PMAC PRINCE MAHIDOL AWARD CONFERENCE 2025

E-ISSN 2228-8082
Volume 77, Number 2, February 2025

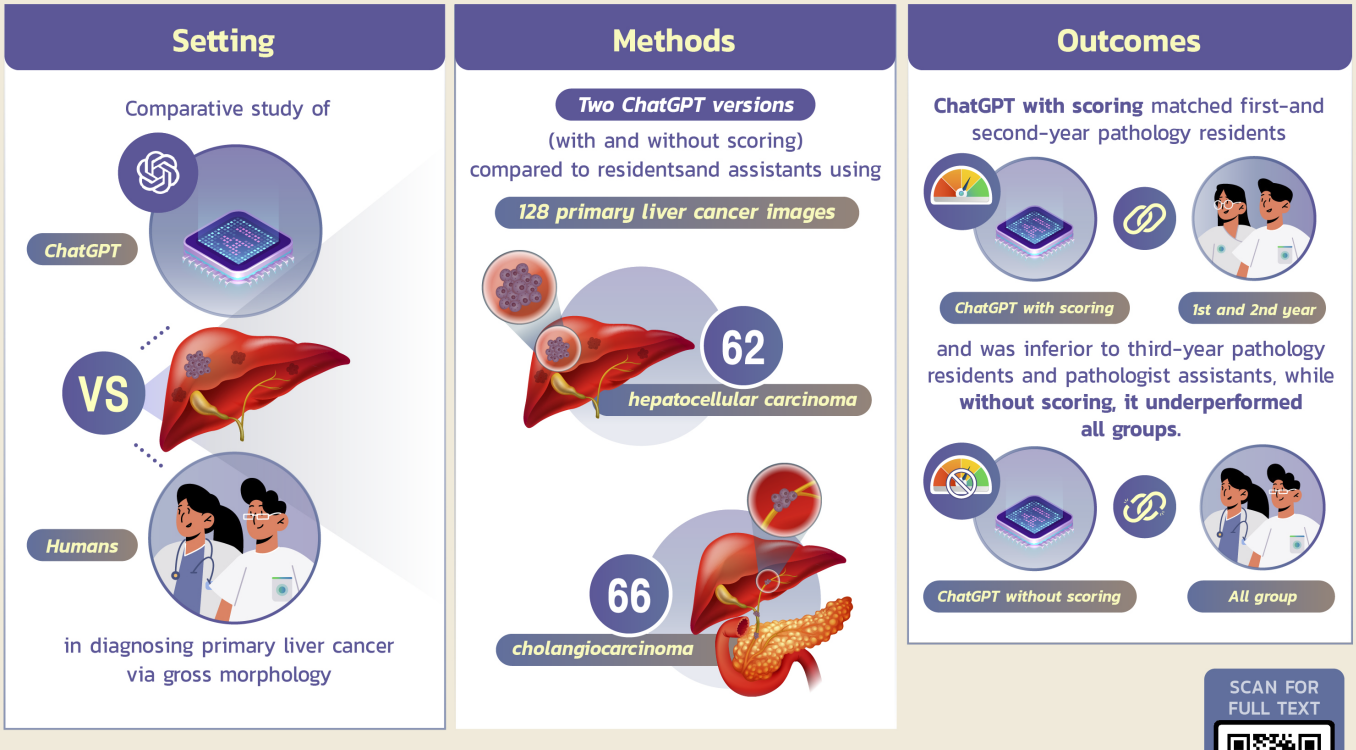
SMJ
Siriraj Medical Journal

The world-leading biomedical science of Thailand

MONTHLY

ORIGINAL ARTICLE

ChatGPT for Gross Morphology of Primary Liver Cancer



SMJ SIRIRAJ MEDICAL JOURNAL | Laohawetwanit, et al. *Siriraj Med J* 2025;77(2):119-129.

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 2
February 2025

ORIGINAL ARTICLE

SIRIRAJ MEDICAL JOURNAL

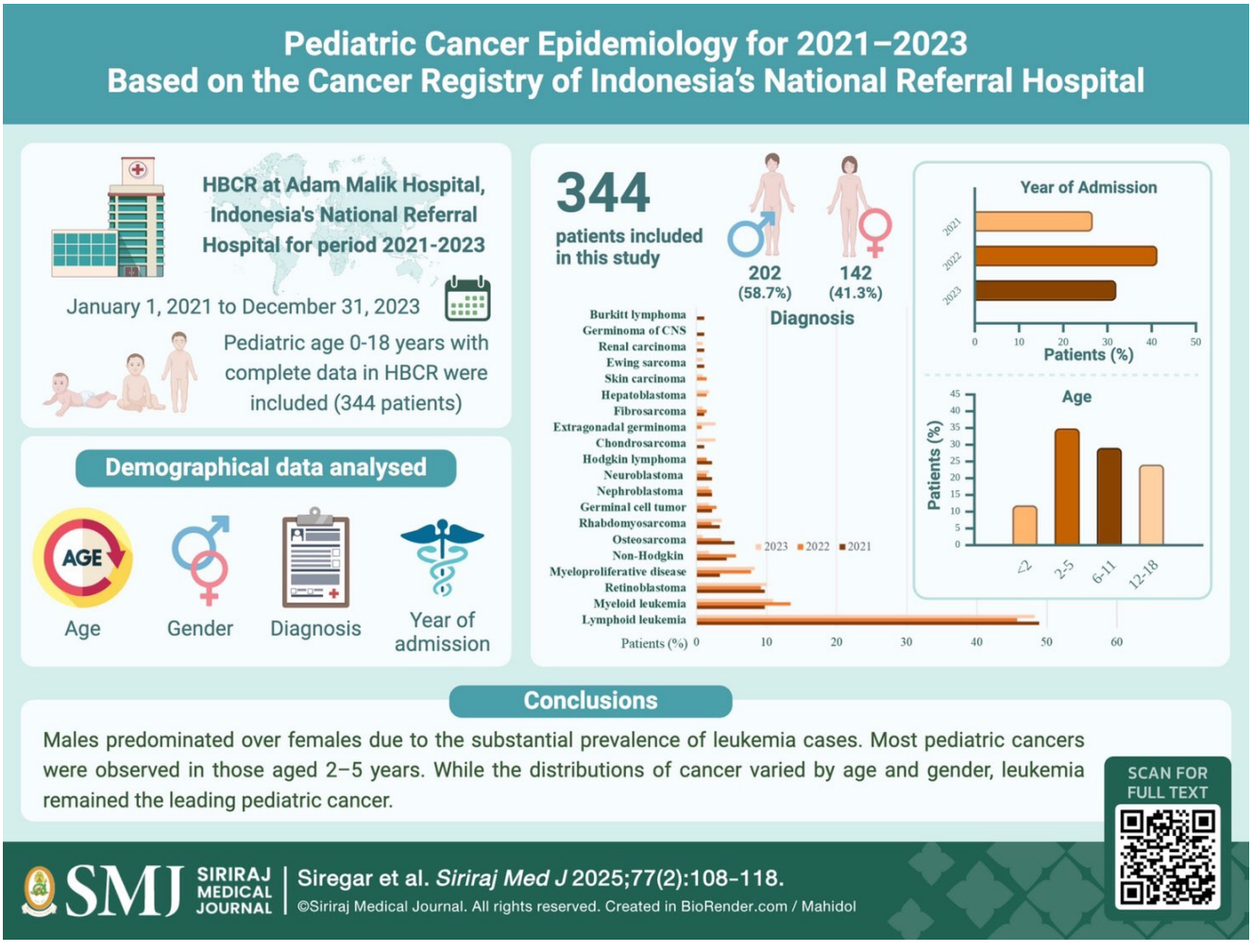
https://he02.tci-thaijo.org/index.php/sirirajmedj/index

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

Pediatric Cancer Epidemiology for 2021–2023 Based on the Cancer Registry of Indonesia's National Referral Hospital

Olga Rasiyanti Siregar[1,*], **Bidasari Lubis**[1], **Peggy Liberty**[2]

[1]*Department of Pediatrics, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia,* [2]*Medical Study Program, Faculty of Medicine, Universitas Sumatera Utara, Medan, Indonesia.*



*Corresponding author: Olga Rasiyanti Siregar
E-mail: olga.rasiyanti@usu.ac.id
Received 7 October 2024 Revised 1 November 2024 Accepted 15 November 2024
ORCID ID:http://orcid.org/0000-0002-6939-5739
https://doi.org/10.33192/smj.v77i2.271534

ABSTRACT

Objective: This study aimed to present epidemiological data on pediatric cancers based on the hospital-based cancer registry (HBCR) at Indonesia's National Referral Hospital.

Material and Methods: This descriptive study analyzed data on pediatric cancers from the HBCR for 2021–2023. Demographical data, including age, gender, diagnosis, and admission year, were extracted from the HBCR data. Univariate data analysis was conducted.

Results: This study included 344 patients, consisting of 202 males and 142 females. Of those patients, 92 (26.7%) were admitted in 2021, 142 (41.3%) in 2022, and 110 (32%) in 2023. The highest cancer incidence was observed in patients aged 2–5 years (34.9%) and the lowest in those aged <2 years (11.9%). The three most common reported cancers were lymphoid leukemia (47.4%), myeloid leukemia (11.6%), and retinoblastoma (9.6%).

Conclusion: Males predominated over females due to the substantial prevalence of leukemia cases. Most pediatric cancers were observed in those aged 2–5 years. While the distributions of cancer varied by age and gender, leukemia remained the leading pediatric cancer.

Keywords: Cancer; data registry; epidemiology; malignancy; pediatric (Siriraj Med J 2025; 77: 108-118)

INTRODUCTION

Cancer is one of the leading causes of mortality in children. In 2020, over 200,000 new cancer cases and over 80,000 cancer-related deaths were reported in children. The global incidence of childhood cancer was 10.5 per 100,000 population, and the mortality rate was 4.1 per 100,000 population.[1] While the global incidence and mortality rate of pediatric cancer had declined from those in 2017 (incidence rate = 16.2 per 100,000 population, mortality rate = 5.5 per 100,000 population), the mortality rate remained higher in low- and middle-income countries, including Indonesia, compared to high-income countries.[2] One previous study reported that the mortality rate for pediatric cancer in low-and middle-income countries was twice that in high-income countries, with Southeast Asia reporting the highest mortality among low- and middle-income countries (mortality rate = 4.5 per 100,000 population).[1]

Indonesia, as a part of low- and middle-income countries in Southeast Asia, also reported a notable prevalence of pediatric cancer. According to Basic Health Research of Indonesia (RISKESDAS), the prevalence of pediatric cancer in Indonesia was 0.42 per 1,000 population in 2018.[3] The Global Cancer Observatory (GLOBOCAN) estimated that 291 per 100,000 Indonesians were diagnosed with cancer in 2018, and the number is expected to increase in the future.[4] According to data from the Social Insurance Administration Organization of Indonesia, approximately 2.2 trillion rupiahs were allocated to treating cancers, including pediatric cancers, in 2016, increasing by about 30% in 2017 and reaching 7.6 trillion rupiahs in 2020.[5] Despite the considerable funding designated for cancer treatment in Indonesia, the treatment success rate remains low, contributing to the high mortality rate. This unsuccessful treatment has been attributed to the delayed diagnosis, leading to advanced disease at diagnosis and delays in treatment.[6]

Understanding the characteristics and distribution of pediatric cancers should facilitate their accurate early diagnosis and prompt treatment. To our knowledge, no valid data has been reported for pediatric cancers. Therefore, this study aimed to report epidemiological data on pediatric cancers based on the hospital-based cancer registry (HBCR) at the National Referral Hospital of Indonesia for the period 2021–2023.

MATERIALS AND METHODS

This cross-sectional descriptive study analyzed data from the HBCR of Adam Malik Hospital, Indonesia's National Referral Hospital. In 2016, Adam Malik Hospital, a tertiary healthcare in Indonesia, had been appointed by the Ministry of Health of Indonesia as one of the national cancer registry hospitals. All pediatric cancer cases detected in primary healthcare were referred to secondary healthcare for confirmation of diagnosis by histopathological examinations. The primary healthcare in this study included district hospitals and general practitioner clinics, while the secondary healthcare included regional referral hospitals. Confirmed cases were referred to tertiary healthcare, including Adam Malik Hospital, for diagnostic stratification and treatment decision. All pediatric cancer cases referred to Adam Malik Hospital will be included in the HBCR. This study was conducted between July and August 2024 and used data collected from the HBCR for all patients treated in the Pediatric Oncology Center at Adam Malik Hospital from January

1, 2021, to December 31, 2023. Its inclusion criteria were pediatric patients aged 0–18 years newly diagnosed with cancer based on pathological anatomy criteria or other diagnostic procedures at Adam Malik Hospital between 2021 and 2023 and recorded in the HBCR from January 1, 2021, to December 31, 2023. The exclusion criteria were patients with incomplete or missing registry data.

