluated using funnel plots.

## RESULTS

### Study selection

A literature search across three databases (Embase: 1,363 records, PubMed: 4,905 records, and CINAHL: 283 records) yielded a total of 6,551 records. After removing duplicates, 332 records remained. Of these, 219 records were excluded based on title and abstract screening. After assessing 113 full-text studies, 44 were excluded for not focusing exclusively on physician populations. Out of 69 identified studies, 44 were eligible, leaving 25 for qualitative synthesis, though two lacked outcome measurements. For quantitative synthesis, 23 studies were included, with 13 undergoing an in-depth review. Among these, 10 studies did not comprehensively measure burnout. Details are shown in Fig 1.

### Study characteristics

The quantitative synthesis included 23 studies, with 13 focusing on stress and burnout reduction interventions. These studies examined the effectiveness of various interventions, including professional coaching to enhance well-being, group-coaching to reduce burnout, training in humanism and professionalism, and mindfulness-based approaches. Ten studies assessed specific components of burnout such as emotional exhaustion, depersonalization,

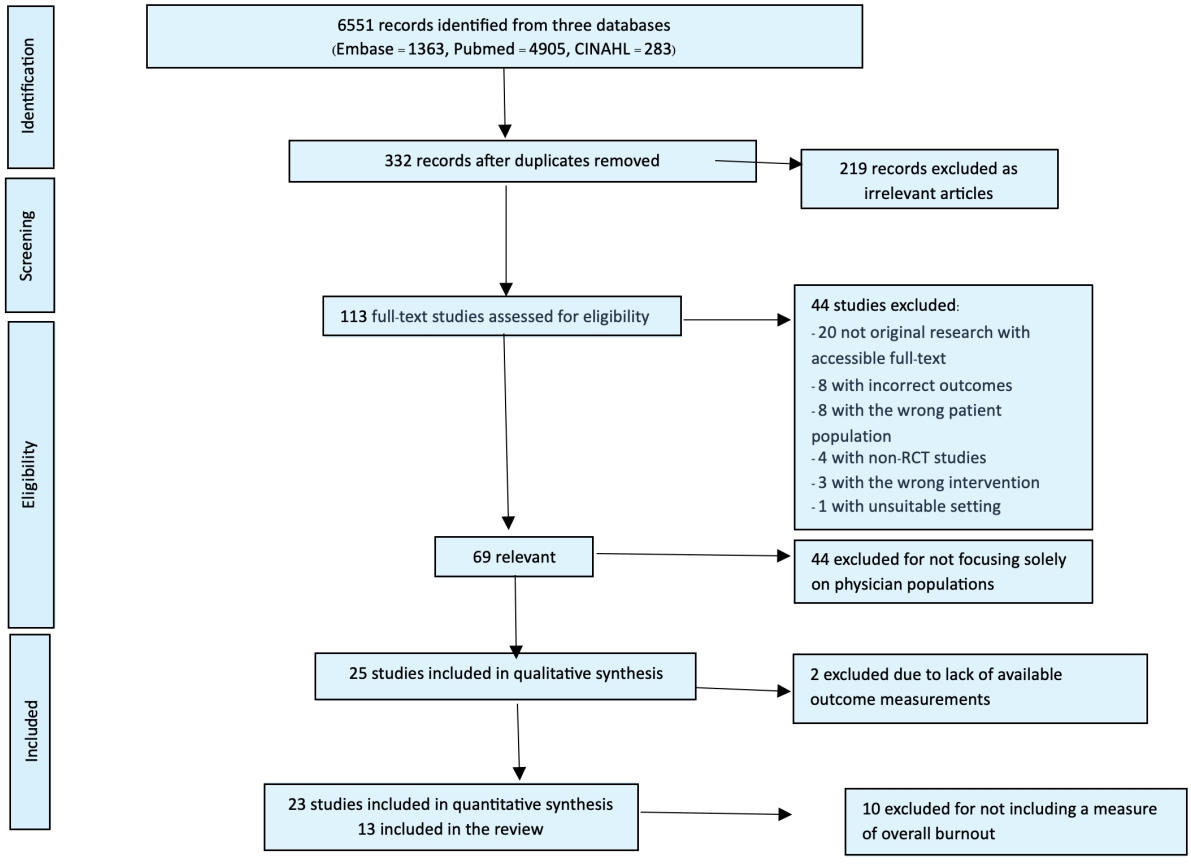



**Fig 1.** Study Inclusion PRISMA flowchart.

**TABLE 1.** Details of physicians, interventions and outcomes in each study.

| Study (Year) (Sample size) | Types of Physicians | Comparison | Intervention | Post-intervention follow-up |
|---|---|---|---|---|
| *Bragard et al.* (2010) (n=72) | Oncologists | Waitlist control | Basic Training for 1 month, followed by 2 months of bi-monthly consolidation workshops | 5 months |
| *West et al.* (2014) (n=74) | Internal medicine physicians | No treatment | 18 theme-based discussion sessions, held twice monthly and led by expert facilitators | 12 months |
| *Amutio et al.* (2015) (n=42) | Physicians in eight departments | Waitlist control | Mindfulness-Based Stress Reduction (MBSR) program: 8 weekly sessions (2.5 hours each) + 8-hour retreat; 10 months of maintenance involving self-practice and mindfulness exercises. | 12 months |
| *Gunasingam et al.* (2015) (n=31) | Intern doctors | No treatment | Four debriefing sessions on stress-related topic discussion, led by facilitators | - |
| *Bernburg et al.* (2016) (n=54) | Physicians in a pediatric department | No treatment | Psychosocial Competency Training (PCT): 12 weekly sessions (1.5 hours each), including cognitive-behavioral and solution- focused approaches. | 6 months |
| *Dyrbye et al.* (2016) (n=290) | Medicine and surgical physicians | No treatment | A 10-week online intervention featuring independent, task-based activities | after 1 month |
| *Ireland et al.* (2017) (n=44) | Interns in an emergency department | 1-hour extra break in a week | A 10-week mindfulness training program | 10 weeks |
| *Mache et al.* (2017) (n=78) | Gynecology and obstetrics physicians | No treatment | 1.5-hour discussions on problem-solving and emotion regulation via Lazarus's stress model over 3 months. | 9 months |
| *Mache et al.* (2018) (n=70) | Junior physicians in an emergency department | Waitlist control | 12 weekly 1.5-hour sessions combining coping skills, cognitive-behavioral strategies, and solution-focused counseling. | after 6 months |
| *Verweij et al.* (2018) (n=148) | Medical, surgical and primary care residents | Waitlist control | Mindfulness-Based Stress Reduction (MBSR) program: 8 weekly 2.5-hour sessions with a silent day | 3 months |
| *Axisa et al.* (2019) (n=46) | Physician trainees in internal medicine and pediatrics | No treatment | 4.5-hour activities covering stress management, group discussion, feedback and a meal. | after 3 months |
| *Dyrbye et al.* (2019) (n=88) | Physicians in medicine, family medicine, and pediatrics | Waitlist control | 6 professionally facilitated coaching sessions | 5 months |
| *Lebares et al.* (2019) (n=21) | 1st-year surgical residents | Active control (different content, same structure) | A 2-hour Mindfulness-Based Stress Reduction (MBSR) session each week for 8weeks, with daily 20-minute home practice. | 12 months |

**TABLE 1.** Details of physicians, interventions and outcomes in each study. (Continue)

| Study (Year) (Sample size) | Types of Physicians | Comparison | Intervention | Post-intervention follow-up |
|---|---|---|---|---|
| *Medisauskaite et al.* (2019) | Practicing clinicians (n=91) | Waitlist control | Four intervention groups studied stress-related modules: Group 1 focused on stress and burnout, Group 2 on coping with patient death, Group 3 on distress management, and Group 4 covered all three combined. | - |
| *Huang et al.* (2020) (n=36) | 1st-year residents | Waitlist control | 10 Balint sessions and 2 lectures over 6 months focusing on distress management | - |
| *Kesselheim et al.* (2020) (n=100) | Pediatric Hematology-Oncology Fellows | Active control (routine educational practices) | Case-based discussion in clinical practice | 6 months |
| *Lee et al.* (2020) (n=112) | Primary care doctors | Waitlist control | Short-term psychotherapy 9-12 sessions (50 minutes each) based on the coping strategy (Asimov method) over a 1-year period | 12 months |
| *Taylor et al.* (2020) (n=21) | Junior doctors | Active control (group fitness training) | 8 weekly private yoga sessions (1 hour each), including breathing, relaxation, and meditation exercises | - |
| *Fendel et al.* (2021) (n=150) | Residents | Active control (received a coursebook) | German MBSR program: 8 weekly sessions (135 minutes each), 1 full-day silent retreat, and 3 monthly booster sessions | 12 months |
| *West et al.* (2021) (n=123) | Practicing physicians in the Department of Medicine | Waitlist control | 12 group sessions (15 minutes of topic discussion + 45 minutes of shared meal and further discussion) | after 6 months |
| *Lebares et al.* (2021) (n=89) | 1st-year surgical residents | Active control (stress-related group discussion) | Mindfulness meditation skills with 2 groups of interventions<br>-Group1: 8 weekly 2-hour classes<br>-Group2: 6 weekly 90-minute classes with more explicitly applied to real-life situations | 10.5 months for ESRT1<br>6.5 months for ESRT2 |
| *A. Brazier et al.* (2022) (n=153) | Trainee anesthetists | No treatment | Weekly text messages with mindfulness-based and cognitive-behavioral prompts over a 6-month period | - |
| *Fainstad et al.* (2022) (n=101) | Residents | Waitlist control | The “Better Together” curriculum: Web-based group coaching with weekly self-study modules (videos and worksheets) | 6 months |

and personal accomplishment, including those by *Bragard et al.* (2010)[25], *Gunasingam et al.* (2015)[29], *Dyrbye et al.* (2016)[27], *Verweij et al.* (2018)[32], *Dyrbye et al.* (2019)[26], *Medisauskaite et al.* (2019)[24], *Huang et al.* (2020)[30], *Lee et al.* (2020)[31], *West et al.* (2021)[33] and *Fainstad et al.* (2022).[28]

Fig 2.1 shows ten studies in which interventions effectively reduced stress, with a small but statistically significant benefit for the intervention group (Hedges' g = -0.20, 95% CI: -0.37 to -0.02). Stress was assessed using tools such as the Depression, Anxiety, Stress Scale (DASS-21)[36], the Perceived Stress Scale (PSS)[38], and the Perceived Stress Questionnaire (PSQ).[39] *Mache et al.* (2017)[10] reported significant stress reduction among junior gynecologists in German hospitals (Hedges' g = -0.60, 95% CI: -1.11 to -0.08). The three-month intervention consisted of 1.5-hour sessions conducted during off-duty hours to enhance job performance and reduce distress. In contrast, *West et al.* (2014)[11], *Bernburg et al.* (2016)[12], *Axisa et al.* (2019)[13] found no significant effects, while *Ireland et al.* (2017)[14], *Mache et al.* (2018)[15], *Lebares et al.* (2019)[16], *Fendel et al.* (2021)[17], and *Lebares et al.* (2021)[18] supported intervention effectiveness.

Fig 3 highlights eight studies on burnout reduction, using tools such as the Copenhagen Burnout Inventory (CBI)[35], the Professional Quality of Life Scale (ProQOL)[37], and the Maslach Burnout Inventory (MBI).[40] For example, *Ireland et al.* (2017)[19] studied a 10-week mindfulness training program, which significantly reduced burnout (Hedges' g = -0.65, 95% CI: -1.26 to -0.04). Other studies, including *Amutio et al.* (2015)[20], *Lebares et al.* (2019)[16], *Fendel et al.* (2021)[17], and *Brazier et al.* (2022)[21], also