Data collection

The HBCR data were collected and processed in several steps: (1) raw data collection, (2) data abstraction and coding, (3) data verification (first), (4) data input, and (5) data verification (second). The Pediatric Oncology Center at Adam Malik Hospital prepared a list of patients, which was filtered to remove duplicates, and patients' medical records were manually tracked to collect the raw data. The raw data then underwent data abstraction and coding, and was verified by intervariable matching. After verification, the data were entered into the HBCR, and a second verification was performed. This study included all patients registered in the HBCR who met the eligibility criteria. The data collection process was susceptible to missing data and reporting bias. A pilot test was conducted prior to the study to mitigate such bias.

Data analyses

Demographical data, including age, gender, diagnosis, and admission year, were extracted from the HBCR. The data underwent univariate analysis using Microsoft Excel and IBM SPSS (version 25.0).[7]

Ethical approval

This study was reviewed and approved by the Research Ethical Committee of the Faculty of Medicine, Universitas Sumatera Utara (clearance number: 1011/KEPK/USU/2024) and was conducted according to the Declaration of Helsinki.

RESULTS

Patients' characteristics

This study included 344 pediatric patients with cancer, consisting of 202 (58.7%) males and 142 (41.3%) females (Fig 1). Of those patients, 92 (26.7%) were admitted in 2021, 142 (41.3%) in 2022, and 110 (32%) in 2023. Their characteristics are summarized in Table 1, and their age distribution is shown in Table 2, stratified by gender and year.

Cancer incidence was highest in patients aged 2–5 years (n = 120, 34.9%) and lowest in patients aged <2 years (n = 41, 11.9%). Regarding sex, cancer incidence was highest among males aged 2–5 years (n = 80. 39.6%) and females aged 6–11 years (n = 43, 30.3%) and lowest among males and females aged <2 years.

Distribution of pediatric cancers

The distribution of pediatric cancers is presented in Table 3 (stratified by admission year) and Table 4 (stratified by age). The most common pediatric cancer was lymphoid leukemia (n = 163, 47.4%), followed by myeloid leukemia (n = 40, 11.6%) and retinoblastoma (n = 33, 9.6%). The least common pediatric cancers were germinoma of the central nervous system and Burkitt lymphoma, each with only one case reported within the three-year study period. The most common pediatric cancer in all age groups was lymphoid leukemia. Myeloid leukemia was diagnosed most often among those aged 5–11 years (n = 16, 16%). The incidence of retinoblastoma was higher among those aged <2 (n = 8, 19.5%) and 2–5 (n = 24, 20.0%) years. Burkitt lymphoma, chondrosarcoma, and renal carcinoma were reported only among those aged 12–18 years. A rare case of germinoma of the central

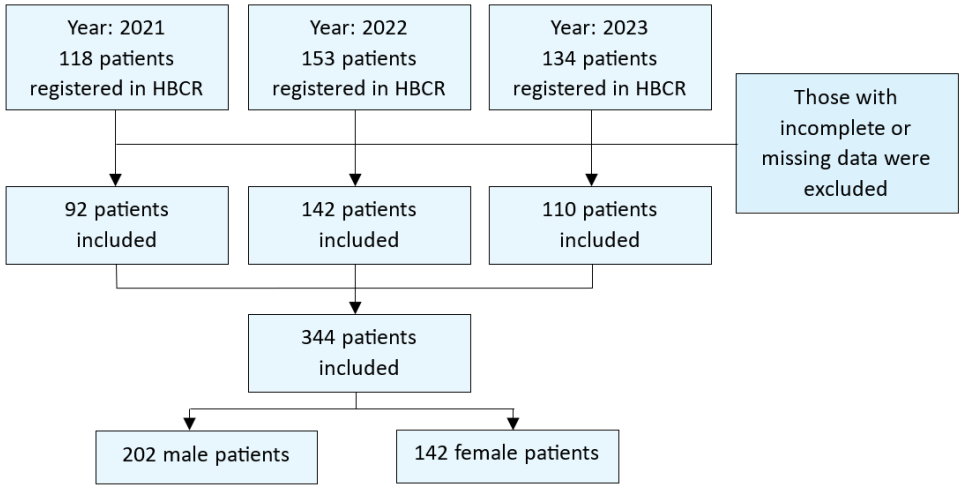


Fig 1. Flowchart of patient inclusion.

TABLE 1. Characteristics of the pediatric patients with cancer.

Characteristics	2021		2022		2023		Total	
	n: 92	%	*n*: 142	%	*n*: 110	%	*n*: 344	%
Gender								
Male	46	50	84	59.2	72	65.5	202	58.7
Female	46	50	58	40.8	38	34.5	142	41.3
Age (years)								
<2	7	7.6	23	16.2	11	10.0	41	11.9
2–5	31	33.7	50	27.6	39	35.5	120	34.9
6–11	28	30.4	41	22.7	31	28.2	100	29.1
12–18	26	28.3	28	15.5	29	26.4	83	24.1

TABLE 2. Age distribution of the pediatric patients with cancer, stratified by gender.

Age	2021				2022				2023				Total			
	Male		Female		Male		Female		Male		Female		Male		Female	
	n:46	%	*n*:46	%	*n*:84	%	*n*:58	%	*n*:72	%	*n*:38	%	*n*:202	%	*n*:142	%
<2	5	10.9	2	4.3	13	15.5	10	17.2	5	6.9	6	15.8	23	11.4	18	12.7
2–5	15	32.6	16	34.8	36	42.9	14	24.1	29	40.4	10	26.3	80	39.6	40	28.2
6–11	12	26.1	16	34.8	22	26.1	19	32.8	23	31.9	8	21.1	57	28.2	43	30.3
12–18	14	30.4	12	26.1	13	15.5	15	25.9	15	20.8	14	36.8	42	20.8	41	28.8

nervous system (n = 1, 0.8%) was reported in a patient aged 2–5 years. Osteosarcoma and Ewing sarcoma were reported only in patients aged >5 years. Three of the four cases of hepatoblastoma were found in patients aged <2 years.

Pediatric cancers by gender

The top 10 most frequent pediatric cancers by gender are shown in Fig 2. The top four predominant pediatric cancers for both genders were lymphoid leukemia, myeloid leukemia, retinoblastoma, and myelodysplastic and myeloproliferative disease. Non-Hodgkin lymphoma (except Burkitt lymphoma), osteosarcoma, and rhabdomyosarcoma were in the top ten most common pediatric cancers for both sexes. However, osteosarcoma and rhabdomyosarcoma were slightly more common in females (six cases of osteosarcoma, five cases of rhabdomyosarcoma) than in males (five cases of osteosarcoma, three cases of rhabdomyosarcoma). In contrast, non-Hodgkin lymphoma (except Burkitt lymphoma) was more slightly common in males (n = 8) than in females (n = 6 cases). Nephroblastoma and neuroblastoma were included in the top ten cancers for males, whereas hepatoblastoma and Hodgkin lymphoma were included in the top ten cancers for females.

DISCUSSION

Our study examined 344 pediatric cancer cases reported between 2021 and 2023, with 92 (26.7%) reported in 2021, 142 (41.3%) in 2022, and 110 (32%) in 2023. Throughout the study period, there was no discernible difference in the annual reports of pediatric cancer. The ratio of male to female patients was 1.4: 1 in our study, similar to those reported by Supriyadi *et al.* in Yogyakarta.[8] Previous studies have firmly established that males are more susceptible to childhood cancer than females. Males were reported to have a higher incidence of most pediatric cancers, with an overall male-to-female incidence ratio of 1.19.[9] Various factors have been reported to contribute to a higher incidence of pediatric cancers in males, including congenital abnormalities,[10] birth weight,[11] X-linked gene expression,[12] and immunological responses.[13] In 2020, the GLOBOCAN also reported that

TABLE 3. Distribution of pediatric cancers, stratified by admission year.

Cancer type	ICD-10	2021		2022		2023		Total	
		n:92	%	*n*:142	%	*n*:110	%	*n*:344	%
Lymphoid leukemia	C91	45	48.9	65	45.8	53	48.3	163	47.4
Myeloid leukemia	C92	9	9.7	19	13.4	12	10.9	40	11.6
Retinoblastoma	C69.2	9	9.7	13	9.2	11	10.0	33	9.6
Myelodysplastic and myeloproliferative disease	C94.6	3	3.3	11	7.8	9	8.3	23	6.7
Non-Hodgkin lymphoma (except Burkitt lymphoma)	C82-C86	4	4.3	8	5.6	2	1.8	14	4.1
Osteosarcoma	C41.9	5	5.4	5	3.5	1	0.9	11	3.2
Rhabdomyosarcoma	C49.9	3	3.3	3	2.1	4	3.6	10	2.9
Malignant gonadal germinal cell tumor	C59.9/ C62.9	2	2.2	4	2.8	2	1.8	8	2.3
Nephroblastoma and other nonepithelial renal tumors	C64	2	2.2	3	2.1	2	1.8	7	2.0
Neuroblastoma	C74.9	2	2.2	2	1.4	2	1.8	6	1.7
Hodgkin lymphoma	C81	2	2.2	2	1.4	0	0	4	1.2
Chondrosarcoma	C49.9	1	1.1	0	0	3	2.7	4	1.2
Extragonadal germinoma	C38.3/ C48.8/ C72.9	0	0	1	0.7	3	2.7	4	1.2
Fibrosarcoma	C49	1	1.1	2	1.4	1	0.9	4	1.2
Hepatoblastoma	C22.2	0	0	2	1.4	2	1.8	4	1.2
Skin carcinoma	C44.91/ C44.92	0	0	2	1.4	1	0.9	3	0.9
Ewing sarcoma	C40.0/ C40.1/ C40.2/ C40.3/ C41.2/ C41.3/ C41.4	1	1.1	0	0	1	0.9	2	0.6
Renal carcinoma	C64.1	1	1.1	0	0	1	0.9	2	0.6
Germinoma of the central nervous system	C72.9	1	1.1	0	0	0	0	1	0.2
Burkitt lymphoma	C83.7	1	1.1	0	0	0	0	1	0.2

TABLE 4. Distribution of pediatric cancer cases, stratified by age group.

Cancer type	ICD-10	Age group (years)							
		<2		2–5		5–11		12–18	
		n:41	%	*n*:120	%	*n*:100	%	*n*:83	%
Lymphoid leukemia	C91	14	34.2	63	52.5	53	53.0	33	39.8
Myeloid leukemia	C92	3	7.3	8	6.7	16	16.0	13	15.7
Retinoblastoma	C69.2	8	19.5	24	20.0	1	1.0	0	0
Myelodysplastic and myeloproliferative disease	C94.6	2	4.9	3	2.5	7	7.0	11	13.3
Non-Hodgkin lymphoma (except Burkitt lymphoma)	C82-C86	2	4.9	2	1.7	6	6.0	4	4.8
Osteosarcoma	C41.9	0	0	0	0	6	6.0	5	6.0
Rhabdomyosarcoma	C49.9	3	7.3	2	1.7	3	3.0	2	2.4
Malignant gonadal germinal cell tumor	C59.9/ C62.9	2	4.9	3	2.5	1	1.0	2	2.4
Nephroblastoma and other nonepithelial renal tumors	C64	1	2.4	4	3.3	1	1.0	1	1.2
Neuroblastoma	C74.9	2	4.9	3	2.5	1	1.0	0	0
Hodgkin lymphoma	C81	0	0	1	0.8	1	1.0	2	2.4
Chondrosarcoma	C49.9	0	0	0	0	0	0	4	4.8
Extragonadal germinoma	C38.3/ C48.8/ C72.9	0	0	2	1.7	0	0	2	2.4
Fibrosarcoma	C49	1	2.4	2	1.7	1	1.0	0	0
Hepatoblastoma	C22.2	3	7.3	1	0.8	0	0	0	0
Skin carcinoma	C44.91/ C44.92	0	0	1	0.8	2	2.0	0	0
Ewing sarcoma	C40.0/ C40.1/ C40.2/ C40.3/ C41.2/ C41.3/ C41.4	0	0	0	0	1	1.0	1	1.2
Renal carcinoma	C64.1	0	0	0	0	0	0	2	2.4
Germinoma of the central nervous system	C72.9	0	0	1	0.8	0	0	0	0
Burkitt lymphoma	C83.7	0	0	0	0	0	0	1	1.2

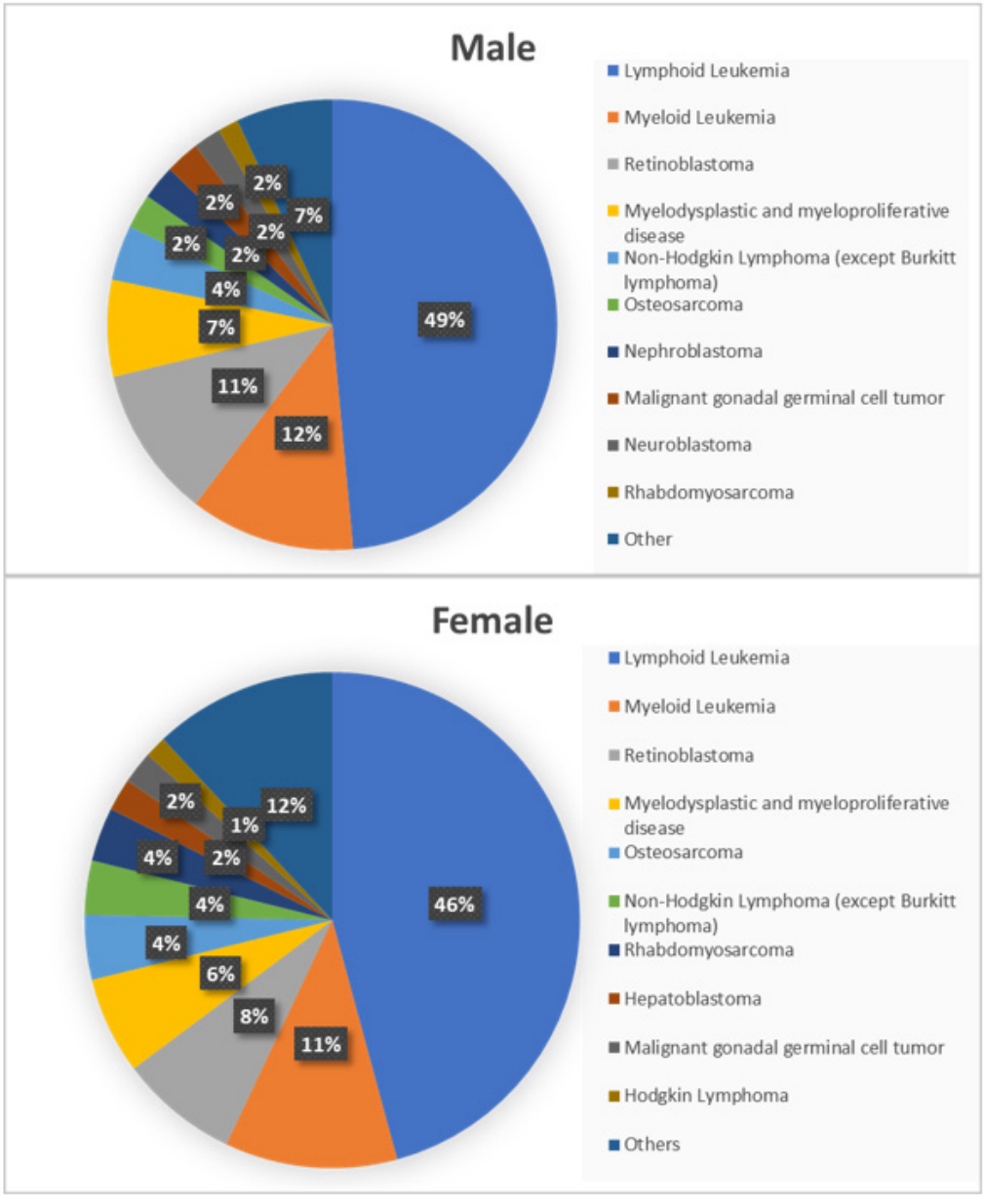


Fig 2. Top 10 cancers in male and female pediatric patients.

leukemia was more common in males than females.[14] Leukemia was the most common pediatric cancer in our study, potentially explaining the higher number of observed cases in males than females.

Cancer was reported most often in patients aged 2–5 years and least often in patients aged <2 years in our study. Cancer incidence among children and adolescents is known to vary by age.[15] Leukemia and brain tumors were found to be most commonly reported in children aged <9 years, and lymphoma, germinoma, and epithelial cancer in adolescents.[16] Leukemia was the most frequently reported cancer in our study. Therefore, the age group with the most cancer reports aligned with the age group with the highest leukemia incidence, which was reported as age 1–4 years.[17] The relatively small proportion of pediatric cancers reported in patients aged <2 years may reflect undiagnosed cases. Diagnosing cancer in very young patients is challenging for medical professionals due to the presence of nonspecific indications and symptoms, frequently resulting in misdiagnosis and delayed diagnosis.[18] There had been numerous initiatives to improve the early diagnosis of pediatric cancer. Both genetic and clinical screenings for retinoblastoma were deemed useful for early diagnosis.[19] Nonetheless, screening for nephroblastoma through the evaluation of homovanillic acid and vanillylmandelic acid excretion frequently resulted in overdiagnosis. Thus, risk-based screening had become the preferable option.[20] To comprehensively implement risk-based screening in a country, health policies governing such screening are necessary. This study demonstrated

the prevalence of pediatric cancers by age and gender, particularly among children in Indonesia. Based on this epidemiological data, a well-structured health policy regarding early diagnosis might be accomplished.[21]

The top three pediatric cancers reported in our study were lymphoid leukemia, myeloid leukemia, and retinoblastoma. This finding is consistent with data from the World Health Organization's (WHO) Global Initiative for Childhood Cancer, which reported hematopoietic neoplasms, such as leukemia, and embryonal tumors, including retinoblastoma, to be the leading pediatric malignancies.[22] Leukemia has been acknowledged as the most common pediatric cancer worldwide throughout past decades. Previous studies on childhood cancer from 2003 to 2019 reported that leukemia had the highest incidence among pediatric cancers in the United States.[16] Other studies also reported leukemia as the most common in other continents, including Asia and Australia.[23,24]

Our finding is also consistent with data from the WHO's Cancer Country Profile 2020, which indicated that leukemia was the most common cancer in those aged 0–14 years in Indonesia (2251 cases).[25] Acute lymphoblastic leukemia was the predominant leukemia subtype in pediatric patients.[26] In our study, lymphoid leukemia was the most common pediatric cancer reported in all age groups, while myeloid leukemia was most frequently reported in those aged 5–11 years. One study on leukemia trends from 1990 to 2019 demonstrated that while leukemia had the highest incidence among those aged <5 years, its overall incidence was considerably high in all age groups.[27] Our study reported 33 cases of retinoblastoma, mostly in those aged <5 years. Retinoblastoma is the most common ocular malignancy in children. Globally, 14.1 per 1 million children aged <5 years are diagnosed with retinoblastoma.[28] Retinoblastoma is remarkably diagnosed within early childhood, with a median age at diagnosis of 23.2 months.[29]

The last reported cancers in our study were germinoma of the central nervous system and Burkitt lymphoma, which is consistent with previous studies that reported a low incidence of both cancers. The incidence of germinoma of the central nervous system was 0.1 and 0.4 per 100,000 children annually in the United States and Asia, respectively.[30,31] The incidence of Burkitt lymphoma was also particularly low in Asia at <2 cases per million population.[32] Germinoma of the central nervous system is a rare malignancy of the central nervous system in children, accounting for 3% of all pediatric brain tumors.[33] It is reportedly more common in males than females, with a median age at diagnosis of approximately 13 years.[31] Burkitt lymphoma is an uncommon form of non-Hodgkin lymphoma.[32] One study reported that Burkitt lymphoma was more commonly diagnosed in males than females, with a median age at diagnosis of around 4–5 years.[34]

The distribution of pediatric cancers by age group varied in our study. For those aged <2 years, the top five pediatric cancers were lymphoid leukemia (34.2%), retinoblastoma (19.5%), myeloid leukemia (7.3%), rhabdomyosarcoma (7.3%), and hepatoblastoma (7.3%). For those aged 2–5 years, the top five pediatric cancers were lymphoid leukemia (52.5%), retinoblastoma (20%), myeloid leukemia (6.7%), nephroblastoma (3.3%), and three cancers with the same frequency (2.5%): myelodysplastic and myeloproliferative disease, malignant gonadal germinal cell tumor, and neuroblastoma. For those aged 5–11 years, the top five pediatric cancers were lymphoid leukemia (53%), myeloid leukemia (16%), myelodysplastic and myeloproliferative disease (7%), non-Hodgkin Lymphoma (except Burkitt lymphoma) (6%), and osteosarcoma (6%). For those aged 12–18 years, the top five pediatric cancers were lymphoid leukemia (39.8%), myeloid leukemia (15.7%), myelodysplastic and myeloproliferative disease (13.3%), osteosarcoma (6%), and two cancers with the same frequency (4.8%): non-Hodgkin Lymphoma (except Burkitt lymphoma) and chondrosarcoma.

Regarding gender, lymphoid leukemia, myeloid leukemia, retinoblastoma, and myelodysplastic and myeloproliferative disease were among the top four pediatric cancers reported in both genders, with an overall incidence slightly higher in males. Non-Hodgkin lymphoma (except Burkitt lymphoma), nephroblastoma, and neuroblastoma were slightly more common in males, whereas osteosarcoma, rhabdomyosarcoma, and hepatoblastoma were slightly more common in females. The incidence of lymphoma among male and female was nearly equal, consistent with findings from prior research.[35]

Leukemia, particularly lymphoid leukemia, remained the leading cancer in all age groups and genders, which is consistent with the high incidence of leukemia reported in previous studies.[18, 23-25] Retinoblastoma was mostly reported in males aged <5 years, similar to a study by Barbosa et al. that reported 91% of retinoblastoma patients were aged <4 years with a male-to-female ratio of 1.3:1, with its incidence declining with age (<4 years: 7.02 per million; 5–9 years: 0.46 per million; 10–14 years: 0.05 per million; 15–19 years: 0.03 per million).[36] Myelodysplastic and myeloproliferative disease was reported among patients aged 2–18 years and was slightly more common in males (7%) than females (6%) in our study. These findings are consistent with a previous study that reported

this disease occurred at any age and affected males and females almost equally.[37]

Rhabdomyosarcoma and hepatoblastoma were frequently reported in patients aged <2 years, with the incidence slightly higher in females than males. Both rhabdomyosarcoma and hepatoblastoma were reported to be diagnosed in early childhood. Rhabdomyosarcoma, especially embryonal rhabdomyosarcoma, was reported to be more frequent than other subtypes, with over 40% of cases involving patients aged 0–4 years.[38] Hepatoblastoma was also reported to be diagnosed with a median age of 14 months.[39] However, in our study, more rhabdomyosarcoma and hepatoblastoma cases were reported in females than males, which is inconsistent with previous studies that reported them to be more common in males (male-to-female ratio: 1.51 for rhabdomyosarcoma and 1.57 for hepatoblastoma).[38,40] This disparity was suspected to be caused by the lower overall survival of male than female patients with rhabdomyosarcoma and hepatoblastoma, leading to undiagnosed cases due to death before confirmed diagnosis.[40,41]

Osteosarcoma and Ewing sarcoma were reported in patients aged >5 years in our study, consistent with the epidemiology of these cancers, which are common in adolescents.[42] Osteosarcoma was slightly more common in females than males in our study, similar to a previous study that reported 51.2% and 48.8% of patients with osteosarcoma were female and male, respectively.[43] This study reported that nephroblastoma was more common in males (5/8 cases), and more than half of the cases involved patients aged <5 years. On average, nephroblastoma is more common in females than in males. Nonetheless, one study demonstrated that the male-to-female ratio was 1:1 in early childhood and increased with age.[44] In our study, two-thirds of neuroblastoma cases involved males, consistent with a previous study that demonstrated male preponderance, with a male-to-female ratio of 1.2.[45]

Limitations

This study was based on data from the HBCR of Indonesia's National Referral Hospital, which cancer cases were referred from primary and secondary healthcare to tertiary healthcare. Thus, we excluded cases that were not referred to the tertiary healthcare. The potential of underreporting was a limitation in this study.

Strengths and Future Aspects

This study provided important epidemiological data on pediatric cancers in Indonesia. We believe that this study could serve as a fundamental framework for the initiation of a population-based data registry in pediatric cancers. Both hospital-based and population-based data registries in Indonesia are expected to contribute to the GLOBOCAN model, thus improving the global pediatric cancer epidemiological data. By providing this epidemiological data on pediatric cancers, our expectation was that health policy regulators could establish a well-structured policy regarding pediatric cancer screening strategies, thereby improving early diagnosis. Thus, an accurate and early treatment could be given in order to reduce the mortality rate of pediatric cancers.

CONCLUSION

As a conclusion, the three most common pediatric cancers were lymphoid leukemia, myeloid leukemia, and retinoblastoma. No significant disparity was observed in the annual reports of pediatric cancers. Males predominated over females due to the substantial prevalence of leukemia cases. Most pediatric cancers were observed in patients aged 2–5 years. While the distributions of pediatric cancers varied by age and gender, leukemia remained the leading pediatric cancer.