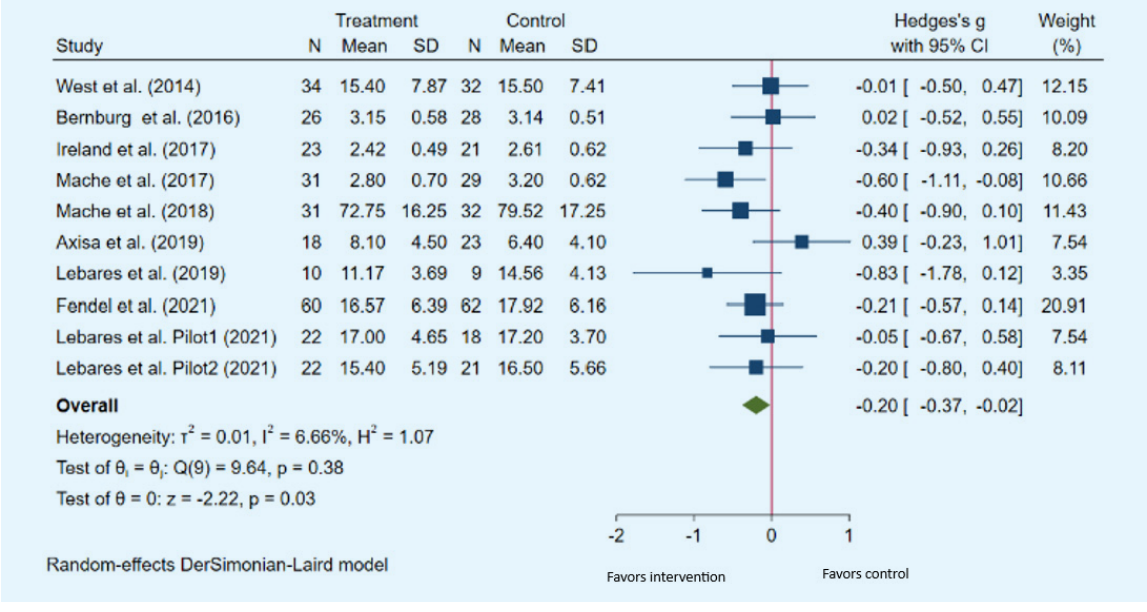



**Fig 2.1.** Final follow-up of stress results.

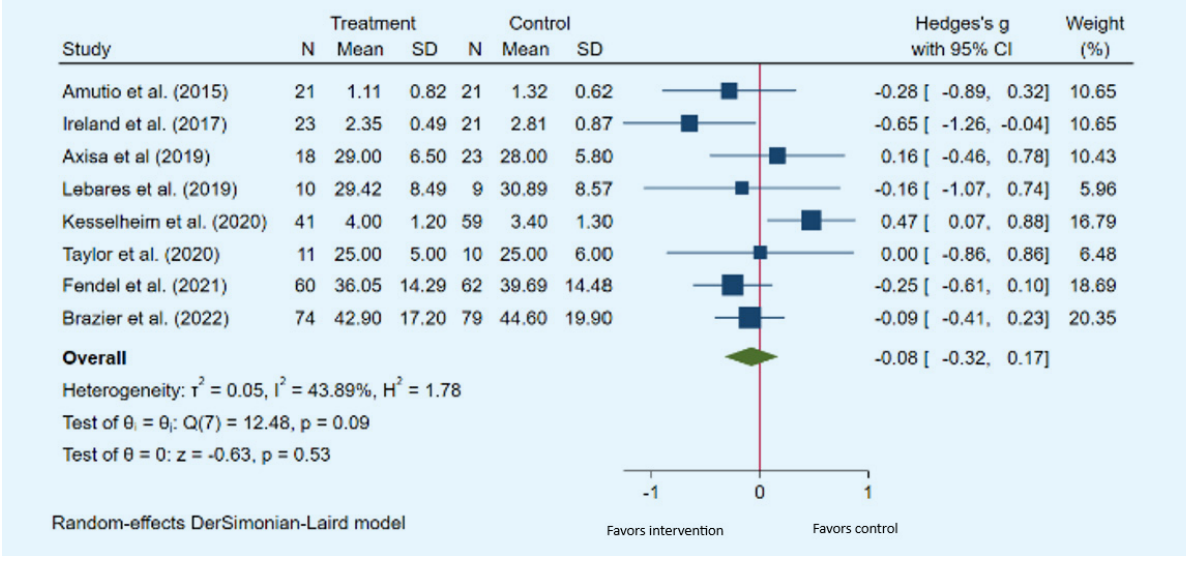



**Fig 2.2.** Final follow-up burnout results.

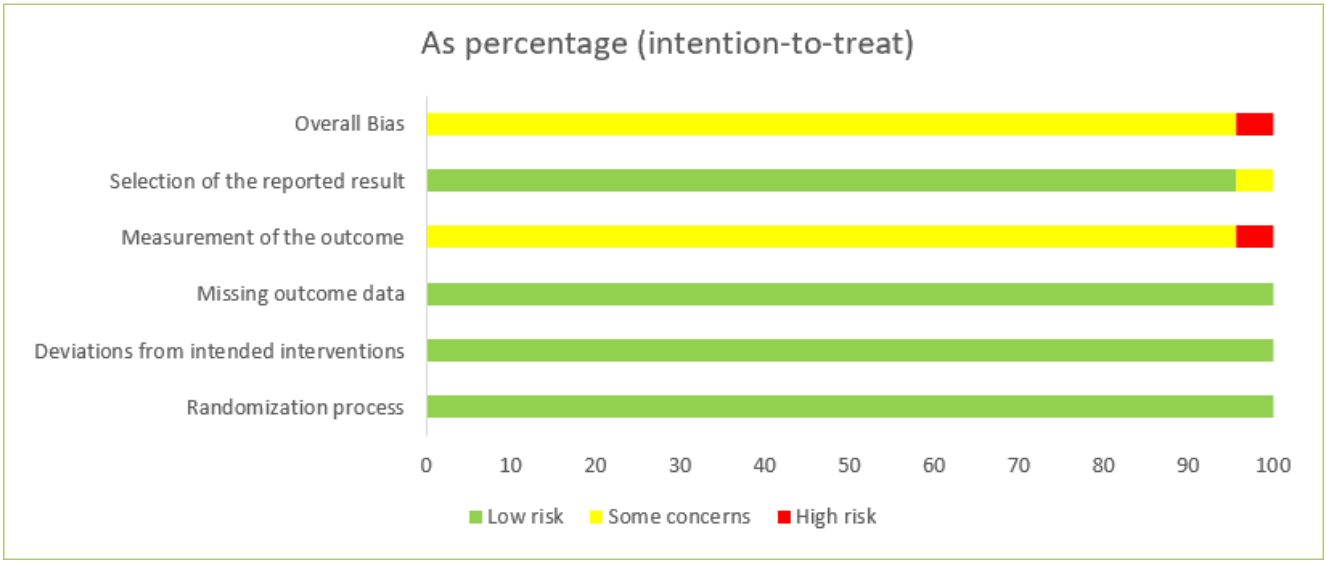



**Fig 3.1.** Risk of Bias Assessment (RoB 2, 2019) of included studies.

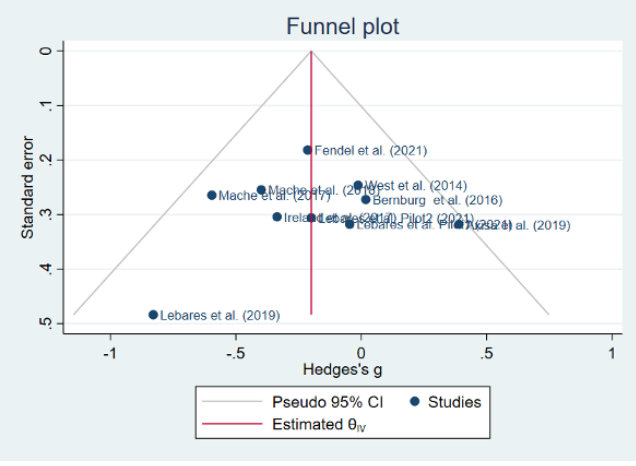



**Fig 3.2.** Funnel plot for final stress analysis

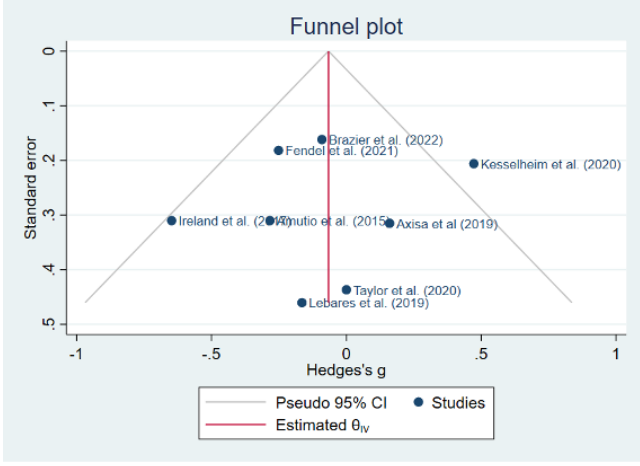



**Fig 3.3.** Funnel plot for final burnout analysis.

supported burnout reduction. However, *Axisa et al.* (2019)[13], *Kesselheim et al.* (2020)[22], and *Taylor et al.* (2020)[23] found no significant effects.

Interventions were assessed at two time points: immediately after the intervention and at final follow-up. Stress-reduction interventions had a statistically significant impact (Fig 2), with Hedges' g = -0.20 (p = 0.03, 95% CI: -0.37 to -0.02) and low heterogeneity ($I^2$ = 6.66%). In contrast, burnout reduction interventions showed a non-significant trend (Fig 3), with Hedges' g = -0.08 (p = 0.53, 95% CI: -0.32 to 0.17) and moderate heterogeneity ($I^2$ = 43.89%), indicating variability in reported effects. Although there was a slight trend toward reduced burnout, the findings were not statistically significant.

**Risk of bias assessment**

Using the Revised Cochrane Risk-of-Bias Tool for Randomized Trials (RoB 2)[41], 22 studies were rated as having some concerns regarding overall risk of bias, mainly due to issues in outcome measurement, as participant awareness of the intervention may have introduced bias in self-reported outcomes. Additionally, one study *Medisauskaite et al.* (2019)[24] showed high risk of bias due to the inclusion of multiple trial groups and incomplete outcome reporting (Fig 3.1). All studies were RCTs and used sufficient randomization methods. Egger's regression test revealed no evidence of publication bias (p=0.69 for the final follow-up stress outcome, p=0.68 for the final follow-up burnout outcome). The funnel plot for stress reduction (Fig 3.2) showed slight asymmetry, suggesting small but significant effects. In contrast, the funnel plot for burnout reduction (Fig 3.3) was symmetrical, indicating limited effectiveness. Studies with larger standard errors such as *Lebares et al.*, 2019[16] reflected smaller sample sizes or higher variability.

## DISCUSSION

This review integrates findings from 13 studies examining interventions aimed at reducing stress and burnout among physicians, highlighting the effectiveness of both mindfulness-based and non-mindfulness-based approaches.

Among the nine studies focused on stress reduction, *Mache et al.* (2017)[10] reported a significant decrease in stress levels among junior gynecologists, indicating that structured training programs can enhance well-being in high-pressure environments. However, other studies, such as *West et al.* (2014)[11], did not demonstrate significant effects, suggesting that variations in intervention design and contextual factors may influence outcomes.

For burnout, six out of eight studies reported favorable outcomes. For example, *Ireland et al.* (2017)[14] demonstrated significant improvements in burnout levels following a 10-week mindfulness training program (Hedges' g = -0.65). However, despite these findings, the overall trend toward decreased burnout was non-significant (Hedges' g = -0.08), underscoring the complexity of burnout, which may require more sustained interventions.

The meta-analysis showed a notable reduction in stress levels (Hedges' g = -0.20, p = 0.03), while burnout interventions showed less consistent results. These findings suggest that stress-reduction strategies are generally effective, but addressing burnout requires more comprehensive long-term approaches.

Future research should prioritize sustained follow-ups and systemic changes to address burnout and promote the well-being of healthcare practitioners. This review has several limitations. First, only English-language studies were included, which may introduce language bias and limit the generalizability of the findings. Second, unpublished studies and additional sources were not explored, which may have affected the comprehensiveness of the evidence. Finally, most included studies relied on self-reported measures, which may be subject to bias due to participants' awareness of their group allocation.

## CONCLUSION

Stress-reduction interventions have been shown to effectively lower stress levels among physicians. However, evidence regarding their impact on burnout remains inconclusive. Differences in study design and outcomes may reduce the strength and generalizability of these findings. Future research should focus on long-term follow-ups, combining systemic changes with individual interventions. A comprehensive approach is essential to address stress and burnout and improve the well-being of healthcare professionals.

### Data Availability Statement

Most of the data used in this study are from published articles. Some data were requested from original authors and can be shared upon request.

## ACKNOWLEDGMENTS

The authors extend their gratitude to Dr. Roongnapa Khampang for her expertise in systematic review and meta-analysis, Dr. Orawan Supapueng for her valuable statistical guidance, and Lakkana Thongchot for her support with research documentation.

## DECLARATIONS

### Grants and Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Conflict of Interest

This research did not receive any specific grant from any public, commercial, or non-profit funding agencies. All authors declare no conflicts of interest regarding the research, authorship, or publication of this manuscript.

### Registration Number of Clinical Trial

Not applicable

### Author Contributions

N.R. and P.P. were responsible for assessing the quality of the selected studies, analyzing the data, and drafting the manuscript. N.R., P.P., and K.W. handled data extraction, with K.W. also involved in the study design and acting as the final decision-maker for resolving disagreements. Every author has reviewed and agreed upon the final draft of the manuscript.

### Use of Artificial Intelligence

The authors used ChatGPT-4o to check and refine the grammar of the manuscript.

## REFERENCES

1. Stress at work - HSE [Internet]. hse.gov.uk. 2018 [cited 30 October 2018]. Available from: http://www.hse.gov.uk/stress/index.htm
2. Michie S. Causes and management of stress at work. Occup Environ Med. 2002;59(1):67–72.
3. Thorsteinsson EB, Brown RF, Richards C. The Relationship between Work-Stress, Psychological Stress and Staff Health and Work Outcomes in Office Workers. Psychology. 2014;05(10):1301–11.
4. Oginska-Bulik N. Occupational stress and its consequences in healthcare professionals: The role of type D personality. Int J Occup Med Environ Health. 2006;19(2):113–22.

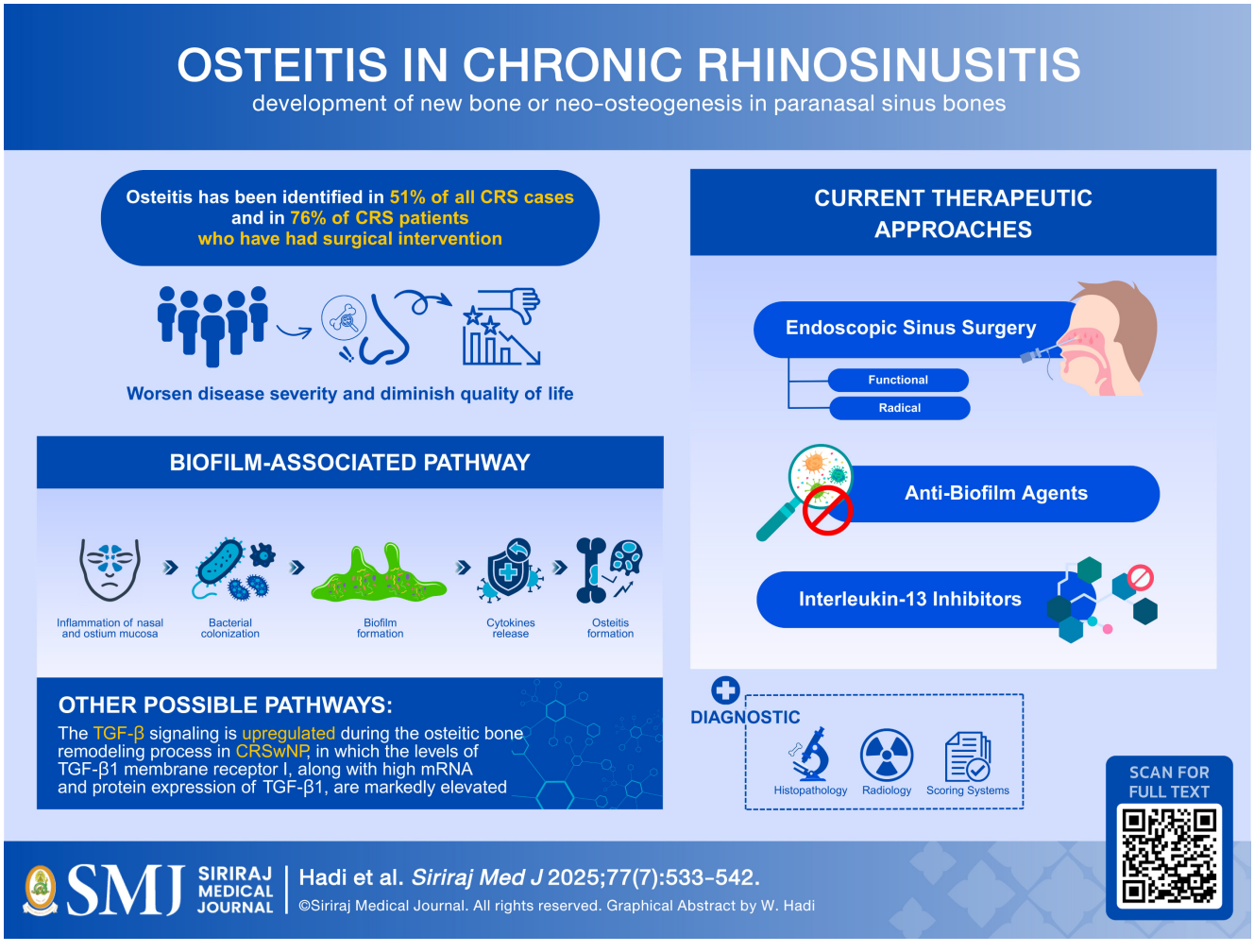
5. Weinberg A, Creed F. Stress and psychiatric disorder in healthcare professionals and hospital staff. Lancet. 2000;355(9203):533–7.
6. Giga SI, Noblet AJ, Faragher B, Cooper CL. The UK perspective: A review of research on organisational stress management interventions. Australian Psychologist. 2003.
7. Ruotsalainen JH, Verbeek JH, Mariné A, Serra C. Preventing occupational stress in healthcare workers. Cochrane Database of Systematic Reviews. 2015, Issue 4.
8. Regehr C, Glancy D, Pitts A, Leblanc VR. Interventions to reduce the consequences of stress in physicians: a review and meta-analysis. The Journal of Nervous and Mental Disease. 2014; 202(5):353–9.
9. Grime PR. Computerized cognitive behavioural therapy at work: a randomized controlled trial in employees with recent stress-related absenteeism. Occup Med (Lond). 2004; 54:353–359.
10. Mache S, Baresi L, Bernburg M, Vitzthum K, Groneberg D. Being prepared to work in Gynecology Medicine: evaluation of an intervention to promote junior gynecologists' professionalism, mental health and job satisfaction. Arch Gynecol Obstet. 2017; 295(1):153-162.
11. West CP, Dyrbye LN, Rabatin JT, Call TG, Davidson JH, Multari A, Romanski SA, Hellyer JM, Sloan JA, Shanafelt TD. Intervention to promote physician well-being, job satisfaction, and professionalism: a randomized clinical trial. JAMA Intern Med. 2014;174(4):527-33.
12. Bernburg M, Baresi L, Groneberg D, Mache S. Does psychosocial competency training for junior physicians working in pediatric medicine improve individual skills and perceived job stress. Eur J Pediatr. 2016;175:1905-12.
13. Axisa C, Nash L, Kelly P, Willcock S. Burnout and distress in Australian physician trainees: Evaluation of a wellbeing workshop. Australas Psychiatry. 2019;27(3):255-61.
14. Ireland MJ, Clough B, Gill K, Langan F, O'Connor A, Spencer L. A randomized controlled trial of mindfulness to reduce stress and burnout among intern medical practitioners. Med Teach. 2017;39(4):409-14.
15. Mache S, Bernburg M, Baresi L, Groneberg D. Mental health promotion for junior physicians working in emergency medicine: evaluation of a pilot study. Eur J Emerg Med. 2018;25(3):191-8.
16. Lebares CC, Guvva EV, Olaru M, Sugrue LP, Staffaroni AM, Delucchi KL, et al. Efficacy of Mindfulness-Based Cognitive Training in Surgery: Additional Analysis of the Mindful Surgeon Pilot Randomized Clinical Trial. JAMA Netw Open. 2019; 2(5):e194108.
17. Fendel JC, Aeschbach VM, Schmidt S, Göritz AS. The impact of a tailored mindfulness-based program for resident physicians on distress and the quality of care: A randomised controlled trial. J Intern Med. 2021;290(6):1233-48.
18. Lebares CC, Coaston TN, Delucchi KL, Guvva EV, Shen WT, Staffaroni AM, et al. Enhanced Stress Resilience Training in Surgeons: Iterative Adaptation and Biopsychosocial Effects in 2 Small Randomized Trials. Ann Surg. 2021;273(3):424-32.
19. Ireland MJ, Clough B, Gill K, Langan F, O'Connor A, Spencer L. A randomized controlled trial of mindfulness to reduce stress and burnout among intern medical practitioners. Med Teach. 2017;39(4):409-14.
20. Amutio A, Martínez-Taboada C, Delgado LC, Hermosilla D, Mozaz MJ. Acceptability and Effectiveness of a Long-Term Educational Intervention to Reduce Physicians' Stress-Related Conditions. J Contin Educ Health Prof. 2015;35(4):255-60.
21. Brazier A, Larson E, Xu Y, Judah G, Egan M, Burd H, Darzi A. 'Dear Doctor': a randomised controlled trial of a text message intervention to reduce burnout in trainee anaesthetists. Anaesthesia. 2022;77(4):405-15.
22. Kesselheim J, Baker JN, Kersun L, Lee-Miller C, Moerdler S, Snaman JM, et al. Humanism and professionalism training for pediatric hematology-oncology fellows: Results of a multicenter randomized trial. Pediatr Blood Cancer. 2020;67(11):e28308.
23. Taylor J, McLean L, Richards B, Glozier N. Personalised yoga for burnout and traumatic stress in junior doctors. Postgrad Med J. 2020;96(1136):349-57.
24. Medisauskaite A, Kamau C. Reducing burnout and anxiety among doctors: Randomized controlled trial. Psychiatry Res. 2019;274: 383-90.
25. Bragard I, Libert Y, Etienne AM, Merckaert I, Delvaux N, Marchal S, et al. Insight on variables leading to burnout in cancer physicians. J Cancer Educ. 2010;25(1):109-15.
26. Dyrbye LN, Shanafelt TD, Gill PR, Satele DV, West CP. Effect of a Professional Coaching Intervention on the Well-being and Distress of Physicians: A Pilot Randomized Clinical Trial. JAMA Intern Med. 2019;179(10):1406-14.
27. Dyrbye LN, West CP, Richards ML, Ross HJ, Satele D, Shanafelt TD. A Randomized, Controlled Study of an Online Intervention to Promote Job Satisfaction and Well-Being among Physicians. Burnout Research. 2016;3(3):69-75.
28. Fainstad T, Mann A, Suresh K, Shah P, Dieujuste N, Thurmon K, et al. Effect of a Novel Online Group-Coaching Program to Reduce Burnout in Female Resident Physicians: A Randomized Clinical Trial. JAMA Netw Open. 2022;5(5):e2210752.
29. Gunasingam N, Burns K, Edwards J, Dinh M, Walton M. Reducing stress and burnout in junior doctors: the impact of debriefing sessions. Postgrad Med J. 2015;91(1074):182-7.
30. Huang L, Harsh J, Cui H, Wu J, Thai J, Zhang X, Cheng L, Wu W. A Randomized Controlled Trial of Balint Groups to Prevent Burnout Among Residents in China. Front Psychiatry. 2020;10:957.
31. Lee S, Rozybakieva Z, Asimov M, Bagiyarova F, Tazhiyeva A, Ussebayeva N, et al. Coping strategy as a way to prevent emotional burnout in primary care doctors: a randomized controlled trial. Arch Balk Med Union. 2020;55(3):398-409.
32. Verweij H, van Ravesteijn H, van Hooff MLM, Lagro-Janssen ALM, Speckens AEM. Mindfulness-Based Stress Reduction for Residents: A Randomized Controlled Trial. J Gen Intern Med. 2018;33(4):429-36.
33. West CP, Dyrbye LN, Satele DV, Shanafelt TD. Colleagues Meeting to Promote and Sustain Satisfaction (COMPASS) Groups for Physician Well-Being: A Randomized Clinical Trial. Mayo Clin Proc. 2021;96(10):2606-14.
34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372:n71.
35. Kristensen TS, Borritz M, Villadsen E, Christensen KB. The Copenhagen burnout inventory: a new tool for the assessment of burnout. Work & Stress. 2005;19(3):192–207.
36. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. 2nd ed. Sydney: Psychology Foundation, 1995.
37. Stamm BH. The Concise ProQOL Manual. 2nd ed. Pocatello, ID: ProQOL.org, 2010.
38. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24(4):385-96.

39. Levenstein S, Prantera C, Varvo V, Scribano ML, Berto E, Luzi C. Development of the Perceived Stress Questionnaire: a new tool for psychosomatic research. J Psychsom Res. 1993;37(1):19–32.
40. Mind Garden website. Available from: http://www.mindgarden.com. Accessed February 18, 2019.
41. Khanngern N, Ratta-apha W, Wannarit K. Burnout among Mental Health Professionals in a Tertiary University Hospital. Siriraj Med J [internet]. 2022 Mar. 1 [cited 2025 Apr. 17];74(3):185-92. available from: https://he02.tci-thaijo.org/index.php/sirirajmedj/article/view/256431
42. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
43. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.3 (updated February 2022). Cochrane; 2022. Available from: https://training.cochrane.org/handbook/current

# Osteitis in Chronic Rhinosinusitis: A State-of-the-Art Review

**Wiyono Hadi, M.D., Ph.D.[1,*], Widia Isa Aprillia Sujana, M.D.[1], Gilang Hadi Fiandani, M.D.[2]**

*[1]Department of Otorhinolaryngology – Head and Neck Surgery, Faculty of Medicine, Universitas Katolik Widya Mandala Surabaya, Surabaya, Indonesia, [2]Department of Otorhinolaryngology – Head and Neck Surgery, Faculty of Medicine, Universitas Airlangga, Dr. Soetomo General Academic Hospital, Surabaya, Indonesia.*




**Corresponding author: Wiyono Hadi*
*E-mail: wiyono@ukwms.ac.id*
*Received 3 March 2025 Revised 17 April 2025 Accepted 20 April 2025*
*ORCID ID:http://orcid.org/0009-0007-4866-1683*
*https://doi.org/10.33192/smj.v77i7.274010*

## ABSTRACT

Chronic rhinosinusitis (CRS) is characterized by inflammation of the nasal mucosa and paranasal sinuses. Mucosal inflammation can induce bone remodeling leading to osteitis. Paranasal sinus osteitis is a relatively new clinical finding, describing changes in the bone due to inflammation that will damage the lamellar bone structure and new bone formation. As many as 36-53% of CRS patients have osteitis, which makes CRS difficult to cure despite adequate management and is related to a substantial number of revision surgeries. Understanding all the evidence regarding osteitis in CRS is critical in discovering effective treatments for this incurable disease. Hence, we summarize the essential yet well-established features of osteitis in CRS, including the updated definition, the role of biofilm formation along with cytokines and runt-related transcription factor 2 axis as its possible underlying pathogenesis, the transforming growth factor beta signaling pathway of the disease, histopathological bone changes, radiographic staging and scoring systems, disease impacts on the CRS severity and quality of life, as well as the most up-to-date treatment strategies for osteitis in CRS. These include functional or radical endoscopic sinus surgery, interleukin-13 inhibitors, anti-biofilm agents, high-dose intranasal corticosteroids, and other potential therapies.


**Keywords:** Bone remodeling; chronic rhinosinusitis; hyperostosis; neo-osteogenesis; osteitis (Siriraj Med J 2025; 77: 533-542)

## INTRODUCTION

Chronic rhinosinusitis (CRS) is a long-term, multifactorial inflammatory disease of the nose and paranasal sinuses, characterized by the presence of at least two such symptoms as: olfactory dysfunction, facial pain, one of which must be anterior/posterior nasal discharge or nasal blockage that lasts for 12 weeks or greater.[1] CRS can be classified into primary and secondary CRS. The primary CRS is divided into two subtypes based on the type of inflammation, consisting of eosinophilic or type 2 and neutrophilic or non-type 2 CRS.[1-3] CRS is also phenotypically divided into CRS with nasal polyps (CRSwNP) and CRS without nasal polyps (CRSsNP).[1-3] CRS prevalence ranged from 3% to 8.7% of the global population.[4,5]

Experiences of persistent or even deteriorating symptoms of CRS after having adequate management have become a global concern. Moreover, there are cases where the symptoms fail to improve after sinus surgery.[6,7] These conditions are called recalcitrant or refractory CRS (RCRS).[8] Osteitis, the involvement of the newly developed paranasal sinus bones from the inflammatory process in CRS, had been considered the critical cause of RCRS and the high rate of revision surgeries.[9] The presence of osteitis may also be associated with a high risk of increasing intraoperative bleeding due to the challenge of removing the thick bone. Osteitis is also associated with CRS severity.[10] Osteitis has been identified in 51% of all CRS cases and 76% of patients with CRS who have had surgical intervention.[11]

Osteitis in CRS, together with its biomolecular mechanism, remains poorly understood owing to the wide variety of statements in the literature. Its existence is also often unrecognized by clinicians. In this review, we summarize the updated definition, the role of biofilm formation along with cytokines and runt-related transcription factor 2 (RUNX2) axis as its possible underlying pathogenesis, the transforming growth factor beta (TGF-β) signaling pathway of the disease, histopathological bone changes, radiographic staging and scoring systems, disease impacts on the CRS severity and quality of life (QoL), as well as the most up-to-date treatment strategies for osteitis in CRS.

### Definition

Osteitis is the accepted term for the bone thickening of paranasal sinuses based on computed tomography (CT) because the underlying sinuses bones, except the frontal sinuses, are void of bone marrow.[12] Osteitis is first described as the inflammation of bone that lacks marrow space.[13] An inflammatory involvement and the expansion of the Haversian canal system have been demonstrated in animal studies.[14-16] Since then, several human studies have contributed further to define osteitis better.

Some human histopathological studies have revealed no involvement of inflammatory cytokines in the osteitic bone.[9,17,18] In contrast, other newest study demonstrated the eosinophils involvement in human osteitic bone and reported the different roles of this signaling pathway between CRSwNP and CRSsNP.[19] Recently, there was also a reported association between IL-13 and osteitis in type 2 CRS.[20] These studies suggest the involvement of inflammatory agents in the underlying sinus bones and indicate that the osteitis formation varies depending on the CRS classification.

Eventually, based on recent evidence, osteitis in CRS has been temporarily defined as the development of new bone or neo-osteogenesis and bone remodeling in paranasal sinus bones. Further larger and specific studies need to be done to reach a consensus due to the controversy of whether: Osteitis is a form of isolated new bone formation resulting from the locally infected sinus mucosa rather than direct inflammation of underlying sinus bone; or the inflammation process in the underlying sinus bone serves as a storage for inflammatory cytokines, resulting from an expansion of the Haversian canal system and the infiltration of inflammatory cells into an augmented vascular network.

## Epidemiology

A study by Lee et al.[9] in 2006, with a sample of 121 CRS patients, showed that 36-53% of CRS patients were found to have osteitis using radiological and histopathological criteria. Mild osteitis was found in 73% of samples, moderate osteitis in 45%, and severe osteitis in 18% of samples. The incidence of osteitis was also shown to increase with increasing Lund-Mackay (LM) score. Georgalas et al.[7] in 2010 also stated that the prevalence of bone involvement in CRS cases was 33% in the group that had never undergone surgery, but increased to 75% in CRS patients with a history of previous surgery. Hyperostosis and heterogeneity of the sinus wall were seen in 63.7% of CRS patients and 10% in the control group, but those who met the significant criteria of the Global Osteitis Scoring Scale (GOSS), which scores more than 5, was 40% in CRS group and 0 in the control group. CRSwNP were also seen to have a greater percentage of osteitis than CRSsNP (73% vs 55%, $p = 0.04$).[7]

## Pathogenesis and molecular mechanism

### *Role of biofilm formation, cytokines, and RUNX2 Axis*

The endotypes of CRS were classified based on the pattern of cytokine production from regulatory immune cells. Eosinophilic CRS, referred to as type 2 CRS, consists of several interleukins (ILs), including IL-4, IL-5, IL-9, and IL-13 cytokines regulated by Th2 cells, mast cells, group 2 innate lymphoid cells (ILC2s), and eosinophils. Non-eosinophilic or neutrophilic CRS consists of type 1 and type 3 CRS. Type 1 CRS consists of interferon-gamma (IFN-γ) regulated by Th1 cells, cytotoxic T cells (CD8+ T), natural killer (NK) cells, and group 1 innate lymphoid cells (ILC1s). Type 3 CRS consists of IL-17 and IL-22, regulated by Th17 cells and group 3 innate lymphoid cells (ILC3s).[1-3]

Various variables, including mucociliary clearance dysfunction, dysbiosis of sinus microbiota, epithelial barrier impairment, and immune system abnormalities, may facilitate the invasion of pathogens, hence fostering infection.[21-23] Infection induces an inflammatory response of the sinus mucosa, leading to the narrowing or complete obstruction of the sinus entry. Obstruction of the ostium reduces oxygen tension within the sinus, causing sinusitis.[24] Ostium sinus obstruction also enhances the trapping and buildup of mucus secretions from goblet cells. Interestingly, certain bacteria, particularly *Pseudomonas aeruginosa*, have special preference for living in mucus.[25] Retention of mucus together with the bacteria will trigger inflammatory reactions, thereby enhancing the bacterial activity within the sinus. Eight mucin genes have been detected in CRS tissues, including two membrane-bound mucins called MUC1 and MUC4, and six secreted mucins called MUC2, MUC5AC, MUC5B, MUC6, MUC7, and MUC8. Research have established a positive association between IL-13 and MUC5AC expression, showing that IL-13 can directly regulate mucin hypersecretion in a type 2 CRS.[26]

An environment with a crowded bacterial population in the paranasal sinuses encourages a shift in bacterial states from a planktonic form (free-living bacteria) to a sessile form (attached-living bacteria), which supports the process of biofilm formation.[27] The primary biofilm formation process occurs in several stages, including attachment, microcolony formation, matrix production, and detachment.[27,28] The extracellular polymeric substance (EPS) matrix in biofilm causes bacteria to become resistant to antibiotics and protects them from the host's immune system.[27-29] These cause the exhibit of the biofilm life cycle, thus promoting more bacterial colonies to grow and ultimately worsening the CRS. Two major bacteria that often play a role in biofilm formation in CRS are *Staphylococcus aureus* and *P. Aeruginosa*.[30-32]

High levels of pro-inflammatory cytokines that accumulate due to biofilm-infected sinus promote its infiltration process into the bone around. The interaction of these cytokines in the bone affects the remodeling process.[33] Bone remodeling is a physiological process that equilibrates the activity of osteoclasts and osteoblasts. In chronic or persistent inflammation, such as in CRS, there exists an imbalance between osteogenesis by osteoblasts and bone resorption by osteoclasts.[33]

IL-13 releases, particularly in eosinophilic CRS, elevates the expression of runt-related transcription factor 2 (RUNX2), which enhances osteoblast proliferation and activation, thereby augmenting the neo-osteogenesis process.[34] IL-13 also inhibits the synthesis of cyclooxygenase-2 (COX-2)-dependent prostaglandins (PG), thereby

reducing osteoclast production and its function in the bone resorption process. Increased osteoblast activity and decreased osteoclast production cause bone thickening, a typical sign of osteitis in CRS.[35] IL-13 and IL-4 in cultured human osteoblast (hOBs) have also found to have a regulation effect in osteoblasts activity by increasing the osteoblast differentiation, collagen secretion, alkaline phosphatase (ALP) expression, and mineralization, thereby increasing neo-osteogenesis.[36] However, further studies analyzed the role of IL-13 and other cytokines on osteitis formation in CRS with a large population, and multi-centered research, including assessing all CRS phenotypes, need to be investigated to avoid potential biases.

Research revealed that of all the cytokines produced by each endotype in CRS, only IL-4, IL-13, and IL-17 play an essential role in osteitis in CRS through the mechanism of neo-osteogenesis.[37] However, research has shown that IL-17 also induces the receptor activator of nuclear factor kappa-B ligand (RANKL) in osteoblasts. This results in inflammation and bone destruction by activating osteoclasts in other chronic inflammatory diseases, such as rheumatoid arthritis. This finding resulted in a mixture of mechanisms between neo-osteogenesis and osteolysis in non-eosinophilic CRS.[37,38] Therefore, it also explains why the prevalence of neo-osteogenesis is higher in eosinophilic CRS than in non-eosinophilic.[20]

### *Role of TGF-β signaling pathway*

Transforming growth factor beta (TGF-β) is a multifunctional cytokine with three isoforms: TGF-β1, TGF-β2, and TGF-β3, which interact with membrane R types I, II, and III.[39] The TGF-β1 is involved in airway pathologies and plays a different role in mucosal and bone tissues in CRS patients.[40] TGF-β1 in mucosal tissue is elevated in CRSsNP but diminished in CRSwNP.[41,42] In contrast, TGF-β1 levels were elevated in CRSwNP bone specimens and exhibited a correlation with osteitis bone grading when compared to CRSsNP, as determined by reverse transcription polymerase chain reaction (RT-PCR) and immunohistochemistry (IHC) staining.[19]

Previously in 2003, Park et al.[43] reported that TGF-β could significantly reduce COX-2 mRNA levels. Moreover, Takai et al.[44] in 2013 demonstrated that TGF-β1 induced downregulation of COX-2, leading to decreased synthesis of PGE2. These results indicate that the TGF-β signaling pathway is upregulated during the osteitic bone remodeling process in CRSwNP. Wang et al.[19] in 2015 noted markedly elevated levels of TGF-β1 RI together with high messenger ribonucleic acid (mRNA) and protein expression of TGF-β1, in the osteitic ethmoid bone of CRSwNP patients relative to control and CRSsNP groups. TGF-β1 expression was observed in eosinophils and osteoblasts of the CRSwNP group, but was infrequently detected in the control and CRSsNP groups. However, this preliminary study needs further research with larger sample sizes and specific analyses. The summary of the pathogenesis of osteitis in CRS can be seen in Fig 1.