Data Availability Statement

The data of this study is available upon request.

ACKNOWLEDGEMENTS

None

DECLARATION

Grants and Funding Information

None

Conflict of Interest

None

Author Contributions

Conceptualization and methodology, O.R.S, B.L., P.L.; Investigation, O.R.S, B.L., P.L.; Formal analysis, O.R.S, B.L., P.L.; Visualization and writing – original draft, O.R.S, B.L., P.L.; Writing – review and editing, O.R.S, B.L., P.L. All authors have read and agreed to the final version of the manuscript.

REFERENCES

1. Huang J, Chan SC, Ngai CH, Lok V, Zhang L, Lucero-Prisno DE, et al. Global incidence, mortality and temporal trends of cancer in children: A joinpoint regression analysis. Cancer Med. 2023;12(2):1903-11.
2. Force LM, Abdollahpour I, Advani SM, Agius D, Ahmadian E, Alahdab F, et al. The global burden of childhood and adolescent cancer in 2017: an analysis of the Global Burden of Disease Study 2017. Lancet Oncol. 2019;20(9):1211-25.

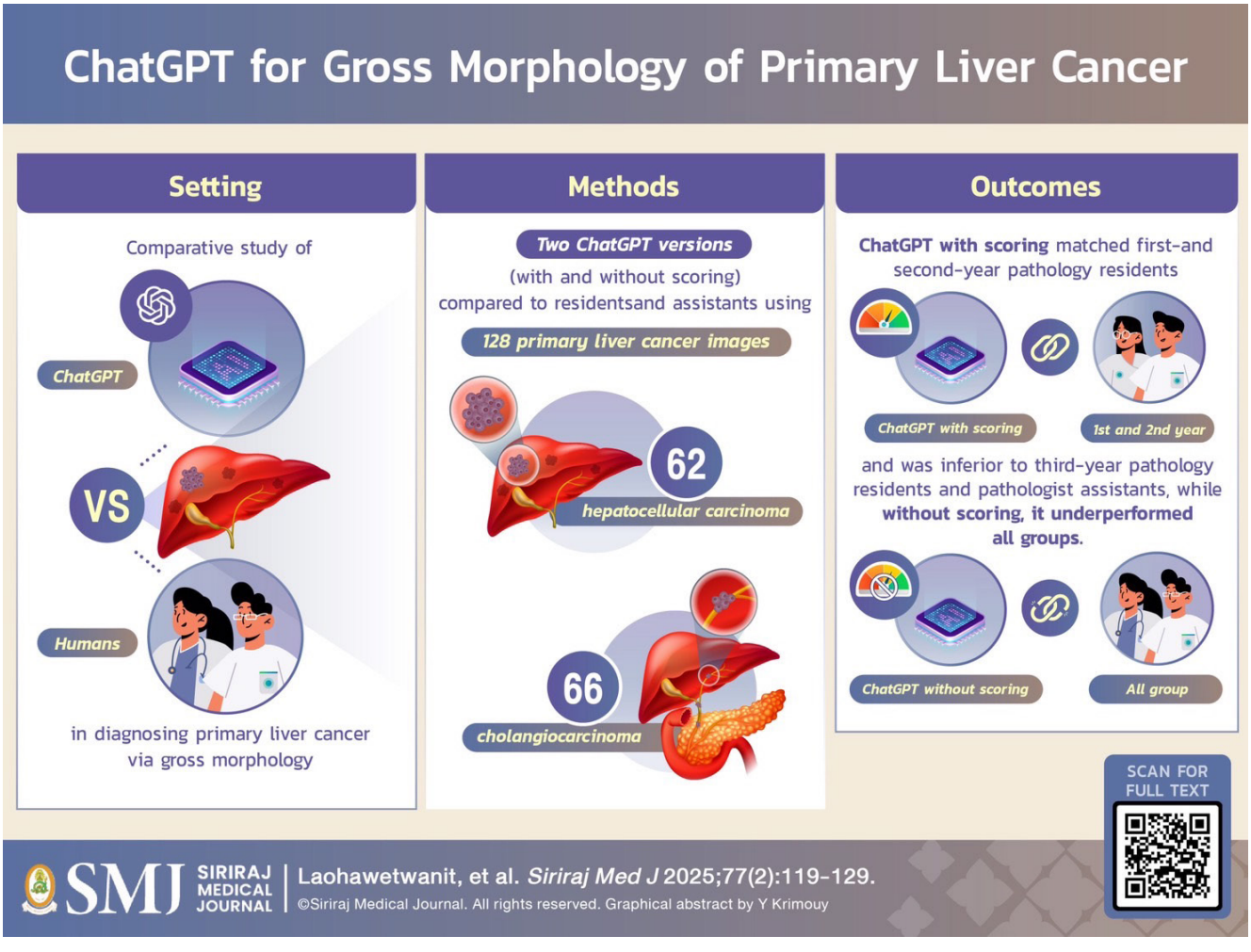
3. Health Research and Development Agency: Basic Health Research 2018. Jakarta, Indonesia, Ministry of Health Republic of Indonesia, 2018. Available from: https://repository.badankebijakan.kemkes.go.id/id/eprint/3514/1/Laporan%20Riskesdas%202018%20Nasional.pdf
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
5. Prasetya D, Layyinah A, Putri S, Rosita E, Nurjanah AI, Maftuchan A. Financial Consequences of Cancer Treatment in Indonesia: A Study Case in the Capital City of Jakarta. Jakarta: Perkumpulan PRAKARSA, 2023.
6. Gondhowiardjo S, Hartanto S, Wirawan S, Jayalie VF, Astiti IAP, Panigoro SS, et al. Treatment delay of cancer patients in Indonesia: a reflection from a national referral hospital. Med J Indones. 2021;30(2):129-37.
7. Songmuang J, Akatat P, Wattanagulanurak S, Rittayamai N, Sripalakit S. Comparison Study between Using Disposable and Non-Disposable Ventilator Circuits on Ventilator-Associated Pneumonia and Health Care Costs at a Respiratory Care Unit, Siriraj Hospital. Siriraj Med J. 2023;75(3):191-9.
8. Supriyadi E, Purwanto I, Armytasari I, Ritter J, Widjajanto PH, Veerman AJP. A province-wide childhood malignancy profiles in Indonesia (2010-2019): Yogyakarta Pediatric Cancer Registry. Paediatr Indones. 2023;63(4):226-7.
9. Williams LA, Richardson M, Marcotte EL, Poynter JN, Spector LG. Sex ratio among childhood cancers by single year of age. Pediatr Blood Cancer. 2019;66(6):e27620.
10. Marcotte EL, Schraw JM, Desrosiers TA, Nembhard WN, Langlois PH, Canfield MA, et al. Male Sex and the Risk of Childhood Cancer: The Mediating Effect of Birth Defects. JNCI Cancer Spectr. 2020;4(5):pkaa052.
11. O'Neill KA, Murphy MF, Bunch KJ, Puumala SE, Carozza SE, Chow EJ, et al. Infant birthweight and risk of childhood cancer: international population-based case control studies of 40 000 cases. Int J Epidemiol. 2015;44(1):153-68.
12. Dunford A, Weinstock DM, Savova V, Schumacher SE, Cleary JP, Yoda A, et al. Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. Nat Genet. 2017;49(1):10-16.
13. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016;16:626-38.
14. Amini M, Sharma R, Jani C. Gender differences in leukemia outcomes based on health care expenditures using estimates from the GLOBOCAN 2020. Arch Public Health. 2023;81(1):151.
15. Siegel DA, Li J, Henley SJ, Wilson RJ, Lunsford NB, Tai E, et al. Geographic Variation in Pediatric Cancer Incidence - United States, 2003-2014. MMWR Morb Mortal Wkly Rep. 2018;67(25):707-13.
16. Siegel DA, King JB, Lupo PJ, Durbin EB, Tai E, Mills K, et al. Counts, incidence rates, and trends of pediatric cancer in the United States, 2003-2019. J Natl Cancer Inst. 2023;115(11):1337-54.
17. Miranda-Filho A, Piñeros M, Ferlay J, Soerjomataram I, Monnereau A, Bray F. Epidemiological patterns of leukaemia in 184 countries: a population-based study. Lancet Haematol. 2018;5(1):e14-e24.
18. Chen J, Mullen CA. Patterns of Diagnosis and Misdiagnosis in Pediatric Cancer and Relationship to Survival. J Pediatr Hematol Oncol. 2017;39(3):e110-e15.
19. Gupta H, Malaichamy S, Mallipatna A, Murugan S, Jeyabalan N, Babu VS, et al. Retinoblastoma genetics screening and clinical management. BMC Med Genomics. 2021;14:188.
20. Mullen CJR, Barr RD, Franco EL. Timeliness of diagnosis and treatment: the challenge of childhood cancers. Br J Cancer. 2021;125(12):1612-20.
21. Arun ARK, Kumari BA, Ramachandra C, Vijay C R, Ramesh C, Apajji L, et al. National childhood cancer comprehensive management policy-a road map. Onkol Radioter. 2020;14(1):6-10.
22. Downing JR, Ghebreyesus TA. WHO Global Initiative for Childhood Cancer on the path to bridging the Survival Gap and attaining Universal Health Coverage: a 5-Year Review. WHO 2021.
23. Yang L, Yuan Y, Sun T, Li H, Wang N. Characteristics and trends in incidence of childhood cancer in Beijing, China, 2000-2009. Chin J Cancer Res. 2014;26(3):285-92.
24. Baade PD, Youlden DR, Valery PC, Hassall T, Ward L, Green AC, et al. Trends in incidence of childhood cancer in Australia, 1983-2006. Br J Cancer. 2010;102(3):620-6.
25. World Health Organization. Cancer Indones 2020 Country Profile. Available from: https://www.who.int/publications/m/item/cancer-idn-2020
26. Bhatnagar N, Qureshi A, Hall G. Leukaemias: a review. PCH. 2017;27(11):489-94.
27. Cui Y, Yan Y. The global burden of childhood and adolescent leukaemia and attributable risk factors: An analysis of the Global Burden of Disease Study 2019. J Glob Health. 2024;14:04045.
28. Stacey AW, Bowman R, Foster A, Kivelä TT, Munier FL, Cassoux N, et al. Incidence of Retinoblastoma Has Increased: Results from 40 European Countries. Ophthalmology. 2021;128(9):1369-71.
29. Global Retinoblastoma Study Group. The Global Retinoblastoma Outcome Study: a prospective, cluster-based analysis of 4064 patients from 149 countries. Lancet Glob Health. 202210(8):e1128-e40.
30. Ostrom QT, de Blank PM, Kruchko C, Petersen CM, Liao P, Finlay JL, et al. Alex's Lemonade Stand Foundation Infant and Childhood Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2007-2011. Neuro Oncol. 2015;16(Suppl 10):x1-x36.
31. Chang HY, Chiu CF, Jung SM, Wong AM, Wu CT, Lo FS. Neurological and endocrinological manifestations of 49 children with intracranial pure germinoma at initial diagnosis in Taiwan. Pediatr Neonatol. 2021;62(1):106-12.
32. Mbulaiteye SM, Devesa SS. Burkitt Lymphoma Incidence in Five Continents. Hemato 2022;3(3):434-53.
33. Ostrom QT, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014-2018. Neuro Oncol. 2021;23(12 Suppl 2):iii1-iii105.
34. Cunha KC, Oliveira MC, Gomes AC, de Castro LP, Viana MB. Clinical course and prognostic factors of children with Burkitt's lymphoma in a developing country: the experience of a single centre in Brazil. Rev Bras Hematol Hemoter. 2012;34(5):361-6.
35. Ngamdamrongkiat P, Arromdee E, Vongwiwatana A, Owattanapanich W, Sukpanichnant S. Histopathological and Clinical Features of Methotrexate-Associated Lymphoproliferative

Disorders and Post-Transplant Lymphoproliferative Disorders. Siriraj Med J. 2022;74(9):575-89.
36. Barbosa AC, de Magalhães-Barbosa MC, Moreira JPL, Colombini GNUI, Prata-Barbosa A. Incidence of retinoblastoma in children and adolescents in Brazil: A population-based study. Front Pediatr. 2022;10:1048792.
37. Hofmann, I. Pediatric myelodysplastic syndromes. J Hematopathol. 2015;8:127-41.
38. Skapek SX, Ferrari A, Gupta AA, Lupo PJ, Butler E, Shipley J, et al. Rhabdomyosarcoma. Nat Rev Dis Primers. 2019;5(1):1.
39. Tayfun KF, Akınel AN, Akbaş H, Sivrice Ç, Kintrup GT, Karagüzel G, et al. Multidisciplinary Management of Pediatric Hepatoblastoma: A 20-Year Single-Center Experience. Turk J Gastroenterol. 2022;33(12):1069-78.
40. Kahla JA, Siegel DA, Dai S, Lupo PJ, Foster JH, Scheurer ME, et al. Incidence and 5-year survival of children and adolescents with hepatoblastoma in the United States. Pediatr Blood Cancer. 2022;69(10):e29763.
41. Wang X, Feng J, Li Z, Zhang X, Chen J, Feng G. Characteristics and prognosis of embryonal rhabdomyosarcoma in children and adolescents: An analysis of 464 cases from the SEER database. Pediatr Investig. 2020;4(4):242-9.
42. Balmant NV, De Souza Reis R, De Oliveira Santos M, Maschietto M, Camargo B. Incidence and mortality of bone cancer among children, adolescents and young adults of Brazil. Clinics. 2019;74:e858.
43. Riza M, Salimo H, Wasita B, Idulhaq M, Saputra RD, Widyaningsih V, et al. Survival and prognostic factors in pediatric osteosarcoma: A 5-year single-center experience in Central Java, Indonesia. Paediatr Indones. 2024;64(1):10.
44. Doganis D, Panagopoulou P, Tragiannidis A, Georgakis MK, Moschovi M, Polychronopoulou S, et al. Childhood nephroblastoma in Southern and Eastern Europe and the US: Incidence variations and temporal trends by human development index. Cancer Epidemiol. 2018;54:75-81.
45. Farouk FS, Viqar OA, Sheikh Z, Castro G, Barengo NC. The Association between Race and Survival among Pediatric Patients with Neuroblastoma in the US between 1973 and 2015. Int J Env Res Public Health. 2020;17(14):5119.