## Diagnosis

### *Histopathological bone changes*

Histology is an examination that is considered to be the most accurate diagnosing method for a disease accompanied by structural changes such as osteitis. The histopathological findings of osteitis in CRS include

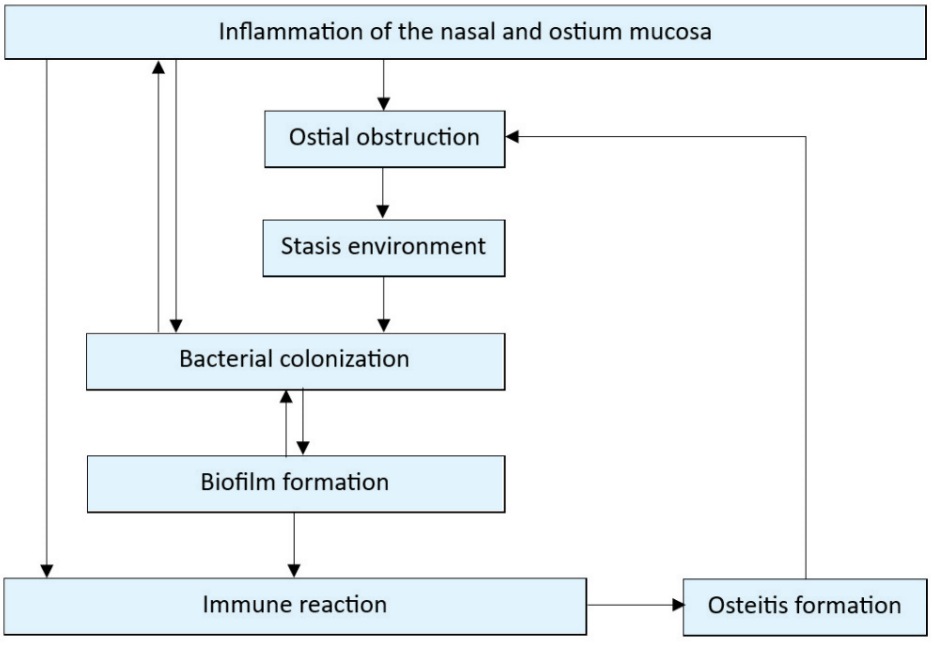



**Fig 1.** Summary of the pathogenesis of osteitis in chronic rhinosinusitis associated with biofilm formation and immune reaction.

thickening of the periosteum, increased activity of osteoblast-osteoclast characterized by the presence of osteoblast or multinucleated osteoclast, dominated woven bone formation of total bone thickness, and fibrosis.[45] Newly formed bone in osteitis is a form of immature woven bone with a weaving arrangement of collagen fibers, making it more flexible and weaker than the mature lamellar bone with strong structured collagen fibers.[46]

### *Radiographic staging systems*

Although histological examination provides a precise way to diagnose osteitis, radiological investigations are generally performed to diagnose osteitis in CRS. The neo-osteogenesis criteria for osteitis in CRS are bone thickening of more than 3 mm with a density of more than 500 Hounsfield units (HU) in CT measures (Table 1).[9,45,47-50] A study by Lee et al. in 2006 reported that CT imaging showed neo-osteogenesis in 36% of patients while 53% showed histopathological evidence of osteitis of bone specimens.[9] An example of a CT image of osteitis in a patient with CRS can be seen in Fig 2.

**TABLE 1.** Osteitis severity based on the bone thickness of computed tomography measurement.

| Osteitis Severity | Bone Thickness (mm) |
|---|---|
| Not Significant | <3 |
| Mild | 3 |
| Moderate | 4-5 |
| Severe | >5 |

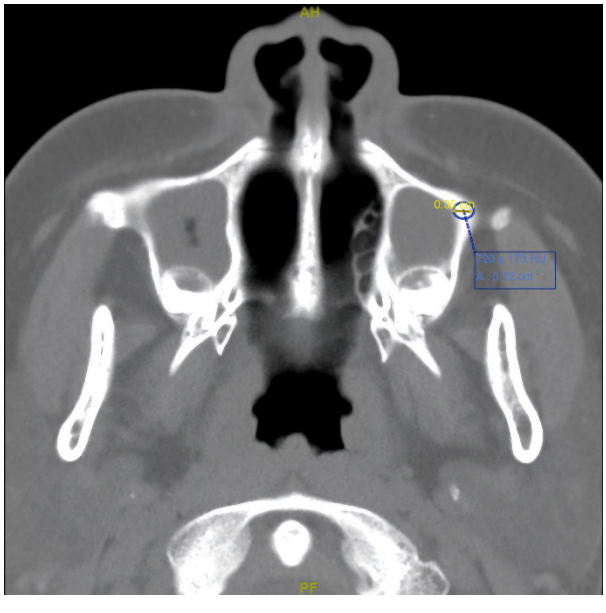



**Fig 2.** A computed-tomography image of osteitis in a chronic rhinosinusitis patient showed 3.5 mm bone thickening with a density of 720±173 Hounsfield units around the maxillary sinus.

Single-photon emission computed tomography (SPECT) is a nuclear medicine imaging technology that employs the intravenous administration of radioactive isotopes prior to imaging.[51] SPECT is a more sensitive method for detecting acute or mild bone remodeling; in contrast, CT observed only a significant thickening of bone density that is usually considered chronic or severe.[51] SPECT may detect milder instances of osteitis, hence averting revision procedures resulting from the ongoing disease progression.[49] Nonetheless, compared to CT, SPECT is probably impractical for a regular application owing to its elevated cost, extended procedural duration, and increased invasiveness resulting from exposure to radioactive isotopes or contrast agents.[51,53] The SPECT imaging had a specificity of 66.7% and positive predictive value of 96.9% in detecting osteitis in CRS, compared to histopathologic bone changes.[53]

### *Scoring systems*

Currently, there are widely accepted and agreed-upon osteitis assessment criteria. Some of them are Kennedy Osteitis Score (KOS), GOSS, and Modified GOSS. Original KOS classified the osteitis as mild, moderate, or severe depending on the extent of bony thickness on all sinuses except the frontal sinuses. GOSS assesses ten sinuses with an addition of the percentage of sinus wall involvement, while Modified GOSS only assesses four sinuses.[54]

The severity of osteitis can be graded using the GOSS, the current most validated radiological grading system based on CT imaging.[7,11] GOSS is the best scoring system in detecting extensive bone changes in the paranasal sinus walls, because in addition to severity, the percentage of bone involvement is also measured. When using GOSS, clinicians must evaluate for loss of bone definition, hyperostosis, new bone formation, or signs of heterogeneity in each paranasal sinus wall.[54] The maximum thickness area of the osteitic bone of each sinus is measured and then assigned as grading from 0 to 5, as seen in Table 2. The possible range of GOSS is between 0 and 50, based on the total or global osteitis score of all ten sinuses.[7] Higher grades indicate greater osteitis severity.

## Disease impacts

### *CRS severity*

Osteitis grade has proven to have a proportional correlation to the increased CRS severity.[55] The LM scoring system is a diagnostic instrument utilized to

**TABLE 2.** Global osteitis scoring scale.

| Osteitis Grade | Involvement of Sinus Wall (%) | Bone Thickness (mm) |
|---|---|---|
| 1 | <50 | <3 |
| 2 | <50 | 3-5 |
| 3 | <50 | >5 |
| 3 | >50 | <3 |
| 4 | >50 | 3-5 |
| 5 | >50 | >5 |

Total score interpretation: <5 = not significant; 5-20 = mild; 20-35 = moderate; >35 = severe

evaluate the severity of CRS.[56] Higher scores indicate greater CRS severity. Studies have demonstrated the association between the LM scoring system and the GOSS score.[9,20,55,57,58] CRS patients with osteitis have been reported to have worse baseline disease severity compared to CRS patients without osteitis, as measured by several diagnostic examinations, including endoscopy, CT imaging, or olfactory function assessment.[55]

### *Quality of life*

An earlier study reported that osteitis significantly affected the QoL along with its severity.[59] The most widely accepted subjective assessment for assessing QoL in CRS patients is the Sinonasal Outcome Test-22 (SNOT-22), a derivative of the SNOT-16 and SNOT-20.[60] Unlike the previous versions, the symptoms of nasal obstruction and loss of smell and/or taste have been added in SNOT-22, with possible total scores between 0 and 110.[61] However, a previous study reported no association between osteitis and the QoL.[55] Potential confounding factors, such as CRS phenotype or comorbidities, should be of greater concern to researchers to obtain more accurate conclusions regarding the relationship between osteitis and QoL.

## Treatment strategies

### *Surgical techniques*

Surgical treatment under endoscopic sinus surgery (ESS) has become the standard treatment for treating osteitis in CRS, with experts recommending removing osteitic bone wherever feasible.[62] The two best-known surgical techniques for osteitis are functional ESS (FESS) and radical ESS (RESS). FESS focused on removing any causes of sinus obstruction, including the osteitic bone, thus improving the ventilation and function of the sinus.[63] RESS focused on resecting all walls between the nasal fossa and paranasal sinuses, resulting in a singular extensive cavity.[64,65] Although RESS is not widely used in clinical practice due to the possible physiological function of sinus impairment, a study conducted by Wang et al. in 2021 has reported the effectiveness of RESS compared to FESS, in which RESS provided a lower LM score than in the FESS group one year after surgery.[65] In contrast, a recent study conducted by Aldajani et al. in 2024 has reported no significant differences between RESS and Primary Endoscopic Sinus Surgery (PESS), a comprehensive full-house FESS, in 12 months follow-up period.[66] These results indicated that a longer follow-up duration and a better study design needs to be conducted to determine the effectiveness of RESS and FESS for osteitis management in CRS.

### *Potential biologic agents*

IL-13 is a natural pro-inflammatory cytokine highly secreted in eosinophilic CRS, including CRSwNP or CRSsNP, but especially in CRSwNP.[67] IL-13 in CRS binds to the IL-13 alpha 2 receptor (IL-13Rα2) in sinonasal epithelial cells.[68] Research has shown that IL-13Rα2 receptors in CRS patients with osteitis are higher than in healthy patients.[37] It implied that IL-13 plays a significant role in osteitis formation in CRS. In addition, it also can be hypothesized that eosinophilic CRS patients with osteitis are more likely to experience RCRS caused by the high level of IL-13 and the large number of IL-13Rα2 receptors, which support the severity of osteitis in CRS. Therefore, inhibition of IL-13 is one of the effective strategies that can be explored and developed to prevent the occurrence of osteitis in eosinophilic CRS, which can ultimately reduce the incidence of RCRS. Khalmuratova et al.[67] in 2020 proved that resveratrol suppressed the expression of IL-13, RUNX2, and osteoblasts in a mouse model of

CRS induced by 3% ovalbumin and 10 ng Staphylococcal enterotoxin B (OVA/SEB). Further studies or clinical trials examining the efficacy of IL-13 inhibitors in CRS patients with osteitis are awaited.

### *Additional possible therapies*

Mukerji et al.[69] in 2009 demonstrated a placebo-controlled trial in 77 CRS patients receiving *Lactobacillus rhamnosus* or placebo orally for four weeks. Results revealed that *L. rhamnosus* is an ineffective adjuvant therapy for CRS, as there were no significant differences in SNOT-20 scores between the *L. rhamnosus* group and the placebo group. Other research, such as the application of baby shampoo irrigation, has shown considerable enhancement in CRS patients following FESS. However, it also resulted in several intolerable side effects, including nasal burning and headaches, making baby shampoo a less feasible treatment option for CRS.[70]

Alandejani et al.[71] in 2009 reported the significant bactericidal effect of Manuka honey against planktonic and biofilm forms of *S. aureus* and *P. aeruginosa*. Methylglyoxal (MGO) is the main active ingredient found in Manuka honey. MGO has been reported to be able to reduce biofilm biomass at 1.8 mg/mL and 3.6 mg/mL in a sheep model for CRS.[72] However, animal studies demonstrated that MGO given alone had more toxic effects, such as metaplasia of respiratory epithelium and severe sinus inflammation, compared with MGO given with Manuka honey. This result suggests that MGO does not have its effects alone and requires the other components in Manuka honey to provide significant advantages.[71,72]

Jervis-Bardy et al.[73] in 2012 performed a randomized controlled clinical trial involving 25 RCRS patients following FESS. Patients were randomized to receive twice-daily nasal rinses for one-month with either 0.05% Mupirocin or oral Augmentin. After one month, negative culture results were observed in 89% of the mupirocin group, whereas the placebo group exhibited 0% findings. However, re-evaluation in the mupirocin cohort at two to six months revealed positive culture results for *S. aureus* in 83.3% of individuals.

Kurasirikul et al.[74] in 2014 performed a retrospective cohort study of 38 children with CRS given gentamicin nasal irrigation. The results showed significant improvements in several parameters, such as nasal congestion, rhinorrhea, itching, post-nasal drip, purulent nasal discharge, halitosis, chronic cough, sneezing, and anosmia. However, further longer follow-up and RCT studies should be performed to better assess the efficacy and toxicity.

Several researchers have reported the concept of biofilm-induced CRS.[75-77] Previous human studies reported the correlation between the volume of bacterial biofilm and the severity of osteitis in CRS.[78] Hadi et al.[79] in 2019 revealed antimicrobial and antibiofilm effects of the Indonesian ethanolic extract of propolis *Trigona spp*. (IEEP-TRI) against *S. aureus* isolated from CRS patients using two-dimensional confocal laser scanning microscopy (CLSM). IEEP-TRI also showed a reduction effect in autoinducer peptides (AIP), the quorum-sensing molecules secreted by *S. aureus* in the biofilm-associated CRS mechanism.

High doses of intranasal corticosteroids (INCS), including Fluticasone 400 µg/200 µL, Budesonide 750-2000 µg/200 µL, and Mometasone 200-400 µg/200 µL, had been shown to have a direct in vitro effect in reducing biofilm biomass by up to 99%.[80] These findings elucidate a novel mechanism of action of INCS against biofilm in CRS, thus preventing the formation of osteitis and necessitating further investigation.

## CONCLUSION

Osteitis is characterized by neo-osteogenesis of woven bone around the infected paranasal sinus and is regarded as a complex condition related to CRS severity and QoL. Osteitis in CRS is linked to biofilm formation, inflammatory cytokine presence, RUNX2 regulation, and TGF-β signaling. Diagnostic tools include both histopathological and radiographic staging and scoring systems. Therapy involves surgical interventions such as FESS or RESS. IL-13 inhibitors, anti-biofilm agents, and high-dose INCS present possible mechanisms of action against osteitis in CRS. Thus, future studies necessitate a better understanding of the inflammatory involvement in the underlying sinus bones of CRS, alongside additional investigations into its pathogenesis and molecular mechanisms in human clinical trial studies in large sample sizes, so that it may aid the discovery of novel therapeutic agents for the disease. Future long-term clinical efficacy studies with well-designed randomized controlled trials (RCTs) are also needed to assess their sustained benefits and safety.

### Data Availability Statement

The data underlying this study are available in the published article.

## ACKNOWLEDGEMENT

None.

## DECLARATION

### Grants and Funding Information

This study received no external funding.

## Conflict of Interest

The authors report no conflict of interest in this study.

## Registration Number of Clinical Trial

Not applicable.

## Author Contributions

Conceptualization and methodology: W.H., W.I.A.S., and G.H.F. ; Data acquisition and writing – review & editing: W.H. and G.H.F. ; Writing – original draft: W.H. and W.I.A.S.

## Use of Artificial Intelligence

This study does not use any artificial intelligence assistant.