Comparative Analysis of ChatGPT and Human Expertise in Diagnosing Primary Liver Carcinoma: A Focus on Gross Morphology

Prakasit Sa-ngiamwibool, M.D.[1], Thiyaphat Laohawetwanit, M.D.[2,*]

[1]Department of Pathology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand, [2]Division of Pathology, Chulabhorn International College of Medicine, Thammasat University, Pathum Thani, Thailand.



**Corresponding author: Thiyaphat Laohawetwanit*
E-mail: thiyapat@tu.ac.th
Received 11 October 2024 Revised 15 November 2024 Accepted 15 November 2024
ORCID ID:http://orcid.org/0000-0003-3805-7291
https://doi.org/10.33192/smj.v77i2.271596

ABSTRACT

Objective: This study aims to compare the diagnostic accuracy of customized ChatGPT and human experts in identifying primary liver carcinoma using gross morphology.

Materials and Methods: Gross morphology images of hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA) cases were assessed. These images were analyzed by two versions of customized ChatGPT (e.g., with and without a scoring system), pathology residents, and pathologist assistants. The diagnostic accuracy and consistency of each participant group were evaluated.

Results: The study analyzed 128 liver carcinoma images (62 HCC, 66 CCA), with the participation of 13 pathology residents (median experience of 1.5 years) and three pathologist assistants (median experience of 5 years). When augmented with a scoring system, ChatGPT's performance was found to align closely with first- and second-year pathology residents and was inferior to third-year pathology residents and pathologist assistants, with statistical significance (p-values < 0.01). In contrast, the diagnostic accuracy of ChatGPT, when operating without the scoring system, was significantly lower than that of all human participants (p-values < 0.01). Kappa statistics indicated that the diagnostic consistency was slight to fair for both customized versions of ChatGPT and the pathology residents. It was noted that the interobserver agreement among the pathologist assistants was moderate.

Conclusion: The study highlights the potential of ChatGPT for augmenting diagnostic processes in pathology. However, it also emphasizes the current limitations of this AI tool compared to human expertise, particularly among experienced participants. This suggests the importance of integrating AI with human judgment in diagnostic pathology.

Keywords: Artificial intelligence; ChatGPT; GPT-4; Liver cancer; Hepatocellular carcinoma; Cholangiocarcinoma
(Siriraj Med J 2025; 77: 119-129)

INTRODUCTION

Primary liver carcinoma, encompassing hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA), represents a significant global health challenge due to its high incidence and mortality rates.[1] Pathologists play a crucial role in diagnosing these malignancies, relying heavily on a combination of histopathological examination and the assessment of gross morphological features. This assessment is vital, as the initial diagnosis often depends on the visual and tactile examination of liver tissue during surgical procedures or post-surgical analysis. The gross morphology of liver carcinoma can provide critical insights into the tumor's type, size, extent, and potential invasiveness, guiding further histological examination. Accurate identification and characterization of these cancers at the gross level are essential, as they directly influence treatment decisions, prognostic evaluations, and patient outcomes.[2,3] Therefore, the precision and expertise brought by pathologists in interpreting these gross morphological features are of paramount importance in the clinical management of primary liver carcinoma.

The uses of artificial intelligence (AI) in medicine have been growing. A chatbot, such as ChatGPT, developed by OpenAI, combines a general-purpose AI system with a chat interface, allowing for natural conversations akin to those between humans. Users interact with ChatGPT by entering prompts, to which it responds contextually and quickly. The effectiveness of these interactions often depends on the precise wording of prompts, a concept known as "prompt engineering". ChatGPT excels in generating accurate responses to straightforward questions.[4] Evidence highlights ChatGPT-4's high success rate in answering USMLE practice questions and its superior performance compared to earlier AI chatbot models.[5] However, this generative AI program demonstrates promising yet limited capabilities in pathology education when being asked to answer board-style questions.[6]

The recent integration of image analysis capabilities into ChatGPT marks a significant advancement in the field of generative AI. This feature enables ChatGPT to comprehend and respond to textual inputs and analyze and interpret images. Users can now upload images directly into the conversation, allowing the AI to provide detailed descriptions, insights, and interpretations based on the visual content.[7] Data on using ChatGPT for image analysis in anatomical pathology is limited. This study aimed to compare the accuracy of the diagnosis of primary liver carcinoma based on gross morphology between customized versions of ChatGPT and human expertise.

MATERIALS AND METHODS

Gross images

Gross images of consecutive cases of HCC and CCA were obtained from the archives of the Department of Pathology, Faculty of Medicine, Khon Kaen University, capturing a broad spectrum of morphological variations in HCC and CCA. Diagnoses were confirmed by a liver pathologist (PS) through histopathologic evaluation using hematoxylin and eosin (H&E)-stained slides, with immunohistochemistry as an adjunct, to establish ground truth. Clinical data for each case was collected from an electronic medical database.

ChatGPT-4

Scoring system

ChatGPT-4 was assigned the task of developing a scoring system to differentiate the gross morphological features of HCC and CCA using the prompt, "Create a scoring system to distinguish gross morphology between HCC and CCA". The initial system included seven parameters: lesion color, texture, nodule presentation, vascular invasion, capsule formation, growth pattern, and presence of satellite lesions. Each parameter was scored on a scale from 0 to 2, representing typical features of HCC and CCA. After a review by two hepatobiliary pathologists, PS and TL, the scoring system was refined, with vascular invasion removed, leaving six criteria. The updated system was validated using 20 images of HCC and CCA.

Customization

For a uniform approach to evaluating image responses, we used two distinct configurations of ChatGPT-4. The first configuration applied a predefined scoring system before arriving at a final diagnosis. In contrast, the second configuration directly determined the final diagnosis, bypassing the scoring system. Snapshots of these configurations are presented in Fig 1.

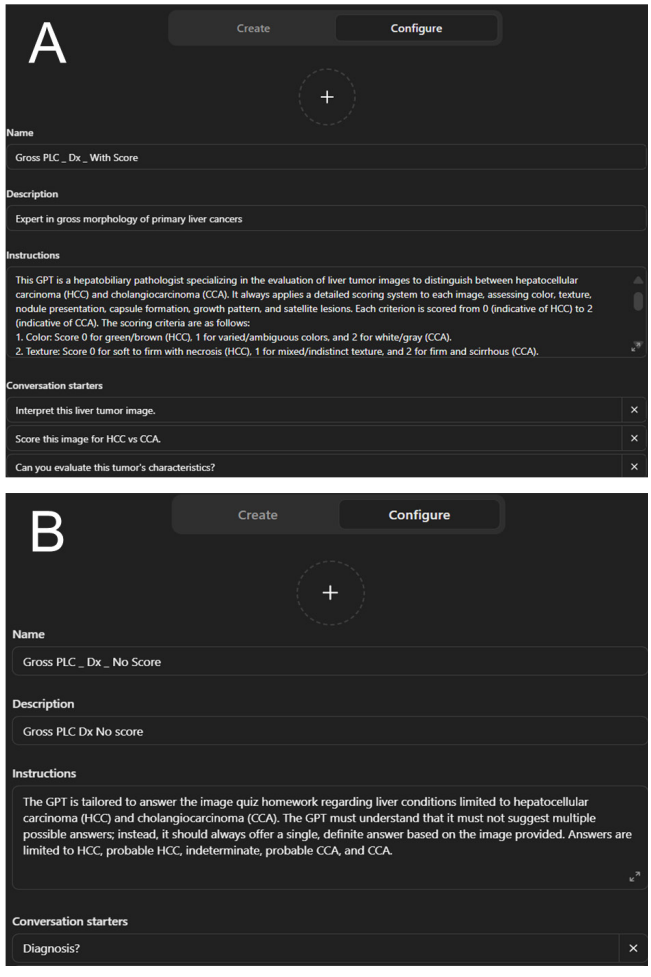


Fig 1. Snapshots of ChatGPT's configurations. (A) With scoring system. (B) Without scoring system.

Detailed guidelines for both configurations of ChatGPT-4 are provided in the following sections.

1. ChatGPT with scoring system

"This GPT is a hepatobiliary pathologist specializing in the evaluation of liver tumor images to distinguish between hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA). It always applies a detailed scoring system to each image, assessing color, texture, nodule presentation, capsule formation, growth pattern, and satellite lesions. Each criterion is scored from 0 (indicative of HCC) to 2 (indicative of CCA). The scoring criteria are as follows:

1. Color: Score 0 for green/brown (HCC), 1 for varied/ambiguous colors, and 2 for white/gray (CCA).

2. Texture: Score 0 for soft to firm with necrosis (HCC), 1 for mixed/indistinct texture, and 2 for firm and scirrhous (CCA).

3. Nodule presentation: Score 0 for single/multiple nodules (HCC), 1 for multiple nodules with unclear margins, and 2 for absence of nodules (CCA).

4. Capsule formation: Score 0 for the presence of a fibrous capsule (HCC), 1 for partial/incomplete capsule, and 2 for no capsule (CCA).

5. Growth pattern: Score 0 for expansive growth (HCC), 1 for mixed/uncertain growth patterns, and 2 for infiltrative growth (CCA).

6. Satellite lesions: Score 0 for numerous satellite lesions (HCC), 1 for a moderate number, and 2 for absent/few (CCA).

The score interpretation is as follows: 0-2 indicates definite HCC, 3-5 suggests probable HCC, 6 is indeterminate, 7-9 points towards probable CCA, and 10-12 indicates definite CCA. This method ensures a thorough and accurate analysis for each case."

2. ChatGPT without scoring system

"The GPT is tailored to answer the image quiz homework regarding liver conditions limited to hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA). The GPT must understand that it must not suggest multiple possible answers; instead, it should always offer a single, definite answer based on the image provided. Answers are limited to HCC, probable HCC, indeterminate, probable CCA, and CCA".

To evaluate the diagnostic consistency of ChatGPT, each image was sent to ChatGPT for 5 times. The responses from ChatGPT for each submission of the image were then recorded.

Human participants

Pathology residents and pathologist assistants at the Department of Pathology, Faculty of Medicine, Khon Kaen University, were tasked with diagnosing primary liver carcinoma based on gross images. While not briefed on ChatGPT-4's specific scoring system, these participants were instructed to classify each image according to categories derived from ChatGPT-4's analysis. These categories included definite HCC, probable HCC, indeterminate, probable CCA, and definite CCA.

Marking system

The accuracy of ChatGPT's responses and those of human participants were assessed against ground truths (e.g., histopathological evaluation). A marking system was established to measure the proficiency of ChatGPT and human experts in diagnosing primary liver carcinoma (Table 1).

TABLE 1. Marking system to measure the proficiency of ChatGPT and human experts in diagnosing primary liver carcinoma.

Ground truth	Response	Mark
HCC	Definite HCC	2
	Probable HCC	1
	Indeterminate	0
	Probable CCA	-1
	Definite CCA	-2
CCA	Definite CCA	2
	Probable CCA	1
	Indeterminate	0
	Probable HCC	-1
	Definite HCC	-2

Abbreviations: HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma.

Sample size justification

A sample size calculation was conducted to determine the number of images required per group for reliable comparisons of diagnostic accuracy between ChatGPT and human evaluators. Using an independent t-test with a significance level of 0.05, a power of 80%, and a moderate effect size (0.5), the analysis indicates that approximately 64 images per group are needed. This sample size provides an 80% probability of detecting a true difference in diagnostic accuracy if one exists, with a 5% chance of a Type I error (false positive).