## REFERENCES


1. Fokkens WJ, Lund VJ, Hopkins C, Hellings PW, Kern R, Reitsma S, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. Rhinology. 2020;58(Suppl S29):1-464.
2. Kato A, Peters AT, Stevens WW, Schleimer RP, Tan BK, Kern RC. Endotypes of chronic rhinosinusitis: Relationships to disease phenotypes, pathogenesis, clinical findings, and treatment approaches. Allergy. 2022;77(3):812-26.
3. Kato A, Schleimer RP, Bleier BS. Mechanisms and pathogenesis of chronic rhinosinusitis. J Allergy Clin Immunol. 2022;149(5): 1491-1503.
4. Dietz de Loos D, Lourijsen ES, Wildeman MAM, Freling NJM, Wolvers MDJ, Reitsma S, et al. Prevalence of chronic rhinosinusitis in the general population based on sinus radiology and symptomatology. J Allergy Clin Immunol. 2019;143(3):1207-14.
5. Min HK, Lee S, Kim S, Son Y, Park J, Kim HJ, et al. Global incidence and prevalence of chronic rhinosinusitis: A systematic review. Clin Exp Allergy. 2025;55(1):52-66.
6. van der Veen J, Seys SF, Timmermans M, Levie P, Levie P, Jorissen M, Fokkens WJ, et al. Real-life study showing uncontrolled rhinosinusitis after sinus surgery in a tertiary referral centre. Allergy. 2017;72(2):282-90.
7. Georgalas C, Videler W, Freling N, Fokkens W. Global Osteitis Scoring Scale and chronic rhinosinusitis: A marker of revision surgery. Clin Otolaryngol. 2010;35(6):455-61.
8. Kong IG, Kim DW. Pathogenesis of recalcitrant chronic rhinosinusitis: The emerging role of innate immune cells. Immune Netw. 2018;18(2):e6.
9. Lee JT, Kennedy DW, Palmer JN, Feldman M, Chiu AG. The incidence of concurrent osteitis in patients with chronic rhinosinusitis: A clinicopathological study. Am J Rhinol. 2006; 20(3):278-82.
10. Israchmadi A, Ningrum FH, Baskoro N, Mailasari A. Correlation between the severity of chronic rhinosinusitis and the degree of osteitis based on computerized tomography evaluation. Medica Hospitalia J Clin Med. 2024;11(2):209-13.
11. Snidvongs K, McLachlan R, Chin D, Pratt E, Sacks R, Earls P, et al. Osteitic bone: A surrogate marker of eosinophilia in chronic rhinosinusitis. Rhinology. 2012;50(3):299-305.
12. Saladin KS. Anatomy & physiology: The unity of form and function. New York, NY: McGraw-Hill Companies; 2007.
13. Kennedy DW, Senior BA, Gannon FH, Montone KT, Hwang P, Lanza DC. Histology and histomorphometry of ethmoid bone in chronic rhinosinusitis. Laryngoscope. 1998;108(4 Pt 1): 502-7.
14. Westrin KM, Norlander T, Stierna P, Carlsöö B, Nord CE. Experimental maxillary sinusitis induced by Bacteroides fragilis: A bacteriological and histological study in rabbits. Acta Otolaryngol. 1992;112(1):107-14.
15. Perloff JR, Gannon FH, Bolger WE, Montone KT, Orlandi R, Kennedy DW. Bone involvement in sinusitis: An apparent pathway for the spread of disease. Laryngoscope. 2000;110(12): 2095-9.
16. Antunes MB, Feldman MD, Cohen NA, Chiu AG. Dose-dependent effects of topical tobramycin in an animal model of Pseudomonas sinusitis. Am J Rhinol. 2007;21(4):423-7.
17. Cho SH, Min HJ, Han HX, Paik SS, Kim KR. CT analysis and histopathology of bone remodeling in patients with chronic rhinosinusitis. Otolaryngol Head Neck Surg. 2006;135(3):404-8.
18. Snidvongs K, Earls P, Dalgorf D, Sacks R, Pratt E, Harvey RJ. Osteitis is a misnomer: A histopathology study in primary chronic rhinosinusitis. Int Forum Allergy Rhinol. 2014;4(5): 390-6.
19. Wang M, Ye T, Liang N, Huang Z, Cui S, Li Y, et al. Differing roles for TGF-β/Smad signaling in osteitis in chronic rhinosinusitis with and without nasal polyps. Am J Rhinol Allergy. 2015;29(5): e152-9.
20. Zhang Z, Zuo Q, Du Y, Jiang H, Ma F, Zhang Y. Correlation of eosinophils and type 2 inflammatory mediators with osteitis in chronic rhinosinusitis with nasal polyps. J Inflamm Res. 2024; 17:4055-64.
21. Jiao J, Wang C, Zhang L. Epithelial physical barrier defects in chronic rhinosinusitis. Expert Rev Clin Immunol. 2019;15(6): 679-88.
22. Cho DY, Hunter RC, Ramakrishnan VR. The microbiome and chronic rhinosinusitis. Immunol Allergy Clin North Am. 2020;40(2):251-63.
23. He Y, Fu Y, Wu Y, Zhu T, Li H. Pathogenesis and treatment of chronic rhinosinusitis from the perspective of sinonasal epithelial dysfunction. Front Med (Lausanne). 2023;10:1139240.
24. Kim YJ, Cho HJ, Shin WC, Song HA, Yoon JH, Kim CH. Hypoxia-mediated mechanism of MUC5AC production in human nasal epithelia and its implication in rhinosinusitis. PLoS One. 2014;9(5):e98136.
25. Hansson GC. Mucins and the microbiome. Annu Rev Biochem. 2020;89:769-93.
26. Zhai Z, Shao L, Lu Z, Yang Y, Wang J, Liu Z, et al. Characteristics of mucin hypersecretion in different inflammatory patterns based on endotypes of chronic rhinosinusitis. Clin Transl Allergy. 2024;14(1):e12334.
27. Rice DH. Biofilm in chronic rhinosinusitis: A review. Med Res Arch. 2020;8(2).
28. Sauer K, Stoodley P, Goeres DM, Hall-Stoodley L, Burmølle M, Stewart PS, et al. The biofilm life cycle: Expanding the conceptual model of biofilm formation. Nat Rev Microbiol. 2022;20(10): 608-20.
29. Maina IW, Patel NN, Cohen NA. Understanding the role of biofilms and superantigens in chronic rhinosinusitis. Curr Otorhinolaryngol Rep. 2018;6(3):253-62.

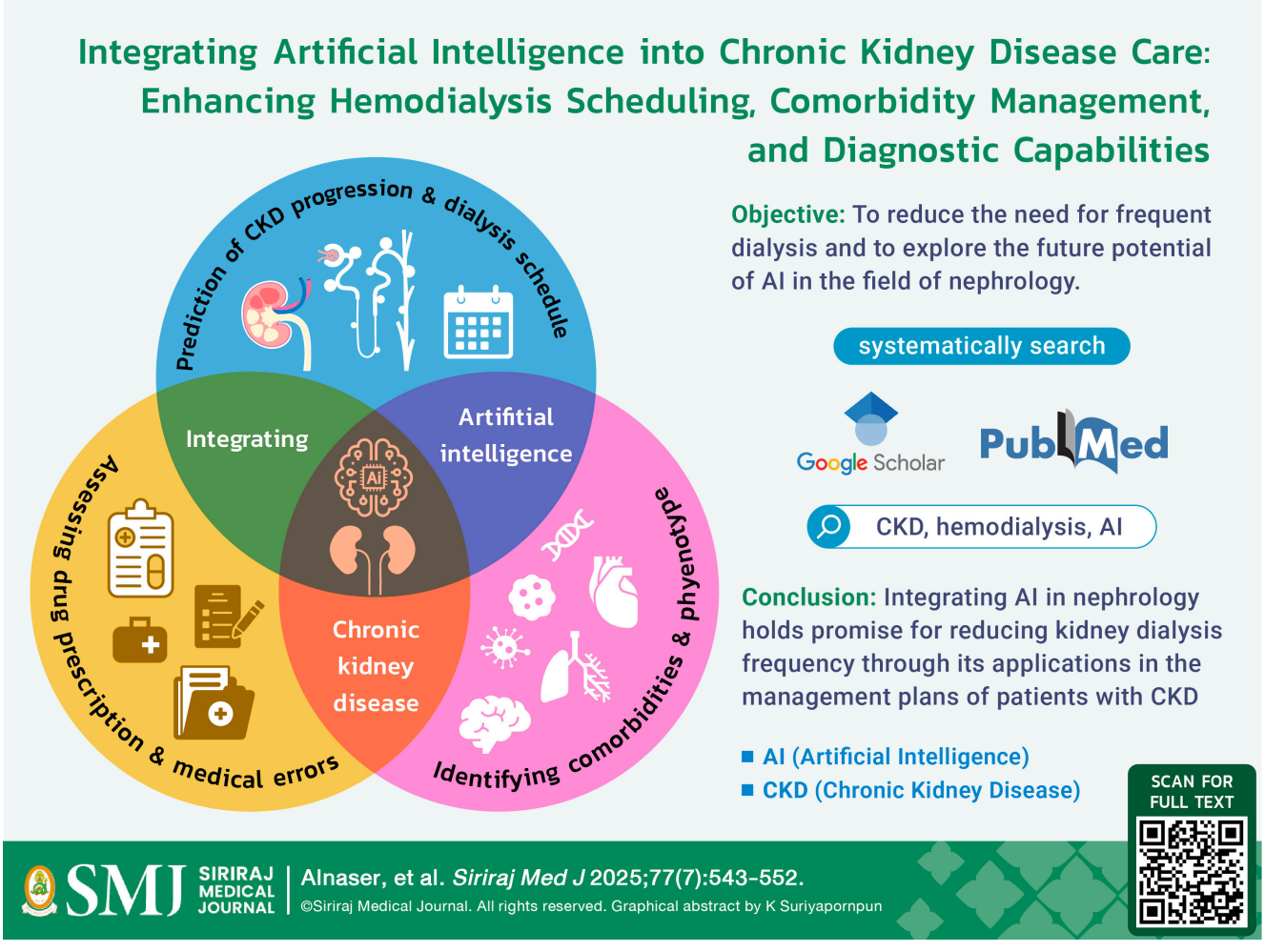
30. Chegini Z, Noei M, Hemmati J, Arabestani MR, Shariati A. The destruction of mucosal barriers, epithelial remodeling, and impaired mucociliary clearance: Possible pathogenic mechanisms of Pseudomonas aeruginosa and Staphylococcus aureus in chronic rhinosinusitis. Cell Commun Signal. 2023;21(1):306.
31. Tantilipikorn P, Bunnag C, Srifuengfung S, Dhiraputra C, Tiensasitorn C, Jaroencharsri P, et al. A surveillance study of bacteriologic profile in rhinosinusitis. Siriraj Med J. 2007;59(4): 177-80.
32. Bedavanija A, Tantilipikorn P, Banditsing C, Kiratisin P, Bunnag C. A comparison of the bacterial culture results of maxillary sinus mucosa and pus collections for chronic maxillary rhinosinusitis. Siriraj Med J. 2019;71(2):95-101.
33. Xu J, Yu L, Liu F, Wan L, Deng Z. The effect of cytokines on osteoblasts and osteoclasts in bone remodeling in osteoporosis: A review. Front Immunol. 2023;14:1222129.
34. Khalmuratova R, Park JW, Shin HW. Immune cell responses and mucosal barrier disruptions in chronic rhinosinusitis. Immune Netw. 2017;17(1):60-67.
35. Onoe Y, Miyaura C, Kaminakayashiki T, Nagai Y, Noguchi K, Chen QR, et al. IL-13 and IL-4 inhibit bone resorption by suppressing cyclooxygenase-2-dependent prostaglandin synthesis in osteoblasts. J Immunol. 1996;156(2):758-64.
36. Silfverswärd CJ, Penno H, Frost A, Nilsson O, Ljunggren O. Expression of markers of activity in cultured human osteoblasts: effects of interleukin-4 and interleukin-13. Scand J Clin Lab Invest. 2010;70(5):338-42.
37. Khalmuratova R, Shin HW, Kim DW, Park JW. Interleukin (IL)-13 and IL-17A contribute to neoosteogenesis in chronic rhinosinusitis by inducing RUNX2. EBioMedicine. 2019;46: 330-41.
38. Beringer A, Miossec P. Systemic effects of IL-17 in inflammatory arthritis. Nat Rev Rheumatol. 2019;15(8):491-501.
39. Vander Ark A, Cao J, Li X. TGF-β receptors: In and beyond TGF-β signaling. Cell Signal. 2018;52:112-20.
40. Salib RJ, Howarth PH. Transforming growth factor-beta in allergic inflammatory disease of the upper airways: Friend or foe? Clin Exp Allergy. 2009;39(8):1128-35.
41. Van Zele T, Claeys S, Gevaert P, Van Maele G, Holtappels G, Van Cauwenberge P, et al. Differentiation of chronic sinus diseases by measurement of inflammatory mediators. Allergy. 2006;61(11):1280-9.
42. Van Bruaene N, Derycke L, Perez-Novo CA, Gevaert P, Holtappels G, De Ruyck N, et al. TGF-beta signaling and collagen deposition in chronic rhinosinusitis. J Allergy Clin Immunol. 2009;124(2): 253-9, 259.e1-2.
43. Park YG, Kang SK, Kim WJ, Lee YC, Kim CH. Effects of TGF-beta, TNF-alpha, IL-beta and IL-6 alone or in combination, and tyrosine kinase inhibitor on cyclooxygenase expression, prostaglandin E2 production and bone resorption in mouse calvarial bone cells. Int J Biochem Cell Biol. 2004;36(11):2270-80.
44. Takai E, Tsukimoto M, Kojima S. TGF-β1 downregulates COX-2 expression leading to decrease of PGE2 production in human lung cancer A549 cells, which is involved in fibrotic response to TGF-β1. PLoS One. 2013;8(10):e76346.
45. Snidvongs K, Sacks R, Harvey RJ. Osteitis in Chronic Rhinosinusitis. Curr Allergy Asthma Rep. 2019;19(5):24.
46. Shapiro F, Wu JY. Woven bone overview: structural classification based on its integral role in developmental, repair and pathological bone formation throughout vertebrate groups. Eur Cell Mater. 2019;38:137-67.
47. Pokharel M, Khadilkar MN, Sreedharan S, Pai R, Shenoy V, Bhojwani K, et al. Bone changes in chronic rhinosinusitis: Pathological or physiological? Indian J Otolaryngol Head Neck Surg. 2022;74(2):178-84.
48. Cho SH, Shin KS, Lee YS, Jeong JH, Lee SH, Tae K, et al. Impact of chronic rhinosinusitis and endoscopic sinus surgery on bone remodeling of the paranasal sinuses. Am J Rhinol. 2008;22(5): 537-41.
49. Dong Y, Zhou B, Wang X, Huang Z, Wang M, Li Y, et al. Computed tomography and histopathological evaluation of osteitis in rabbit models with rhinosinusitis. Acta Otolaryngol. 2017;137(5):534-40.
50. Che Z, Zhang Q, Zhao P, Lv H, Ding H, Li J, et al. Computed tomography evaluation of unilateral chronic maxillary sinusitis with osteitis. Ear Nose Throat J. 2023;102(5):NP237-NP244.
51. Crișan G, Moldovean-Cioroianu NS, Timaru DG, Andrieș G, Căinap C, Chiș V. Radiopharmaceuticals for PET and SPECT imaging: A literature review over the last decade. Int J Mol Sci. 2022;23(9):5023.
52. Jang YJ, Koo TW, Chung SY, Park SG. Bone involvement in chronic rhinosinusitis assessed by 99mTc-MDP bone SPECT. Clin Otolaryngol Allied Sci. 2002;27(3):156-61.
53. Catalano PJ, Dolan R, Romanow J, Payne SC, Silverman M. Correlation of bone SPECT scintigraphy with histopathology of the ethmoid bulla: Preliminary investigation. Ann Otol Rhinol Laryngol. 2007;116(9):647-52.
54. Li YX, Lin F, Cheng L, Huang Q, Huang ZX, Zhang XQ, Zhou B. Clinical application of modified global osteitis score in chronic rhinosinusitis. Lin Chuang Er Bi Yan Hou Tou Jing Wai Ke Za Zhi. 2017;31(21):1666-70.
55. Bhandarkar ND, Mace JC, Smith TL. The impact of osteitis on disease severity measures and quality of life outcomes in chronic rhinosinusitis. Int Forum Allergy Rhinol. 2011;1(5): 372-8.
56. Brooks SG, Trope M, Blasetti M, Doghramji L, Parasher A, Glicksman JT, et al. Preoperative Lund-Mackay computed tomography score is associated with preoperative symptom severity and predicts quality-of-life outcome trajectories after sinus surgery. Int Forum Allergy Rhinol. 2018;8(6):668-75.
57. Cheng L, Huang ZX, Zhou B, Huang Q, Wang CS, Cui SJ, et al. Clinical implication of global osteitis score system and its role in evaluation of osteitis of chronic rhinosinusitis. Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi. 2013;48(2):119-22.
58. Huang Z, Hajjij A, Li G, Nayak JV, Zhou B, Hwang PH. Clinical predictors of neo-osteogenesis in patients with chronic rhinosinusitis. Int Forum Allergy Rhinol. 2015;5(4):303-9.
59. Aparna S, George S. The impact of osteitis on quality of life in patients with chronic rhinosinusitis. Indian J Otolaryngol Head Neck Surg. 2023;75(Suppl 1):1056-61.
60. Gallo S, Russo F, Mozzanica F, Preti A, Bandi F, Costantino C, et al. Prognostic value of the Sinonasal Outcome Test 22 (SNOT-22) in chronic rhinosinusitis. Acta Otorhinolaryngol Ital. 2020;40(2):113-21.
61. Ting F, Hopkins C. Outcome measures in chronic rhinosinusitis. Curr Otorhinolaryngol Rep. 2018;6(3):271-5.
62. Yim MT, Smith KA, Alt JA, Orlandi RR. The value of endoscopic sinus surgery in chronic rhinosinusitis. Laryngoscope Investig Otolaryngol. 2021;6(1):58-63.