Data analysis

Descriptive statistics were used to detail the gross image and characteristics of the participants. The performance of ChatGPT-4 and human experts was assessed through overall scores. The scores were compiled for each entity: ChatGPT with a scoring system, ChatGPT without a scoring system, residents from each year of their program, and pathologist assistants. The ANOVA test was used to calculate whether there were differences among the groups. To provide insights into the specific groups between which significant differences existed, Tukey's Honestly Significant Difference (HSD) test was used for post-hoc analysis. A p-value of less than 0.05 indicated a statistically significant difference. Interobserver agreement was assessed with kappa statistics, interpreted according to the Landis and Koch scale: values <0.00 indicate poor agreement, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement.[8]

RESULTS

Gross image and participant characteristics

The study included 128 macroscopic images of primary liver carcinoma, consisting of 62 cases of HCC and 66 cases of CCA. Thirteen pathology residents (5 first-year, 3 second-year, and 5 third-year residents) and three pathologist assistants participated in the analysis. The residents had a median of 1.5 years of pathology experience (range: 0.5 to 2.5 years) but lacked specific training in liver pathology and had limited exposure to primary liver carcinoma cases, with no recent educational activities that might have influenced their diagnostic performance. In contrast, the pathologist assistants had a median of 5 years of experience (range: 3 to 7 years), were trained in handling liver specimens, and were familiar with the gross morphology of primary liver carcinomas.

Performance of ChatGPT-4 and human participants

Snapshots of ChatGPT's performance are demonstrated in Fig 2. Participants' scores are illustrated in Fig 3. The performance of each participant group is summarized in Table 2. The customized ChatGPT that did not apply a scoring system performed worst among all evaluators. The highest mean scores were those of pathologist assistants and third-year residents.

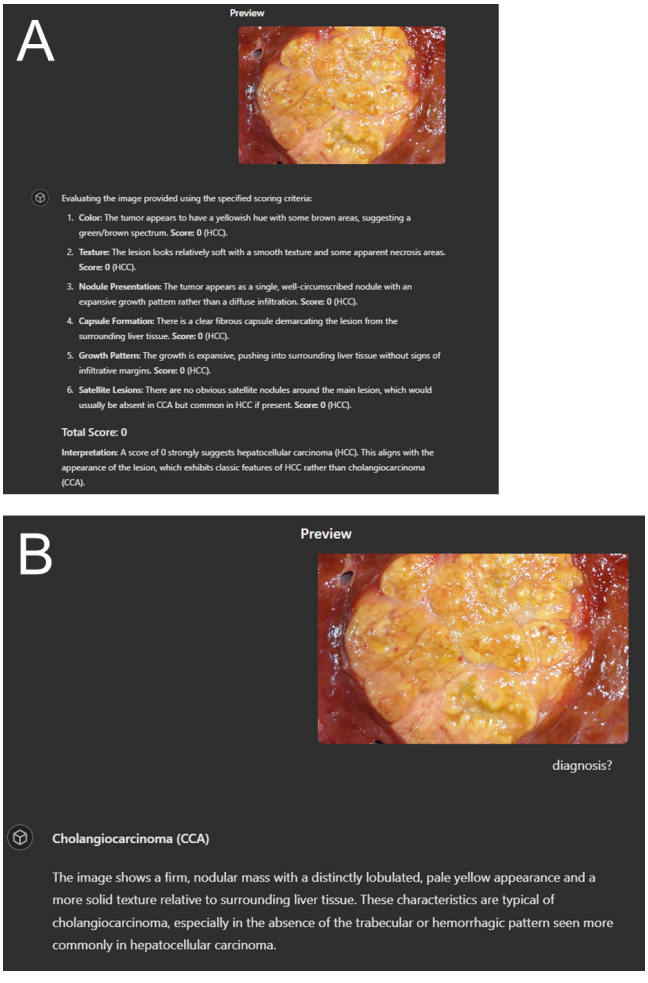


Fig 2. Snapshots of ChatGPT's responses. (A) With scoring system. (B) Without scoring system.

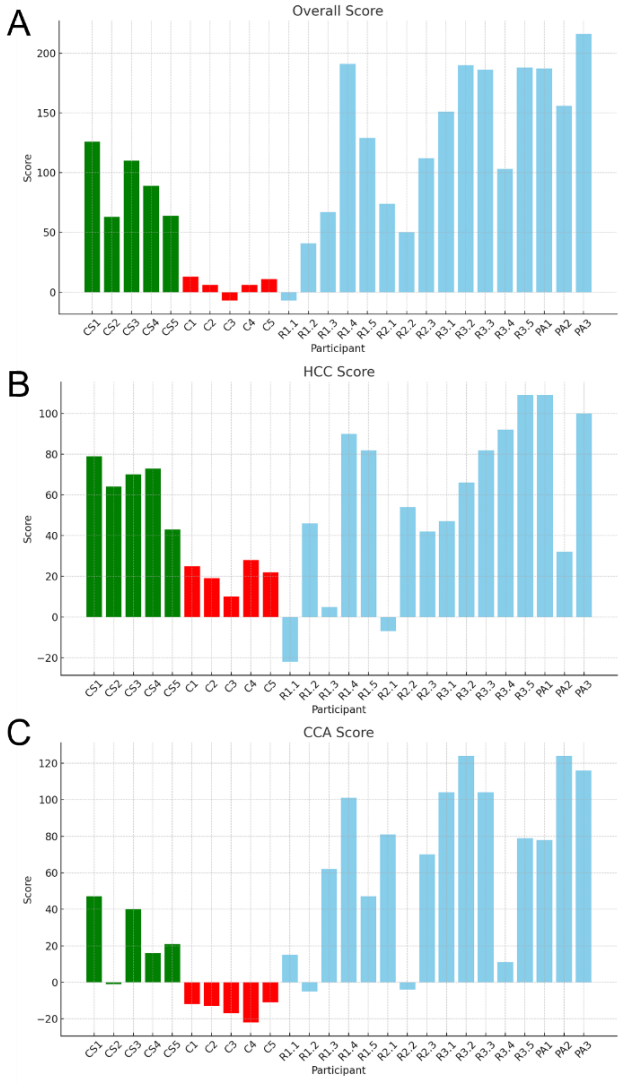


Fig 3. Scores of each participant. Bar charts demonstrating (A) overall scores, (B) hepatocellular carcinoma (HCC) scores, and (C) cholangiocarcinoma (CCA) score.
Abbreviations: CS = ChatGPT with a scoring system; C = ChatGPT without a scoring system; R1 = First-year pathology residents; R2 = Second-year pathology residents; R3 = Third-year pathology residents; PA = Pathologist assistants.

TABLE 2. Performance of ChatGPT 4 and human participants.

Evaluators	Mean	SD	Min	Max
ChatGPT				
Scoring system	90.4	24.9	63	123
No scoring system	5.8	7	-7	13
Human participants				
First-year residents	84.2	69.1	-7	191
Second-year residents	78.7	25.5	50	112
Third-year residents	163.6	33.5	103	190
Pathologist assistants	186.3	24.5	156	216

Abbreviations: SD = Standard deviation; Min = Minimum; Max = Maximum.

Comparison of scores among ChatGPT-4 and human participants

The Shapiro-Wilk test and Levene's test were applied to assess normality and homogeneity of variances, respectively, yielding p-values above 0.05 before performing the ANOVA test. The ANOVA showed an F-statistic of 9.804 and a p-value below 0.01. A post-hoc Tukey's HSD test was then conducted to identify specific group differences (Table 3). In scoring comparisons, ChatGPT's average score was similar to those of first- and second-year pathology residents but was lower than those of third-year residents and pathologist assistants (p-value < 0.01). Notably, when ChatGPT operated without a scoring system, its overall performance was significantly lower than that of all human participant groups (p-value < 0.01). The data also indicated a trend of higher scores among human participants as their pathology experience increased.

A subgroup analysis focusing on comparing the diagnostic abilities of HCC and CCA among different participant groups was performed. In assessing the HCC diagnostic scores, the ANOVA test revealed an F-statistic of 3.10 and a p-value of 0.03. Despite this, none of the group-to-group comparisons showed statistically significant differences in average scores, as all p-values exceeded 0.05. This suggests that while there was some overall variance among the groups, no single group differed significantly.

Regarding CCA diagnostic scores, the ANOVA test indicated a more pronounced difference with an F-statistic of 7.22 and a p-value below 0.01. Detailed pairwise comparisons revealed significant score differences between several groups: ChatGPT with a scoring system and pathologist assistants (p-value = 0.03), ChatGPT without a scoring system and pathologist assistants (p-value < 0.01), and ChatGPT without a scoring system and third-year pathology residents (p-value < 0.01). Conversely, the remaining group comparisons did not show statistically significant differences.

TABLE 3. Multiple comparison of means by the Tukey's HSD test.

Group 1	Group 2	Mean difference	P-value	Lower CI	Upper CI	Reject
CS	C	-84.60	<0.01*	-115.59	-53.61	TRUE
CS	PA	95.93	<0.01*	64.95	126.92	TRUE
CS	R1	-6.20	0.9	-37.19	24.79	FALSE
CS	R2	-11.73	0.9	-42.72	19.26	FALSE
CS	R3	73.20	<0.01*	42.21	104.19	TRUE
C	PA	180.53	<0.01*	149.55	211.52	TRUE
C	R1	78.40	<0.01*	47.41	109.39	TRUE
C	R2	72.87	<0.01*	41.88	103.86	TRUE
C	R3	157.80	<0.01*	126.81	188.79	TRUE
PA	R1	-102.13	<0.01*	-133.12	-71.14	TRUE
PA	R2	-107.67	<0.01*	-138.66	-76.68	TRUE
PA	R3	-22.73	0.69	-53.72	8.26	FALSE
R1	R2	-5.53	0.9	-36.52	25.46	FALSE
R1	R3	79.40	<0.01*	48.41	110.39	TRUE
R2	R3	84.93	<0.01*	53.95	115.92	TRUE

The table shows the mean differences between the scores of each pair of groups, along with the p-value for each comparison. The lower and upper bounds of the 95% confidence interval for the mean difference are presented. If the "reject" column is true, it indicates that there is a statistically significant difference between those groups.

Abbreviations: CS = ChatGPT with scoring system; C = ChatGPT without scoring system; PA = Pathologists assistants; R1 = First-year pathology residents; R2 = Second-year pathology residents; R3 = Third-year pathology residents; CI = Confidence interval.

* = Statistically significant difference.

Diagnostic consistency

Heat maps showing diagnostic consistency among participants are presented in Fig 4. Kappa statistics with interpretations of agreement level are shown in Table 4. Similar to pathology residents, both customized ChatGPT versions provided slight to fair diagnostic consistency. Notably, the level of agreement in the pathologist assistant group was moderate, higher than that of other participant groups. A subgroup analysis was conducted to compare the diagnostic consistency for HCC and CCA across different participant groups. While the use of ChatGPT without a scoring system resulted in moderate diagnostic consistency, other groups demonstrated only slight agreement in diagnosing both HCC and CCA.

Qualitative analysis of images with consistent responses

Images that consistently received definitive diagnoses from participants were subjected to further analysis. Out of these, 35 images (56.5%) of HCC and 20 images (30.3%) of CCA were classified as "definite" for HCC and CCA, respectively, by over half of the participants. Specifically, the three HCC images that received the highest number of "definite" diagnoses (ranging from 18 to 20 out of 26 responses) depicted either a mass in a cirrhotic background (Fig 5A) or a solitary nodule exhibiting expansive growth (Fig 5B). Similarly, the three CCA images with the most definitive diagnoses (16 to 21 responses) displayed characteristics such as an intraductal growth pattern (Fig 5C) or a combination of periductal infiltrating and mass forming growth patterns (Fig 5D). The HCC image that most frequently received reversed diagnoses of CCA (12 responses) presented a large, protruding mass with a minimal amount of non-neoplastic liver tissue (Fig 5E). Conversely, the CCA image most commonly mistaken for HCC (11 responses) predominantly showed a mass forming growth pattern with focal areas of periductal infiltration (Fig 5F).