63. Bassiouni A, Naidoo Y, Wormald PJ. When FESS fails: The inflammatory load hypothesis in refractory chronic rhinosinusitis. Laryngoscope. 2012;122(2):460-6.
64. Wreesmann VB, Fokkens WJ, Knegt PP. Refractory chronic sinusitis: Evaluation of symptom improvement after Denker's procedure. Otolaryngol Head Neck Surg. 2001;125(5):495-500.
65. Wang M, Zhou B, Li Y, Cui S, Huang Q. Radical versus functional endoscopic sinus surgery for osteitis in chronic rhinosinusitis. ORL J Otorhinolaryngol Relat Spec. 2021;83(4):234-41.
66. Aldajani A, Alroqi A, Alrashidi A, Alsaif A, Almeshari S, Aldwaighri M, et al. Outcomes of endoscopic sinus surgery for chronic rhinosinusitis with nasal polyposis and risk factors of recurrence in a tertiary care teaching hospital. Ther Adv Allergy Rhinol. 2024;15:27534030241274764.
67. Khalmuratova R, Lee M, Park JW, Shin HW. Evaluation of neo-osteogenesis in eosinophilic chronic rhinosinusitis using a nasal polyp murine model. Allergy Asthma Immunol Res. 2020;12(2):306-21.
68. Hershey GK. IL-13 receptors and signaling pathways: An evolving web. J Allergy Clin Immunol. 2003;111(4):677-90; quiz 691.
69. Mukerji SS, Pynnonen MA, Kim HM, Singer A, Tabor M, Terrell JE. Probiotics as adjunctive treatment for chronic rhinosinusitis: A randomized controlled trial. Otolaryngol Head Neck Surg. 2009;140(2):202-8.
70. Chiu AG, Palmer JN, Woodworth BA, Doghramji L, Cohen MB, Prince A, et al. Baby shampoo nasal irrigations for the symptomatic post-functional endoscopic sinus surgery patient. Am J Rhinol. 2008;22(1):34-7.
71. Alandejani T, Marsan J, Ferris W, Slinger R, Chan F. Effectiveness of honey on Staphylococcus aureus and Pseudomonas aeruginosa biofilms. Otolaryngol Head Neck Surg. 2009;141(1):114-8.
72. Paramasivan S, Drilling AJ, Jardeleza C, Jervis-Bardy J, Vreugde S, Wormald PJ. Methylglyoxal-augmented manuka honey as a topical anti-Staphylococcus aureus biofilm agent: Safety and efficacy in an in vivo model. Int Forum Allergy Rhinol. 2014;4(3):187-95.
73. Jervis-Bardy J, Boase S, Psaltis A, Foreman A, Wormald PJ. A randomized trial of mupirocin sinonasal rinses versus saline in surgically recalcitrant staphylococcal chronic rhinosinusitis. Laryngoscope. 2012;122(10):2148-53.
74. Kurasirikul S, Jirapongsananurak O, Vichyanond P, Visitsunthorn N. Gentamicin nasal irrigation in children with chronic rhinosinusitis: A retrospective cohort of 38 patients. Siriraj Med J. 2014;66(2):28-32.
75. Al-Mutairi D, Kilty SJ. Bacterial biofilms and the pathophysiology of chronic rhinosinusitis. Curr Opin Allergy Clin Immunol. 2011;11(1):18-23.
76. Foreman A, Jervis-Bardy J, Wormald PJ. Do biofilms contribute to the initiation and recalcitrance of chronic rhinosinusitis? Laryngoscope. 2011;121(5):1085-91.
77. Foreman A, Boase S, Psaltis A, Wormald PJ. Role of bacterial and fungal biofilms in chronic rhinosinusitis. Curr Allergy Asthma Rep. 2012;12(2):127-35.
78. Dong D, Yulin Z, Xiao W, Hongyan Z, Jia L, Yan X, et al. Correlation between bacterial biofilms and osteitis in patients with chronic rhinosinusitis. Laryngoscope. 2014;124(5):1071-7.
79. Hadi W, Handoko E, Noorhamdani, Prawiro SR. Effect of ethanolic extract propolis Trigona spp. Malang Indonesia on isolate Staphylococcus aureus biofilm architecture from chronic rhinosinusitis: A confocal laser scanning microscopic study. Int J Pharm Sci Res. 2019;10(6):2711-7.
80. Goggin R, Jardeleza C, Wormald PJ, Vreugde S. Corticosteroids directly reduce Staphylococcus aureus biofilm growth: An in vitro study. Laryngoscope. 2014;124(3):602-7.

# Integrating Artificial Intelligence into Chronic Kidney Disease Care: Enhancing Hemodialysis Scheduling, Comorbidity Management, and Diagnostic Capabilities

**Raniah I. Alnaser, M.Sc.[1,*], Mohammed N. Abed, Ph.D.[2], Fawaz A. Alassaf, Ph.D.[3], Mohammad H. Alsaaty, M.D.[4]**

*[1]Nineveh Health Directorate, Mosul, Nineveh province, 41006, Iraq, [2]University of Mosul, College of Pharmacy, Department of Clinical Laboratory Sciences, Mosul, Nineveh province, 41002, Iraq, [3]University of Mosul, College of Pharmacy, Department of Pharmacology and Toxicology, Mosul, Nineveh province, 41002, Iraq, [4]Department of Medicine, College of Medicine, University of Mosul, Mosul, Nineveh province, 41002, Iraq.*




*Corresponding author: Raniah I. Alnaser
E-mail: raniahimad@yahoo.com
Received 23 March 2025 Revised 15 April 2025 Accepted 24 April 2025
ORCID ID:http://orcid.org/0009-0007-8832-5342
https://doi.org/10.33192/smj.v77i7.274440

## ABSTRACT

Chronic kidney disease (CKD) is one of the most common and serious illnesses affecting individuals worldwide, potentially leading to kidney failure. Various strategies exist to manage CKD, with hemodialysis being the most effective. However, this treatment comes with numerous limitations that can significantly affect patients' quality of life. Therefore, it is crucial to explore new approaches to address these challenges. Recently, artificial intelligence (AI) has emerged as a promising tool in nephrology. This review aims to reduce the need for frequent dialysis and to explore the future potential of AI in the field of nephrology. The frequency of hemodialysis refers to the regular, scheduled dialysis sessions mainly prescribed for patients with CKD, in addition to the unplanned or premature initiation of hemodialysis through predictive and preventive interventions. This narrative review systematically searched Google Scholar and PubMed using keywords related to CKD, hemodialysis, and AI. AI is used in kidney disease to predict CKD progression, evaluate drug prescriptions, detect medical errors, adjust dialysis schedules, and identify unknown comorbidities and phenotypes. Integrating AI in nephrology holds promise for reducing kidney dialysis frequency through its applications in the management plans of patients with CKD.


**Keywords:** Chronic kidney disease; hemodialysis; artificial intelligence (Siriraj Med J 2025; 77: 543-552)

## INTRODUCTION

Chronic kidney disease (CKD) is one of the most common and fatal disorders that affect individuals globally and it is expected to be the fifth leading cause of death worldwide by 2040.[1] It is characterized by a gradual decline in kidney function, which results from several causes including diabetes, hypertension, obesity, polycystic kidney disease, glomerulonephritis, and chronic obstruction of the urinary tract.[2] CKD is classified into five stages based on the estimated glomerular filtration rate (eGFR), where the fifth stage is referred to as end-stage kidney disease (ESKD) or kidney failure in which the eGFR is less than 15 milliliters per minute. This stage could progress to multiple overlapping complications and usually needs kidney dialysis or kidney transplantation.[3] Hemodialysis is the most effective treatment modality for CKD patients, and it has a role in the longevity of the patient's life. However, this treatment has several modifications and restrictions impacting the quality of life for patients with kidney failure. Particularly, hemodialysis impacts the social and economic status of these patients as well as their psychological standing leading to a large number of psychological illnesses.[4] Moreover, numerous complications such as cardiovascular issues, infections, and problems related to the dialysis access sites. Frequent hemodialysis sessions could be associated with cardiovascular strain, due to rapid shifts in the fluids and electrolytes.[5] This may exacerbate left ventricular hypertrophy and augments the risk of arrhythmias. Likewise, repeated access to the ventricle increases the tendency of infections, thrombosis and stenosis.[6] Accordingly, it is important to implement a new strategy to adjust the impact and frequency of hemodialysis to improve the quality of patients' life and to overcome the aforementioned challenges.[7]

Artificial intelligence (AI) is recently implemented as an interesting strategy in various fields of medicine, including nephrology. It could be applied to monitor hemodialysis, improve clinical care, and follow-up transplant recipients.[8] The potential of AI in nephrology aims to improve kidney dialysis for CKD patients by reducing frequency and enhancing quality of life. This review discusses the role of AI in dialysis treatment, addressing challenges and limitations while exploring future applications in nephrology.

### 1. Current strategies in chronic kidney disease treatment

The management of CKD involves several approaches aimed at reducing symptoms, slowing progression, and improving quality of life.[9]

#### 1.1 Management of the coexistent diseases

##### 1.1.1 Reducing the risk of cardiovascular disease

Cardiovascular disease is more prevalent in patients with CKD than in those without, often leading to life-threatening outcomes.[10] Therefore, it is important to reduce the risk of cardiovascular diseases by encouraging patients to stop smoking, intensive blood pressure control (target blood pressure is less than 120 millimeters of mercury (mmHg)), and treating elderly patients who have CKD with low to moderate doses of statin regardless of the levels of low-density lipoprotein cholesterol.[11]

##### 1.1.2 Management of hypertension

Antihypertensive drugs such as angiotensin II

receptor blockers (ARBs) and angiotensin-converting enzyme inhibitors (ACEIs) are valuable in patients with CKD.[12] Renin-angiotensin-aldosterone system blockade can reduce systolic blood pressure by approximately 20 mmHg in patients with hypertension and CKD, the same as calcium channel blockers (CCBs) and diuretics.[13] However, in patients with diabetes and proteinuria, ARBs and ACEIs show superiority in comparison to other groups.[14]

Diuretic therapy can alleviate volume overload in CKD patients, helping reduce left ventricular mass and arterial stiffness. For non-proteinuric CKD, thiazide (e.g., bendroflumethiazide) or thiazide-like diuretics (e.g., indapamide) may be first-line treatments. While loop diuretics (e.g., furosemide) are useful, they require higher doses for patients with lower eGFR rates.[15] CCBs, particularly dihydropyridines like amlodipine, are effective for managing hypertension in non-proteinuric CKD patients. For proteinuric CKD, a combination of ACEIs and CCBs is recommended, while non-dihydropyridines like verapamil can reduce proteinuria and control blood pressure.[16]

Beta-blockers can be safely used in all stages of renal impairment with potential dosage adjustments. Carvedilol, a liver-excreted beta-blocker with vasodilatory effects, is especially beneficial for CKD patients.[17,18]

### 1.1.3 Management of diabetes mellitus

Controlling blood sugar delays the progression of CKD with a goal of 7% hemoglobin A1c as recommended by most studies.[19-21] Moreover, adjusting the dose of oral antidiabetics will also be helpful. For instance, the dose of drugs that are metabolized by the liver and partially excreted by the kidney (such as metformin and some sodium-glucose cotransporter 2 inhibitors (SGLT2) and dipeptidyl peptidase inhibitors) may require adjustment or cessation, especially when eGFR falls below 30 mL/min/1.73 $m^2$.[22,23] Drugs cleared by the kidney, like glyburide, should be avoided. SGLT2 inhibitors are recommended for patients with severe albuminuria.[24,25]

The nephrology and clinical evaluation (CREDENCE) trial found that patients with type 2 diabetes and CKD taking canagliflozin had a 30% lower risk of primary renal complications compared to placebo, indicating cardiovascular benefits.[26,27]

### 1.1.4 Nephrotoxins

Nephrotoxic medications should be avoided for patients with CKD for example, non-steroidal anti-inflammatory drugs, especially among patients who are on ARBs or ACEIs therapy.[28,29] Phosphate-based bowel preparations can cause acute phosphate nephropathy, so it's important to advise caution. Herbal remedies are concerning as they lack FDA standardization; some, particularly those with anthraquinones and aristolochic acid, can lead to kidney issues like acute and chronic interstitial nephritis, nephrolithiasis, acute tubular necrosis, Fanconi syndrome, rhabdomyolysis, and hypokalemia.[30-32] Proton pump inhibitors have been linked to acute interstitial nephropathy in individuals with CKD. Their use in this population should be evaluated in primary care units.[33,34]

### 1.1.5 Drug dosing

Patients with CKD require dose adjustment of certain drugs due to the high risk of adverse drug reactions.[35] These drugs include antibiotics, opiates, and anticoagulants.[36] Furthermore, contrast agents such as gadolinium-based agents are contraindicated in patients with acute kidney injury with an eGFR rate of less than 30 mL/min/1.73$m^2$ or ESKD due to the potential of nephrogenic systemic fibrosis.[37] New macrocyclic chelate formulations like gadobutrol, gadoteritol, and gadoterate have a lower risk of fibrosis. However, avoiding gadolinium entirely is the best way to prevent this complication. Patients should be informed of the fibrosis risk if it must be used, and a nephrologist may be consulted for post-exposure hemodialysis.[38]

## 1.2 Monitoring and treating of chronic kidney disease complications

The Kidney Disease Improving Global Outcomes (KDIGO) guidelines recommend annual monitoring of kidney function, eGFR, and albuminuria. High-risk patients should be monitored twice a year, while very high-risk patients require thrice a year.[39,40] Patients with moderate to severe CKD are at an increased risk of minerals, electrolytes, bone abnormalities, and anemia. Laboratory abnormalities are assessed according to the stage of CKD and include evaluations of blood count, lipid panel, phosphate, serum albumin, 25-hydroxyvitamin D, parathyroid hormone, and metabolic panel.[39]

### 1.2.1 Anemia

It is important to monitor anemia in CKD patients by assessing iron stores, and in case of iron deficiency; oral or intravenous replacement therapy might be assisted.[41] However, patients with very low levels of hemoglobin less than 10 g/dL might need additional therapy, such as erythropoietin-stimulating agents after weighing the potential benefits of such therapy against the risks that include stroke, venous thromboembolism, and death.[42]

### 1.2.2 Electrolytes, minerals, and bone abnormalities

Electrolyte abnormalities occur in 3-11% of CKD patients. Early management includes dietary restrictions for hyperphosphatemia and hyperkalemia. Oral bicarbonate supplementation is recommended for serum bicarbonate levels below 22 mmol/L to reduce the risk of metabolic acidosis and slower CKD progression.[43]

In addition, mineral and bone disorders are also common, and many nephrologists have agreed to address the concomitant hypocalcemia, hyperphosphatemia, and vitamin D deficiency by elemental calcium intake, a low-phosphate diet and/or phosphate binders, and vitamin D supplementation.[44]

## 2. Limitations of the current strategies in chronic kidney disease treatment

Managing CKD with traditional strategies addresses several challenges including, economic and healthcare burdens in which significant costs associated with medications, dialysis, and hospitalization limit the access to optimal care for some patients.[45] The side effects of certain medications that are used to manage CKD such as ARBs and ACEIs which can lead to hyperkalemia and kidney dysfunction, will complicate the treatment and require further monitoring and medical interventions.[46] Patients with CKD may struggle to maintain medication adherence, lifestyle changes, regular dialysis sessions, and dietary restrictions which can impact the overall health and the effectiveness of treatment.[47] Finally, the quality of life will be reduced due to the physical and emotional burdens of CKD and its management through which patients experience depression, fatigue, and social isolation that would complicate their treatment adherence.[48]

The limitations highlight the need to integrate advanced technologies like AI into traditional CKD management strategies, potentially enhancing patient care and quality of life.

## 3. The applications of artificial intelligence in chronic kidney disease treatment

The use of AI in nephrology is crucial for improving management plans, aiding clinician decision-making, and reducing the need for hemodialysis and hospital visits. AI applications in kidney disease include predicting CKD progression, assisting treatment, identifying medical errors, adjusting dialysis schedules, and detecting unknown comorbidities as illustrated in Fig 1.[49]

### 3.1 Prediction of chronic kidney disease progression

Patients with ESKD have high hospitalization and mortality rates. Therefore, early detection of CKD and controlling its progression are essential for better patient outcomes. Several machine learning techniques in AI generally exist, including logistic regression, linear regression, ensembles like random forest and XGBoost, Lasso, Ridge, support vector machines, k-means clustering, k-nearest neighbors, decision trees, and principal component analysis.[50]

The development of AI techniques nowadays makes them effective choices for CKD prediction.[51] A traditional regression model for predicting kidney failure has been developed by Tangri et al. This model uses clinical and

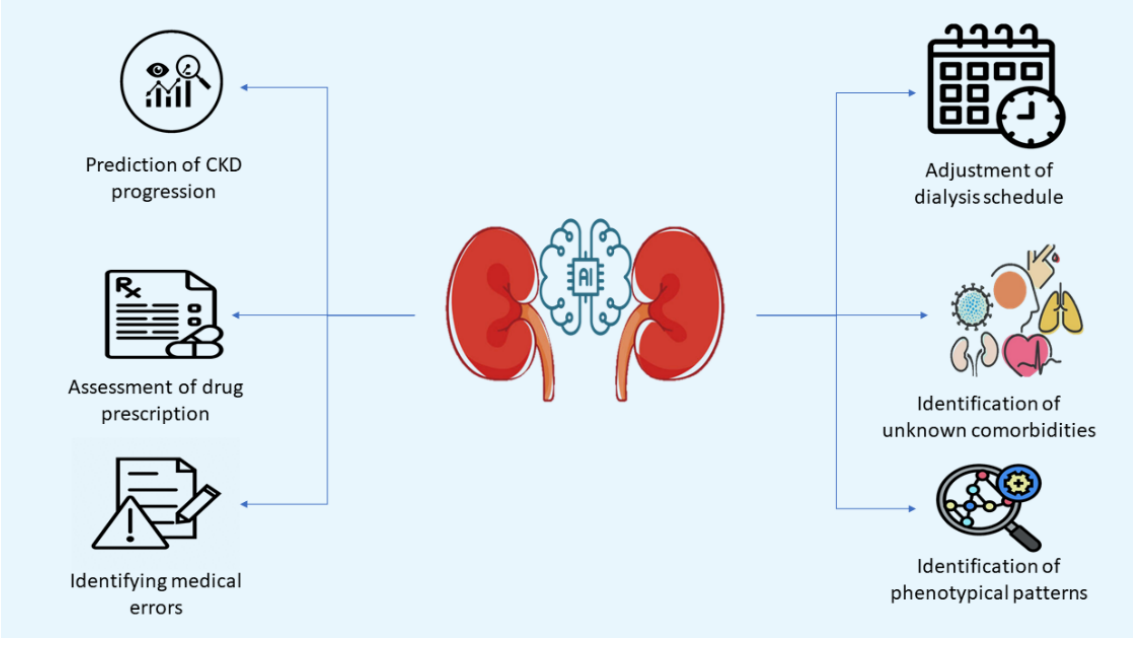



**Fig 1.** Artificial intelligence (AI) applications in chronic kidney disease (CKD).

demographic data from two independent groups of patients with stages 3 to 5 of CKD.[52] In two recent studies, random forest models have been developed to create a prognostic risk score. These models combine data from electronic medical records and circulating biomarkers, such as kidney injury molecule-1 and tumor necrosis factors, to predict the progression of CKD.[53,54]

Predictive interventions using hospitalization and mortality models based on traditional statistical techniques have been performed; utilizing several features including machine learning models. Machine learning is a branch of AI that allows computers to learn from data and make predictions and decisions that might be missed by humans.[55]

Advanced machine learning methods involve developing complex models to enhance the efficacy of machine learning and include several models.[56] The most common model used in nephrology is the random forest, a collaborative method that builds several decision trees using a random set of data to get a more accurate prediction of CKD progression. An early example of employing the random forest models in nephrology was the prediction of sudden cardiac death among elderly hemodialysis patients.[57] Another example was the use of support vector machine model for the prediction of ischemic heart disease in patient with dialysis.[58] The support vector machine is a supervised machine learning algorithm used for the classification and regression of CKD stages.[59]

Xiao et al. compared several machine learning methods to predict the risk of proteinuria in CKD patients using blood biochemical features and demographic data.[60] Moreover, Jamshid Norouzzi et al. established an artificial neural network (ANN) model to predict the progression of kidney failure through the assessment of glomerular filtration rate in patients with CKD, ANN is organized in layers and serves as the foundation for supervised or unsupervised machine learning systems that can replicate complex tasks involved in categorization or prediction procedures.[61]

The Renal Research Institute utilized AI models to predict the progression of CKD, and the results of the models were included in the CKD Forecaster Tool to be used by nephrologists to support clinical decisions. Accordingly, the nephrologists that used this Tool had fewer transitioning patients to hemodialysis providing better care planning for patients with ESKD.[62]

### 3.2 Treatment assessment

AI models provided interesting aid in the prescription of drugs for patients with CKD and identified several medical errors. For example, the erythropoietin prescription could be automated in patients with ESKD to increase its efficacy and enhance patient care. Several approaches have been to reduce the erythropoietin dose and improve patient care.[63] For instance, the use of ANN for the treatment of anemia has minimized the variability of hemoglobin levels and reduced the dose of erythropoietin.[64] Moreover, the historical data of patients is another area of understanding the appropriate drug for specific patient categories, which can guide clinicians in the decision-making process. In this regard, Fresenius Medical Care, is a major dialysis organization that specializes in providing integrated care for kidney disease, has provided nephrologists the option to recommend cinacalcet off-label for in-center administration three times a week, with direct observation.[65] Following observations for the indication of three times per week in-center administration of cinacalcet indicated that it is not superior to the daily administered cinacalcet in regulating the levels of parathyroid hormone, supporting the results of a virtual clinical trial.[66] Despite being theoretical, attempts such as these could improve and customize secondary hyperparathyroidism treatment and increase understanding of the debatable impact of drugs in mineral bone medications on real outcomes.[67]

### 3.3 Identifying medical errors

Medical errors are the third leading cause of death in the United States thus, identifying medical errors is another important strategy to improve patient care which can be ensured by using AI approaches.[68] Several causes of medical errors are available, such as human factors and ergonomics, healthcare system complexity, education, competency, and training. Correcting medical errors using traditional approaches could be accomplished by creating new rules that need to be used in a healthcare system.[69] Nonetheless, the application of AI approaches can be implemented when historical data exist; guiding the clinicians to identify what therapeutic approach is ideal for patients.[70]

Additionally, machine learning algorithms can aid in decision-making when the susceptibility of complex and uncommon medical interactions is expected such as therapeutic duplication and drug-allergy or drug-drug interactions.[71] Several technological companies have facilities that assist in reducing medical errors by supporting physicians in interacting with patients' data. Therefore, integrating the technology with AI models could assist in the decision support system. The prediction of algorithms developed at the University of Stanford has established the best examples of such applications where researchers have created a network to study drug

interactions with over 19,000 proteins in the body. They used deep learning approach of AI to identify patterns in side effects based on drug-targeting proteins. The system supposed patterns about drug interaction side effects and predicted unexpected consequences from taking two drugs together.[72]

### 3.4 Adjustment of dialysis schedule

Fascinatingly, AI algorithms adjust dialysis settings and schedules based on patient-specific data, such as electrolyte balance and fluid retention. This would help in individualization of the dialysis frequency according to the patient's need, which may involve increasing, decreasing or rescheduling sessions, aiding in replacing the standard fixed hemodialysis schedule, reducing the potential complications and customizing treatment regimens and leading to optimizing dialysis treatments and improving quality of life.[73]

### 3.5 Identification of unknown comorbidities

Patients' comorbidities are another area of concern in which the rate of survival will decrease among patients with ESKD and multiple comorbidities. Additionally, the health picture of patients could be complicated by the presence of comorbidities as highlighted by prognostic comorbidity indexes which demonstrate the mortality risk of patients with renal replacement therapy.[74] Moreover, comorbidity information is an important element participating in medical billing. In which patients suffering from ESKD, who have complex health pictures, receive multiple payment coverage for medical services.[75] Therefore, comorbidities must be properly recorded in medical documents in order to appropriate the payment procedures for extra levels of services and support tied to these populations.[76]

In nephrology, one long dialysis organization integrated with a kidney disease care organization has used the machine learning model to identify potentially undocumented comorbidities and to eliminate the expected comorbidities, by finding the patterns in physicians' notes regarding the diseases.[76]

### 3.6 Identification of phenotypical patterns

On the topic of mortality rate, it increases nearly six-fold with the presence of concomitant pathophysiological risk factors including malnutrition, inflammation, and atherosclerosis, which have been detected by traditional statistical methods in patients with ESKD.[77] Additionally, fluid overload has been included in recent studies' pathophysiological patterns. In nephrology, the detection of patterns is based on an unsupervised learning technique, a category of machine learning that teaches a computer to use unlabeled data and allows the program to work without human oversight.[78,79] For example, three distinct phenotypic patterns were found using agglomerative hierarchical clustering, an unsupervised learning technique, in patients with heart failure who had retained ejection fraction. These patterns were based on echocardiographic, laboratory, and clinical characteristics. The mortality risk varied significantly among these groupings. The authors referred to the application of unsupervised learning methods in cardiology to identify phenotypical patterns as "phenomapping", thus monitoring the progress of cardiovascular disease among CKD patients.[80]

In the field of infection medicine, researchers applied k-means clustering, a type of unsupervised learning, to analyze a group of sepsis patients, identifying four distinct phenotypes that exhibited significant differences in outcomes. One of these was the β phenotype, which consisted of older patients with more chronic illnesses and renal failure. Another phenotype, the δ phenotype, included patients experiencing liver dysfunction and septic shock, and it had the highest mortality rate at 40%. In comparison, the β phenotype had a mortality rate of 13%, while the low-risk α phenotype had a rate of just 5%.[81] Additional research is necessary to identify phenotypic patterns in patients with CKD to personalize treatment.

## 4. Clinical applications of artificial intelligence in kidney dialysis

In 2001, Akl and his colleagues addressed the application of AI to the urea kinetic of hemodialysis patients, aiming to predict the adequate dialysis time to reach target urea removal, as a result, they concluded that AI can provide valuable insights for tailoring intradialysis protocols to meet individual clinical needs. This method improved the customization of hemodialysis session prescriptions, especially for patients with varying weight and dietary habits. Instead of using a standard prescription, their dialysis sessions should be tailored to meet their unique needs.[82]

In 2004, Gabutti and his colleagues explored the role of the AI approach to assist nephrologists in accurately recognizing the trend in the evolution of the protein nutritional status and they concluded that the predicted protein catabolic rate is more accurate than the protein catabolic rate established by clinicians, hence enabling the implementation of preventative interventions.[83]

In 2016, an Anemia Control Model (ACM) using artificial intelligence decision support system, was conducted in an aim to improve anemia outcomes for patients

undergoing hemodialysis. The ACM was developed based on patient profiles and was used to recommend appropriate doses of erythropoietic-stimulating agents such as darbepoetin. The study included 752 patients receiving hemodialysis treatment across three NephroCare clinics located in different countries. The primary outcomes measure was the percentage of hemoglobin values that fell within the target range, the individual fluctuations in hemoglobin levels, and the median dose of darbepoetin administered. The results indicated that care guided by the ACM led to a decrease in hemoglobin variability, a significant reduction in darbepoetin usage, and an increase in the percentage of hemoglobin values on target. These findings suggest that the ACM can enhance anemia management, reduce the need for erythropoietic-stimulating agents, and significantly lower treatment costs.[64]

In 2018, Neil and Bastard utilized AI a study to increase the precision of dry weight measurements in hemodialysis patients. Dry weight refers to the minimum weight that patients on hemodialysis can safely tolerate. Accurate estimation of dry weight is essential to reduce morbidity and mortality but can be challenging to achieve. A neural network was created using blood pressure readings, blood volume monitoring, and bio-impedancemetry. Fourteen children were moved from nephrologists to dry-weight AI patients. According to the findings, the dry weight of AI was 28.6%, 50%, or the same as that of nephrologists. Systolic blood pressure was considerably lower and antihypertensive medications were successfully stopped in patients with higher artificial intelligence dry weight. Finally, the study concluded that AI is a powerful tool for predicting dry weight among patients with hemodialysis.[84]

### 5. Future promising issues of AI in CKD and potential limitations

Nephrologists are increasingly collaborating with AI researchers to enhance kidney disease diagnosis and treatment. AI has great promise for CKD research regarding early and precise kidney disease prediction, enabling individualized risk assessments for CKD patients.[50]

AI can analyze vast patient data and identify complex relationships that traditional prediction models often miss. AI can create detailed risk profiles by integrating additional information such as biomarkers, socioeconomic factors, medical images, genetic markers, and comorbidities. This enables the personalization of treatment plans for patients.[50] For example, deep learning models can detect early kidney disease in patients with type 2 diabetes mellitus by analyzing retinal fundus photographs and clinical metadata, including sex, age, weight, height, blood pressure, and body mass index.[85] Furthermore, external photographs of the eyes can help identify poor blood glucose control, serving as a warning for the progression of diabetic complications, including kidney failure.[86]

In a significant project, Google DeepMind collaborated with the United States Department of Veterans Affairs to develop an AI system capable of predicting acute kidney injury up to 2 days before it becomes clinically apparent. This system demonstrates how data science can be applied to nephrology to develop effective tools for the prevention and early detection of kidney disease, optimizing the use of medical resources, and potentially saving lives.[87]

Using AI-based approaches in clinical practice presents significant challenges, particularly because deep learning models operate as "black boxes", due to their complexity, it is often impossible to trace the path from input to output, making it difficult for doctors to assess and understand the predictions made by these models and complicating the understanding of how decisions are made.[88]

Researchers are increasingly focusing on Explainable AI (ExAI) to address important issues in artificial intelligence. ExAI aims to enhance the transparency of AI models by providing clear explanations for their decisions. Techniques such as attention mechanisms, feature importance analysis, and model-agnostic approaches are used to identify the factors that influence AI predictions. By improving transparency, ExAI helps clinicians assess the reliability of AI systems, trust their outputs, and integrate them into clinical decision-making processes with greater confidence. Addressing the "black box" problem is essential for complying with regulatory standards, ensuring accountability, and addressing concerns related to patient privacy, safety, and ethical implications.[89] A new reporting approach for medical models is needed to improve healthcare outcomes, focusing on ethical considerations and data-sharing policies.

## CONCLUSION

Despite the challenges regarding data privacy and ethical considerations that need more clinical validation, the integration of AI in nephrology provides a potential value in reducing the frequency of kidney dialysis. This could be obtained through early disease prediction, treatment assessment, and identification of expected comorbidities. Besides, AI approaches could personalize patient care and improve the quality of life by customizing dialysis schedules and settings.

### Data Availability Statement

The data supporting the findings of this review article are available from the cited primary literature sources. No new data were generated or analyzed for this study.

## ACKNOWLEDGMENTS

We are grateful for the support from the Nineveh Health Directorate and the University of Mosul, College of Pharmacy, and College of Medicine.

## DECLARATIONS

### Grants and Funding Information

There are no grants and funding sources in this research.

### Conflict of Interest

The authors declare no competing conflicts of interest.

### Registration Number of Clinical Trial

There is no clinical trial number because this study is not a clinical trial/experimental study.

### Author Contributions

Conceptualization and methodology, RIA, MNA, FAA; Investigation, MNA, FAA, and MHA. Visualization and writing – original draft, RIA; Writing-review and editing, MNA, FAA; Supervision, MNA, FAA, and MHA.; All authors have read and agreed to the final version of the manuscript.

### Use of Artificial Intelligence

This study did not use artificial intelligence.