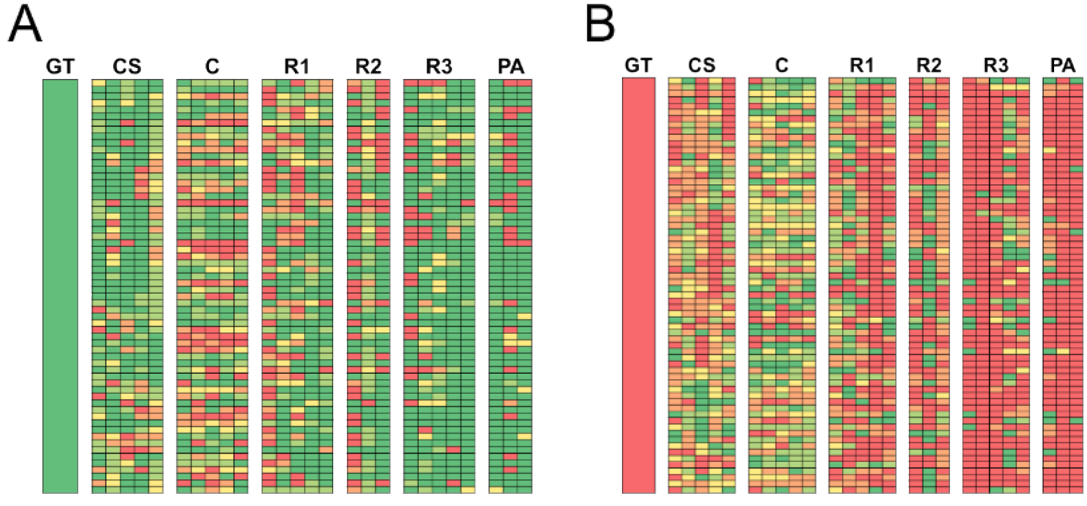


Fig 4. Diagnostic consistency. Heat maps showing diagnostic consistency of (A) hepatocellular carcinoma (HCC) and (B) cholangiocarcinoma (CCA) among each participant group. Different colors were used to show diagnostic confidence levels for HCC and CCA as follows: green for definite HCC, light green for probable HCC, yellow for indeterminate, orange for probable CCA, and red for definite CCA.
Abbreviations: GT = Ground truth; CS = ChatGPT with a scoring system; C = ChatGPT without a scoring system; R1 = First-year pathology residents; R2 = Second-year pathology residents; R3 = Third-year pathology residents; PA = Pathologist assistants.

TABLE 4. Diagnostic consistency of each participant group.

Participant group	Kappa	Level of agreement
ChatGPT with a scoring system	0.07	Slight
ChatGPT without a scoring system	0.25	Fair
First-year residents	0.06	Slight
Second-year residents	0.02	Slight
Third-year residents	0.31	Fair
Pathologist assistants	0.46	Moderate

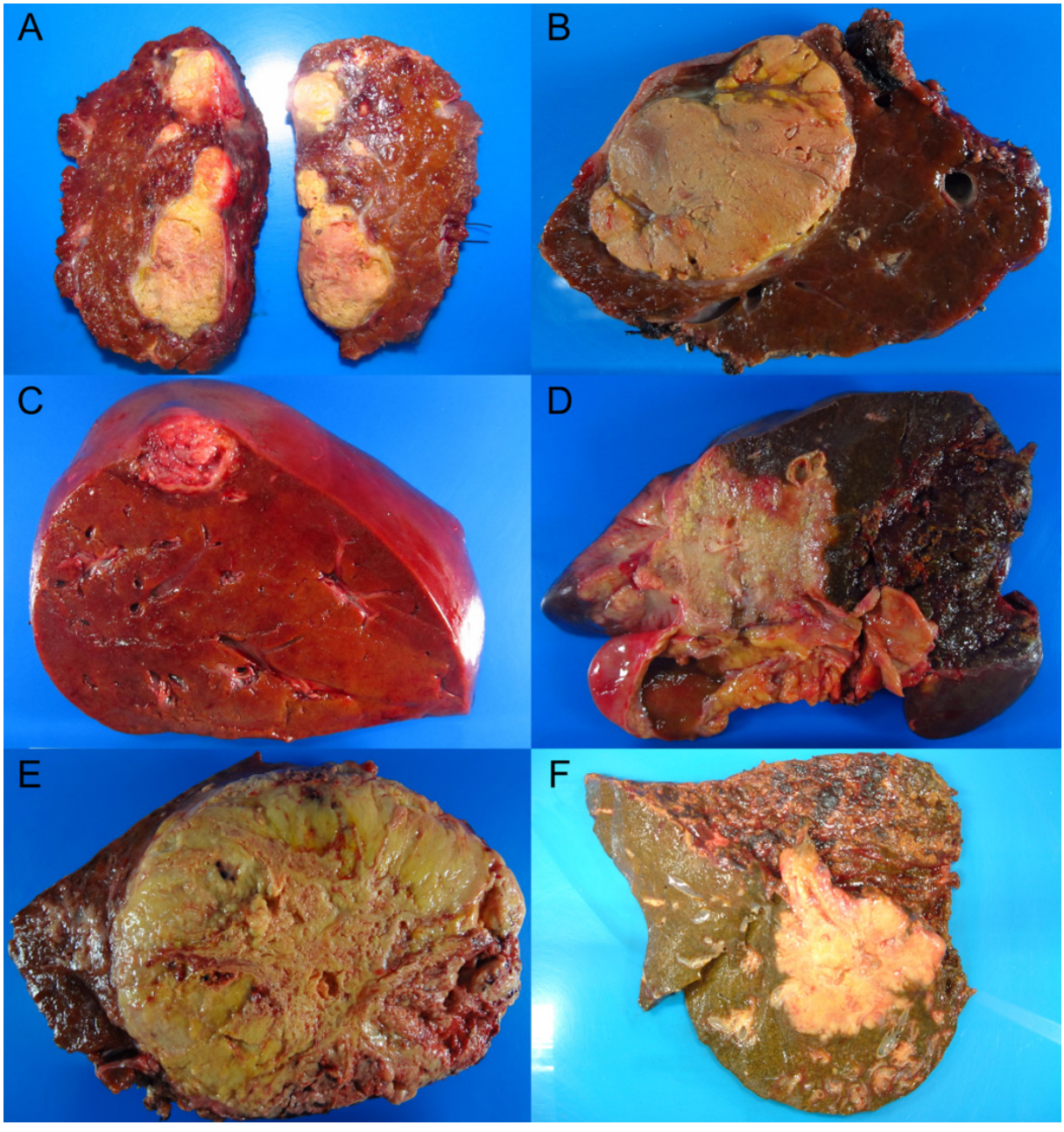


Fig 5. Images with consistent responses. Well-demarcated masses in a cirrhotic background (A) and a single nodule with an expansive growth (B) are usually diagnosed as hepatocellular carcinoma. Masses with intraductal (C) and unequivocal periductal infiltrating (D) growth patterns are commonly regarded as cholangiocarcinoma. A large, bulging mass with limited non-neoplastic liver tissue is sometimes mistaken for cholangiocarcinoma (E). Cholangiocarcinoma with a predominant mass-forming growth can be misidentified as hepatocellular carcinoma (F).

DISCUSSION

Our study provides a comparative analysis of ChatGPT and human experts in diagnosing primary liver carcinoma based on gross morphology. Key findings reveal that both versions of ChatGPT—one using a scoring system and one without—showed limited accuracy and consistency in diagnosing primary liver carcinoma through gross morphology alone. ChatGPT with a scoring system performed at a level similar to that of first- and second-year pathology residents but was less effective than more experienced residents and pathologist assistants. Notably, ChatGPT without a scoring system had the lowest performance among all evaluators. The slight to fair agreement observed among ChatGPT versions is inadequate for practical application, contrasting with the moderate agreement level among pathologist assistants with experience in gross morphology of primary liver carcinoma. These findings highlight the current limitations of AI in interpreting gross morphology in neoplastic diseases, underscoring the essential role of human expertise in this domain.

As of this writing, there are limited data on ChatGPT's use for analyzing pathological images, and the available literature shows inconsistent results. While ChatGPT-4 demonstrates high sensitivity in identifying adenomas, it struggles with low specificity, and its performance in polyp classification remains variable. One major concern is its lack of diagnostic consistency.[9] Additionally, its histopathological evaluations for common diseases are less accurate than those of pathology trainees.[10] However, applying step-by-step reasoning (also known as chain-of-thought prompting) has been shown to improve its performance.[11] In one study, ChatGPT was able to accurately grade and stage fatty liver disease using H&E and Masson's trichrome-stained images. Yet, caution is needed when interpreting these results, as they relied on a small sample of publicly available images (e.g., from PathologyOutlines.com) during a time when ChatGPT had internet browsing capabilities.[12,13] Another study reports that ChatGPT outperformed expert technicians in distinguishing between normal and abnormal blood cells.[14] However, the applicability of these findings is constrained by the specific types of images (e.g., histopathology, peripheral blood smears) utilized in the research. Despite high expectations among pathologists for AI chatbots to excel in image analysis,[15] these expectations has yet to be realized. The current study suggests that integrating a scoring system could improve ChatGPT-4's diagnostic accuracy.

ChatGPT, a cutting-edge chatbot developed by OpenAI, stands at the forefront of natural language processing technology.[16] Built on the sophisticated GPT-4 architecture, this chatbot is adept at generating text that closely resembles human conversation, making it incredibly versatile and user-friendly. Its training involves extensive analysis of diverse text data, enabling it to perform a variety of tasks such as providing informed responses, drafting emails, generating creative content, and even programming.[17] In 2023, ChatGPT underwent significant advancements, most notably the introduction of image processing. These features, enabled by GPT-4V and updated DALL-E models, allowed users to engage with ChatGPT through images, expanding its applications beyond text.[7] Additionally, the custom instructions feature was introduced, allowing users to personalize their interactions with ChatGPT to meet specific needs.[18] These developments not only enhanced user experience but also increased the tool's adaptability across various industries, highlighting its growing importance in AI-driven solutions and its potential for widespread impact in different sectors.

Research has demonstrated that ChatGPT-4 effectively answers a significant proportion of practice questions for the USMLE examinations[5,19] and offers accurate and consistent responses to materials designed for pathology examination preparation.[20] Additionally, a study has shown that ChatGPT can respond to pathology-specific inquiries on par with or surpass the performance of a trained pathologist, depending on the version of the model (GPT-3.5 or GPT-4).[21] Despite these findings, it is crucial to recognize that these outcomes do not imply ChatGPT's equivalence to pathologists in diagnostic abilities. The studies specifically exclude questions that require the interpretation of images. Furthermore, the nature of the examination questions, which tend to have clear and definitive answers, differs significantly from real-world diagnostic scenarios that demand complex decision-making.[22]

In a comparative study between ChatGPT and pathologists regarding histopathology diagnosis and collaborative potential, ChatGPT's contributions were deemed somewhat helpful but their effectiveness varied, contingent on the pathologists' prompts.[23] This finding highlights the importance of prompt engineering. Prompt engineering is essential in the AI field, particularly for advanced language models ChatGPT. It involves crafting questions or instructions to elicit precise and relevant responses from AI. Effective prompts also help mitigate misinterpretation and bias, leading to more reliable AI outcomes.[24]

Our study introduces a novel method for assessing primary liver carcinoma by comparing ChatGPT's image analysis capabilities with human pathologist expertise. A key strength lies in the innovative use of generative AI to create a scoring system, subsequently reviewed by liver pathologists—a first in this field. Similar research in radiology, applying AI to evaluate gross tumor volume and grade prognostic assessments of HCC, has shown promising results.[25,26] Additionally, the study pioneers AI analysis of gross morphological images of liver carcinomas, directly comparing AI and human assessments. This approach highlights AI's potential to support diagnostic processes in pathology while offering valuable insights into AI's current capabilities and limitations in medical image interpretation. By including both AI and human evaluators with diverse experience levels and no standardized training, the research methodology provides a realistic comparison, emphasizing areas where AI can complement human judgment and where human expertise remains indispensable.

While this study introduces a novel approach, several limitations affect its broader applicability. First, its focus on the gross morphology of primary liver carcinomas

may restrict the findings' relevance to other diseases and histopathologic evaluations. Incorporating microscopic findings, expert-provided descriptions, or other AI models might yield more promising outcomes.[10] Second, as the study was conducted within a single institution, the specialized expertise of pathology residents and pathologist assistants in cholangiocarcinoma may have influenced the results. Third, ChatGPT's current lack of annotation capabilities in image analysis could have impacted diagnostic precision; with the tool evolving rapidly, these findings may soon become outdated. Additionally, using other commercially available generative AI models, such as Gemini (formerly Bard) (Google, California, USA) and Claude (Anthropic, California, USA), might yield different outcomes. However, a recently published study has suggested that Bard underperformed compared to ChatGPT in pathology image interpretation.[27] These factors limit the generalizability of the study and highlight areas for further research and improvement.