## REFERENCES

1. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. Lancet. 2018;392(10159):2052–90.
2. Romagnani P, Remuzzi G, Glassock R, Levin A, Jager KJ, Tonelli M, et al. Chronic kidney disease. Nat Rev Dis Prim. 2017;3(1):17088.
3. Ammirati AL. Chronic Kidney Disease. Rev Assoc Med Bras. 2020;66(suppl 1):s03–9.
4. Lateef A. Psychological Impact of Chronic Kidney Disease and Hemodialysis: Narrative Review. Psychosom Med Res. 2022;4(2):9.
5. Bello AK, Okpechi IG, Osman MA, Cho Y, Htay H, Jha V, et al. Epidemiology of haemodialysis outcomes. Nat Rev Nephrol. 2022;18(6):378–95.
6. Zaki MK, Abed MN, Alassaf FA. Antidiabetic Agents and Bone Quality: A Focus on Glycation End Products and Incretin Pathway Modulations. J Bone Metab. 2024;31(3):169–81.
7. Himmelfarb J, Vanholder R, Mehrotra R, Tonelli M. The current and future landscape of dialysis. Nat Rev Nephrol. 2020;16(10):573–85.
8. Niel O, Bastard P. Artificial Intelligence in Nephrology: Core Concepts, Clinical Applications, and Perspectives. Am J Kidney Dis. 2019;74(6):803–10.
9. Chen TK, Knicely DH, Grams ME. Chronic Kidney Disease Diagnosis and Management. JAMA. 2019;322(13):1294.
10. Saran R. US renal data system. 2018 USRDS annual data report: epidemiology of kidney disease in the United States. Vol. 73, Am J Kidney Dis. 2019.p.Svii-Sxxii, S1-S.
11. Anderson TJ, Grégoire J, Pearson GJ, Barry AR, Couture P, Dawes M, et al. 2016 Canadian Cardiovascular Society Guidelines for the Management of Dyslipidemia for the Prevention of Cardiovascular Disease in the Adult. Can J Cardiol. 2016;32(11):1263–82.
12. Sarafidis PA, Khosla N, Bakris GL. Antihypertensive Therapy in the Presence of Proteinuria. Am J Kidney Dis. 2007;49(1):12–26.
13. Banerjee D, Winocour P, Chowdhury TA, De P, Wahba M, Montero R, et al. Management of hypertension and renin-angiotensin-aldosterone system blockade in adults with diabetic kidney disease: Association of British Clinical Diabetologists and the Renal Association UK guideline update 2021. BMC Nephrol. 2022;23(1):9.
14. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. J Am Coll Cardiol. 2018;71(19):e127–248.
15. Burnier M, Bakris G, Williams B. Redefining diuretics use in hypertension. J Hypertens. 2019;37(8):1574–86.
16. Ohno S, Ishii A, Yanagita M, Yokoi H. Calcium channel blocker in patients with chronic kidney disease. Clin Exp Nephrol. 2022;26(3):207–15.
17. Tomiyama H, Yamashina A. Beta-Blockers in the Management of Hypertension and/or Chronic Kidney Disease. Int J Hypertens. 2014;2014:1–7.
18. Hocht C, Bertera FM, Del Mauro JS, Santander Plantamura Y, Taira CA, Polizio AH. What is the Real Efficacy of Beta-Blockers for the Treatment of Essential Hypertension? Curr Pharm Des. 2017;23(31):4658-77.
19. Guideline development group, Bilo H, Coentrao L, Couchoud C, Covic A, De Sutter J, et al. Clinical Practice Guideline on management of patients with diabetes and chronic kidney disease stage 3b or higher (eGFR <45 mL/min). Nephrol Dial Transplant. 2015;30(Suppl 2):ii1–142.
20. Alnaser RI, Alassaf FA, Abed MN. Melatonin as a potential treatment option in diabetes complications. Eur J Transl Clin Med. 2024;7(2):78-91.
21. Alassaf FA, Jasim MHM, Alfahad M, Qazzaz ME, Abed MN, Thanoon IAJ. Effects of Bee Propolis on FBG, HbA1c, and Insulin Resistance in Healthy Volunteers. Turkish J Pharm Sci. 2021;18(4):405-9.
22. Ahmed GM, Abed MN, Alassaf FA. An overview of the effects of sodium-glucose co-transporter-2 inhibitors on hematological parameters in diabetic patients. Iraqi J Pharm. 2023;20(1):65–71.
23. Alnaser RI, Alassaf FA, Abed MN. Adulteration of hypoglycemic products: the silent threat. Rom J Med Pract. 2023;18(4):202–5.
24. Madero M, Chertow GM, Mark PB. SGLT2 Inhibitor Use in Chronic Kidney Disease: Supporting Cardiovascular, Kidney,

and Metabolic Health. Kidney Med. 2024;6(8):100851.
25. Alnaser RI, Alassaf FA, Abed MN. Incretin-Based Therapies: A Promising Approach for Modulating Oxidative Stress and Insulin Resistance in Sarcopenia. J Bone Metab. 2024;31(4): 251-63.
26. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet. 1998;352(9131): 837–53.
27. Neal B, Perkovic V, Mahaffey KW, de Zeeuw D, Fulcher G, Erondu N, et al. Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. N Engl J Med. 2017;377(7):644–57.
28. Ahmed GM, Abed MN, Alassaf FA. Impact of calcium channel blockers and angiotensin receptor blockers on hematological parameters in type 2 diabetic patients. Naunyn Schmiedebergs Arch Pharmacol. 2023;397(3):1817-28.
29. Hammo AA, Althanoon ZA, Ahmad AA. The Protective Effect of Coenzyme Q10 against Doxorubicin-induced Nephrotoxicity in Albino Rats. Rev Electron Vet. 2022;23(3):314–25.
30. Abed MN, Qazzaz ME, Alassaf FA. Investigating the nephrotoxic effects of medroxyprogesterone in female albino rats. Ukr J Nephrol Dial. 2024;2(82):25-33.
31. Markowitz GS, Perazella MA. Acute phosphate nephropathy. Kidney Int. 2009;76(10):1027–34.
32. Hammo AA, Ahmad AA, Althanoon ZA. Role of Gender in the Protection Against Doxorubicin-Induced Oxidative Stress. Pharmacogn J. 2023;14(6):782–8.
33. Abed MN, Alassaf FA, Jasim MHM, Alfahad M, Qazzaz ME. Comparison of Antioxidant Effects of the Proton Pump-Inhibiting Drugs Omeprazole, Esomeprazole, Lansoprazole, Pantoprazole, and Rabeprazole. Pharmacology. 2020;105(11–12):645–51.
34. Alfahad M, Qazzaz ME, Abed MN, Alassaf FA, Jasim MHM. Comparison of Anti-Oxidant Activity of Different Brands of Esomeprazole Available in Iraqi Pharmacies. Syst Rev Pharm. 2020;11(5):330-4.
35. Palacio-Lacambra M, Comas-Reixach I, Blanco-Grau A, Suñé-Negre J, Segarra-Medrano A, Montoro-Ronsano J. Comparison of the Cockcroft–Gault, MDRD and CKD-EPI equations for estimating ganciclovir clearance. Br J Clin Pharmacol. 2018; 84(9):2120-8.
36. Inker LA, Astor BC, Fox CH, Isakova T, Lash JP, Peralta CA, et al. KDOQI US Commentary on the 2012 KDIGO Clinical Practice Guideline for the Evaluation and Management of CKD. Am J Kidney Dis. 2014;63(5):713–35.
37. Bahrainwala JZ, Leonberg-Yoo AK, Rudnick MR. Use of Radiocontrast Agents in CKD and ESRD. Semin Dial. 2017; 30(4):290–304.
38. Perazella MA. Advanced kidney disease, gadolinium and nephrogenic systemic fibrosis: the perfect storm. Curr Opin Nephrol Hypertens. 2009;18(6):519–25.
39. Stevens PE, Ahmed SB, Carrero JJ, Foster B, Francis A, Hall RK, et al. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int. 2024;105(4):S117–314.
40. Mohammed AA, Abdulla AA, Karam AA. Measurement of Inflammation-Related Biomarkers in Different Chronic Kidney Diseases in Humans: Role of Aging and Gender? IIUM Med J Malaysia. 2021;20(4):37–43.
41. Ahmed GM, Alassaf FA, Abed MN. The Interplay of the Angiotensin Receptor Blockers and Haematological Abnormalities: Insights and Implications. J Ayub Med Coll Abbottabad. 2023;35(4 (Suppl 1)):785–92.
42. Inker LA, Grams ME, Levey AS, Coresh J, Cirillo M, Collins JF, et al. Relationship of Estimated GFR and Albuminuria to Concurrent Laboratory Abnormalities: An Individual Participant Data Meta-analysis in a Global Consortium. Am J Kidney Dis. 2019;73(2):206–17.
43. Driver TH, Shlipak MG, Katz R, Goldenstein L, Sarnak MJ, Hoofnagle AN, et al. Low Serum Bicarbonate and Kidney Function Decline: The Multi-Ethnic Study of Atherosclerosis (MESA). Am J Kidney Dis. 2014;64(4):534–41.
44. Abed MN, Alassaf FA, Qazzaz ME. Exploring the Interplay between Vitamin D, Insulin Resistance, Obesity and Skeletal Health. J Bone Metab [Internet]. 2024;31(2):75–89. Available from: http://e-jbm.org/journal/view.php?doi=10.11005/jbm.2024.31.2.75
45. Vareesangthip K, Thanapattaraborisuth B, Chanchairujira K, Wonglaksanapimon S, Chanchairujira T. Assessment of Volume Status in Chronic Hemodialysis: Comparison of Lung Ultrasound to Clinical Practice and Bioimpedance. Siriraj Med J. 2023;75(3): 224–33.
46. Raebel MA. Hyperkalemia Associated with Use of Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers. Cardiovasc Ther [Internet]. 2012;30(3):e156–66. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1755-5922.2010.00258.x
47. Srisuwan W, Charoensri S, Jantarakana K, Chanchairujira T. Increasing Dialysate Flow Rate over 500 ml/min for Reused High-Flux Dialyzers do not Increase Delivered Dialysis Dose: A Prospective Randomized Cross Over Study. Siriraj Med J. 2022;74(3):152–60.
48. Hao W, Tang Q, Huang X, Ao L, Wang J, Xie D. Analysis of the prevalence and influencing factors of depression and anxiety among maintenance dialysis patients during the COVID-19 pandemic. Int Urol Nephrol. 2021;53(7):1453–61.
49. Chaudhuri S, Long A, Zhang H, Monaghan C, Larkin JW, Kotanko P, et al. Artificial intelligence enabled applications in kidney disease. Semin Dial. 2021;34(1):5-16.
50. Simeri A, Pezzi G, Arena R, Papalia G, Szili-Torok T, Greco R, et al. Artificial intelligence in chronic kidney diseases: methodology and potential applications. Int Urol Nephrol. 2025;57(1):159-68.
51. Xie Y, Bowe B, Mokdad AH, Xian H, Yan Y, Li T, et al. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. Kidney Int. 2018;94(3):567–81.
52. Tangri N, Grams ME, Levey AS, Coresh J, Appel LJ, Astor BC, et al. Multinational Assessment of Accuracy of Equations for Predicting Risk of Kidney Failure. JAMA. 2016;315(2):164-74.
53. Chauhan K, Nadkarni GN, Fleming F, McCullough J, He CJ, Quackenbush J, et al. Initial Validation of a Machine Learning-Derived Prognostic Test (KidneyIntelX) Integrating Biomarkers and Electronic Health Record Data To Predict Longitudinal Kidney Outcomes. Kidney360. 2020;1(8):731–9.
54. Chan L, Nadkarni GN, Fleming F, McCullough JR, Connolly P, Mosoyan G, et al. Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict progression of diabetic kidney disease. Diabetologia. 2021; 64(7):1504–15.
55. Mahesh B. Machine Learning Algorithms - A Review. Int J Sci Res. 2020;9(1):381–6.
56. MacEachern SJ, Forkert ND. Machine learning for precision

medicine. Genome. 2021;64(4):416–25.
57. Goldstein BA, Chang TI, Mitani AA, Assimes TL, Winkelmayer WC. Near-Term Prediction of Sudden Cardiac Death in Older Hemodialysis Patients Using Electronic Health Records. Clin J Am Soc Nephrol. 2014;9(1):82–91.
58. Mezzatesta S, Torino C, Meo P De, Fiumara G, Vilasi A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. Comput Methods Programs Biomed. 2019;177:9–15.
59. Ramspek C, Voskamp P, van Ittersum F, Krediet R, Dekker F, van Diepen M. Prediction models for the mortality risk in chronic dialysis patients: a systematic review and independent external validation study. Clin Epidemiol. 2017;9:451–64.
60. Xiao J, Ding R, Xu X, Guan H, Feng X, Sun T, et al. Comparison and development of machine learning tools in the prediction of chronic kidney disease progression. J Transl Med. 2019;17(1):119.
61. Norouzi J, Yadollahpour A, Mirbagheri SA, Mazdeh MM, Hosseini SA. Predicting Renal Failure Progression in Chronic Kidney Disease Using Integrated Intelligent Fuzzy Expert System. Comput Math Methods Med. 2016;2016:6080814.
62. Jiao Y, Kopyt N, Bollu P. Use of kidney disease progression model care planning report associates with lower dialysis catheter rates at the initiation of hemodialysis [abstract SA-PO840]. J Am Soc Nephrol. 2019;30:980.
63. Brier ME, Gaweda AE, Aronoff GR. Personalized Anemia Management and Precision Medicine in ESA and Iron Pharmacology in End-Stage Kidney Disease. Semin Nephrol. 2018;38(4):410–7.
64. Barbieri C, Molina M, Ponce P, Tothova M, Cattinelli I, Ion Titapiccolo J, et al. An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients. Kidney Int. 2016;90(2):422–9.
65. Gudrun Schappacher-Tilp, Doris H. Fuertinge PK. A Multi-Compartment Model Capturing the Pharmacokinetics of the Calcimimetic Cinacalce. Cell Physiol Biochem. 2019;52(2):429–38.
66. Schappacher-Tilp G, Cherif A, Fuertinger DH, Bushinsky D, Kotanko P. A mathematical model of parathyroid gland biology. Physiol Rep. 2019;7(7):e14045.
67. Akizawa T, Kurita N, Mizobuchi M, Fukagawa M, Onishi Y, Yamaguchi T, et al. PTH-dependence of the effectiveness of cinacalcet in hemodialysis patients with secondary hyperparathyroidism. Sci Rep. 2016;6(1):19612.
68. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. BMJ. 2016;353:i2139.
69. Kopec D, Kabir MH, Reinharth D, Rothschild O, Castiglione JA. Human Errors in Medical Practice: Systematic Classification and Reduction with Automated Information Systems. J Med Syst. 2003;27(4):297-313.
70. Paredes M. Can Artificial Intelligence help reduce human medical errors? Two examples from ICUs in the US and Peru. Techpolicy Inst. 2021. Available from: https://techpolicyinstitute.org/wp-content/uploads/2018/02/Paredes-Can-Artificial-Intelligence-help-reduce-human-medical-errors-DRAFT.pdf
71. Stafie CS, Sufaru IG, Ghiciuc CM, Stafie II, Sufaru EC, Solomon SM, et al. Exploring the Intersection of Artificial Intelligence and Clinical Healthcare: A Multidisciplinary Review. Diagnostics (Basel). 2023;13(12):1995.
72. Chaudhuri S, Long A, Zhang H, Monaghan C, Larkin JW, Kotanko P, et al. Applications of Artificial Intelligence (AI) in Kidney Disease. Appl Digit Technol Artif Intell Nephrol. 2021;34(1):21.
73. Khan K, Zameer F, Jain P, KR R, Niranjan V, Manoj S, et al. Artificial Intelligence in Revolutionizing Kidney Care and Beyond: Kid-AI Revolution. J Bio-X Res. 2024;7. Available from: https://spj.science.org/doi/10.34133/jbioxresearch.0022
74. Rattanasompattikul M, Feroze U, Molnar MZ, Dukkipati R, Kovesdy CP, Nissenson AR, et al. Charlson comorbidity score is a strong predictor of mortality in hemodialysis patients. Int Urol Nephrol. 2012;44(6):1813–23.
75. Taesilapasathit C, Spanuchart I, Suppadungsuk S, Sutharattanapong N, Vipattawat K, Sethakarun S, et al. Accumulation of Advanced Glycation End Products Independently Increases the Risk of Hospitalization Among Hemodialysis Patients. Siriraj Med J. 2022;74(5):305-13.
76. Brown R, Thorsteinsson E. Comorbidity: What Is It and Why Is It Important? In: Comorbidity. Cham: Springer International Publishing; 2020.p.1–22.
77. Fayos De Arizón L, Viera ER, Pilco M, Perera A, De Maeztu G, Nicolau A, et al. Artificial intelligence: a new field of knowledge for nephrologists? Clin Kidney J. 2023;16(12):2314–26.
78. Dekker MJE, Konings C, Canaud B, van der Sande FM, Stuard S, Raimann JG, et al. Interactions Between Malnutrition, Inflammation, and Fluid Overload and Their Associations With Survival in Prevalent Hemodialysis Patients. J Ren Nutr. 2018;28(6):435–44.
79. Hung SC, Kuo KL, Peng CH, Wu CH, Lien YC, Wang YC, et al. Volume overload correlates with cardiovascular risk factors in patients with chronic kidney disease. Kidney Int. 2014;85(3):703–9.
80. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiade M, et al. Phenomapping for Novel Classification of Heart Failure With Preserved Ejection Fraction. Circulation. 2015;131(3):269–79.
81. Seymour CW, Kennedy JN, Wang S, Chang CCH, Elliott CF, Xu Z, et al. Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. JAMA. 2019;321(20):2003.
82. Akl AI, Sobh MA, Enab YM, Tattersall J. Artificial intelligence: A new approach for prescription and monitoring of hemodialysis therapy. Am J Kidney Dis. 2001;38(6):1277–83.
83. Gabutti L, Burnier M, Mombelli G, Malé F, Pellegrini L, Marone C. Usefulness of artificial neural networks to predict follow-up dietary protein intake in hemodialysis patients. Kidney Int. 2004;66(1):399–407.
84. Niel O, Bastard P. Artificial intelligence improves estimation of tacrolimus area under the concentration over time curve in renal transplant recipients. Transpl Int. 2018;31(8):940–1.
85. Zhang K, Liu X, Xu J, Yuan J, Cai W, Chen T, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. Nat Biomed Eng. 2021;5(6):533–45.
86. Babenko B, Mitani A, Traynis I, Kitade N, Singh P, Maa AY, et al. Detection of signs of disease in external photographs of the eyes via deep learning. Nat Biomed Eng. 2022;6(12):1370–83.
87. Delrue C, De Bruyne S, Speeckaert MM. Application of Machine Learning in Chronic Kidney Disease: Current Status and Future Prospects. Biomedicines. 2024;12(3):568.
88. Hagemann V, Rieth M, Suresh A, Kirchner F. Human-AI teams—Challenges for a team-centered AI at work. Front Artif Intell. 2023;6:1252897.
89. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion. 2020;58:82-115.