This study offers several implications for pathology practice and future research. First, it highlights ChatGPT's potential to enhance human expertise in pathology, especially in analyzing and interpreting gross morphological features. ChatGPT's development of a scoring system represents a step toward integrating AI into diagnostic frameworks. Future research could benefit from advancing AI-based diagnostic criteria that incorporate both gross and microscopic features for more comprehensive diagnoses. Another promising application is using ChatGPT to suggest or interpret immunohistochemistry in challenging cases.[28] Additionally, expanding research to encompass various neoplastic and non-neoplastic diseases and testing across different institutions with diverse disease prevalences could provide valuable insights into the broader applicability and effectiveness of AI tools in pathology. It is important not to overstate ChatGPT's potential. As a fundamentally text-based large language model, ChatGPT generally excels with text-based tasks, such as generating pathological diagnoses from detailed microscopic descriptions or creating educational content,[10,29] rather than with tasks requiring pure image interpretation.[9-11]

In conclusion, this study underscores the evolving role of AI in pathology, particularly in diagnosing primary liver carcinoma using gross morphology. While ChatGPT exhibits potential in augmenting diagnostic processes, its performance, particularly with a scoring system, remains inferior to experienced pathology residents and pathologist assistants. The findings emphasize the importance of combining AI tools with human judgment in diagnostic pathology.

Data Availability Statement

Data are available on reasonable request.

ACKNOWLEDGEMENTS

The authors would like to thank pathology residents and pathologist assistants at the Department of Pathology, Faculty of Medicine, Khon Kaen University, for providing diagnoses of images used in this study. The authors thank Mr. Suwit Balthaisong, Department of Pathology, Khon Kaen University for his skillful assistance gross images.

DECLARATION

Grants and Funding Information

This study received no funding support.

Conflict of Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Both authors involved in conception, design of the study, acquisition and analysis of data, drafting and reviewing the manuscript.

Use of Artificial Intelligence

ChatGPT-4 was utilized to refine the English language in this manuscript. However, the accuracy and validity of the content are solely the responsibility of the authors.

Ethics approval statement

Ethical approval was provided by the Khon Kaen University Ethics Committee for Human Research based on the Declaration of Helsinki (Number HE671129).

REFERENCES

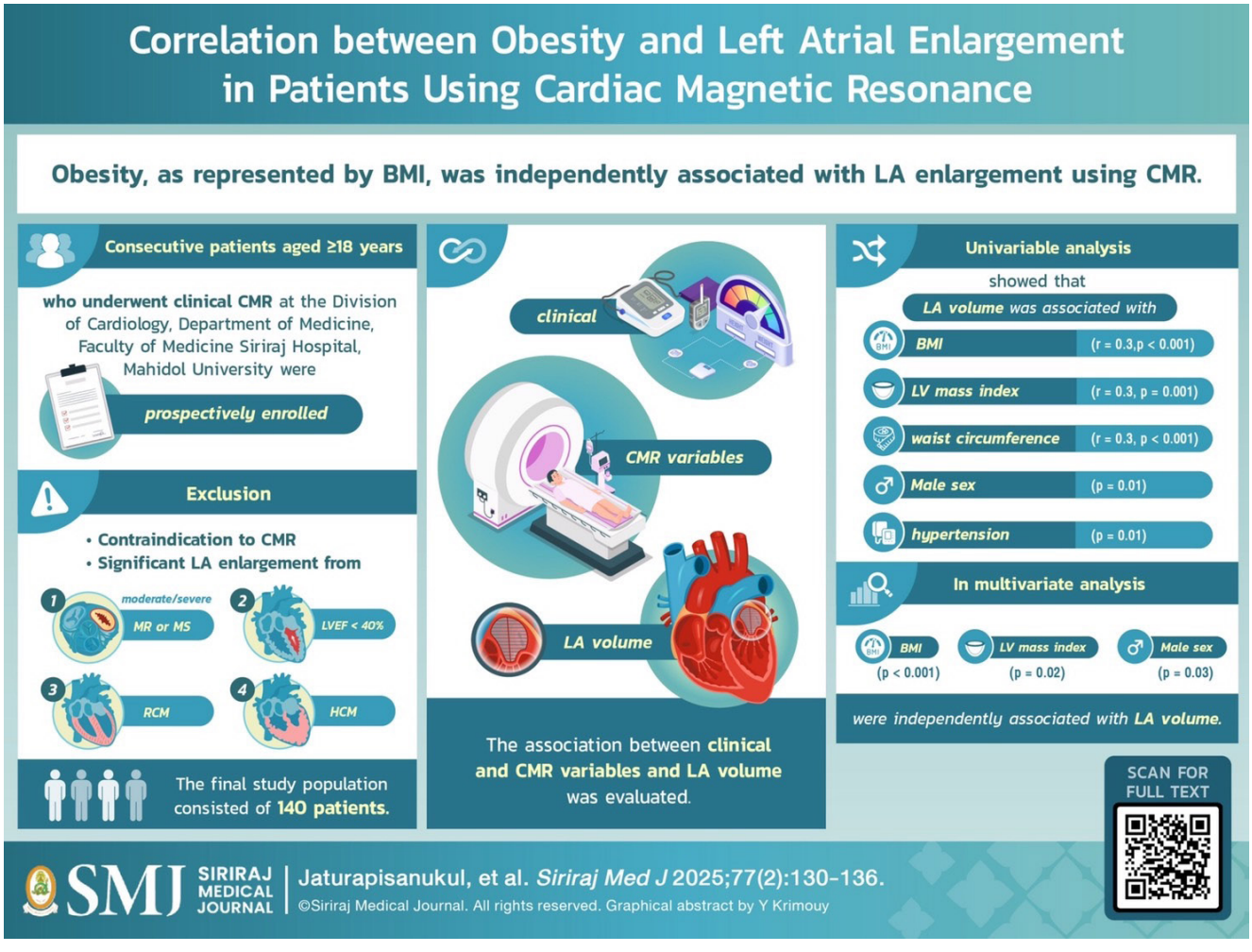
1. Rumgay H, Ferlay J, de Martel C, Georges D, Ibrahim AS, Zheng R, et al. Global, regional and national burden of primary liver cancer by subtype. Eur J Cancer. 2022;161:108-18.
2. Fan Z, Jin M, Zhang L, Wang N, Li M, Wang C, et al. From clinical variables to multiomics analysis: a margin morphology-based gross classification system for hepatocellular carcinoma stratification. Gut. 2023;72(11):2149-63.
3. Zen Y. Intrahepatic cholangiocarcinoma: typical features, uncommon variants, and controversial related entities. Hum Pathol. 2023;132:197-207.
4. Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. N Engl J Med. 2023;388(13): 1233-9.
5. Mihalache A, Huang RS, Popovic MM, Muni RH. ChatGPT-4: An assessment of an upgraded artificial intelligence chatbot in the United States Medical Licensing Examination. Med Teach. 2023:1-7.

6. Geetha SD, Khan A, Khan A, Kannadath BS, Vitkovski T. Evaluation of ChatGPT pathology knowledge using board-style questions. Am J Clin Pathol. 2024;161(4):393-98.
7. OpenAI. GPT-4V(Ision) system card2023. Available from: https://cdn.openai.com/papers/GPTV_System_Card.pdf.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
9. Laohawetwanit T, Namboonlue C, Apornvirat S. Accuracy of GPT-4 in histopathological image detection and classification of colorectal adenomas. J Clin Pathol. 2024:jcp-2023-209304.
10. Apornvirat S, Thinpanja W, Damrongkiet K, Benjakul N, Laohawetwanit T. Comparing customized ChatGPT and pathology residents in histopathologic description and diagnosis of common diseases. Ann Diagn Pathol. 2024;73:152359.
11. Laohawetwanit T, Apornvirat S, Namboonlue C. Thinking like a pathologist: Morphologic approach to hepatobiliary tumors by ChatGPT. Am J Clin Pathol. 2024:aqae087.
12. Zhang Y, Liu H, Sheng B, Tham YC, Ji H. Preliminary Fatty Liver Disease Grading Using General-Purpose Online Large Language Models: ChatGPT-4 or Bard? J Hepatol. 2024;80(6): e279-81.
13. OpenAI. ChatGPT plugins2023. Available from: https://openai.com/blog/chatgpt-plugins.
14. Yang WH, Yang YJ, Chen TJ. ChatGPT's innovative application in blood morphology recognition. J Chin Med Assoc. 2024;87(4): 428-33.
15. Laohawetwanit T, Pinto DG, Bychkov A. A survey analysis of the adoption of large language models among pathologists. Am J Clin Pathol. 2024:aqae093.
16. OpenAI. Introducing ChatGPT2022. Available from: https://openai.com/blog/chatgpt.
17. OpenAI. GPT-4 Technical Report2023. Available from: https://arxiv.org/pdf/2303.08774.pdf.
18. OpenAI. Introducing GPTs2023. Available from: https://openai.com/blog/introducing-gpts.
19. Brin D, Sorin V, Vaid A, Soroush A, Glicksberg BS, Charney AW, et al. Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. Sci Rep. 2023;13(1):16492.
20. Koga S. Exploring the pitfalls of large language models: Inconsistency and inaccuracy in answering pathology board examination-style questions. Pathol Int. 2023;73(12):618-20.
21. Wang AY, Lin S, Tran C, Homer RJ, Wilsdon D, Walsh JC, et al. Assessment of Pathology Domain-Specific Knowledge of ChatGPT and Comparison to Human Performance. Arch Pathol Lab Med. 2024.
22. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model. Res Sq. 2023.
23. Oon ML, Syn NL, Tan CL, Tan KB, Ng SB. Bridging bytes and biopsies: A comparative analysis of ChatGPT and histopathologists in pathology diagnosis and collaborative potential. Histopathology. 2024;84(4):601-13.
24. Mesko B. Prompt Engineering as an Important Emerging Skill for Medical Professionals: Tutorial. J Med Internet Res. 2023; 25:e50638.
25. Zhang HW, Huang DL, Wang YR, Zhong HS, Pang HW. CT radiomics based on different machine learning models for classifying gross tumor volume and normal liver tissue in hepatocellular carcinoma. Cancer Imaging. 2024;24(1):20.
26. Gross M, Haider SP, Ze'evi T, Huber S, Arora S, Kucukkaya AS, et al. Automated graded prognostic assessment for patients with hepatocellular carcinoma using machine learning. Eur Radiol. 2024;34(10):6940-52.
27. Apornvirat S, Namboonlue C, Laohawetwanit T. Comparative analysis of ChatGPT and Bard in answering pathology examination questions requiring image interpretation. Am J Clin Pathol. 2024;162(3):252-60.
28. Sukpanichnant S. Malignancy of the lymph node: How general practitioners and pathologists can achieve a definitive diagnosis. Siriraj Med J. 2022;74(9):604–17.
29. Laohawetwanit T, Apornvirat S, Kantasiripitak C. ChatGPT as a teaching tool: Preparing pathology residents for board examination with AI-generated digestive system pathology tests. Am J Clin Pathol. 2024;162(5):471-9.

Correlation between Obesity and Left Atrial Enlargement in Patients Using Cardiac Magnetic Resonance

Solos Jaturapisanukul, M.D.[1,2], Yodying Kaolawanich, M.D.[1,3], Methat Meechuen, M.S.[1,3], Thananya Boonyasirinant, M.D.[1,3,*]

[1]Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Division of Nephrology, Department of Internal Medicine, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand, [3,]Division of Cardiology, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*



**Corresponding author: Thananya Boonyasirinant*
E-mail: drthananyaa@yahoo.com
Received 28 October 2024 Revised 16 November 2024 Accepted 18 November 2024
ORCID ID:http://orcid.org/0000-0002-9149-3407
https://doi.org/10.33192/smj.v77i2.271916

ABSTRACT

Objective: The correlation between left atrial (LA) enlargement and obesity has been previously reported. However, most studies primarily evaluated LA diameter using echocardiography, which is less accurate and reliable than assessing LA volume with cardiac magnetic resonance (CMR). This study aimed to explore the correlation between obesity and LA volume by using CMR imaging.

Materials and Methods: We prospectively enrolled consecutive eligible patients aged 18 years or older who underwent CMR at a tertiary academic hospital. Clinical variables, including body mass index (BMI), were collected from the medical records. LA volume classification was based on the current guideline recommendations.

Results: A total of 140 patients (41% men), with a mean age of 66.5 ± 10.5 years, were studied. The mean BMI was 25.7 ± 4.2 kg/m^2. CMR parameters revealed an LA volume of 75.7 ± 22.5 mL, a left ventricular (LV) ejection fraction of 71.1 ± 9.5%, and an LV mass index of 46.1 ± 27.4 g/m^2. Univariable analysis indicated that the factors affecting LA volume included BMI ($r = 0.3$, $p < 0.001$), LV mass index ($r = 0.3$, $p = 0.001$), waist circumference ($r = 0.3$, $p < 0.001$), male sex ($p = 0.01$), and hypertension ($p = 0.01$). In stepwise multivariable analysis, BMI ($p < 0.001$), LV mass index ($p = 0.02$), and male sex ($p = 0.03$) were independently associated with LA volume.

Conclusion: Obesity, as represented by BMI, was independently associated with LA enlargement. Other independent factors correlated with LA volume included the LV mass index and male sex.

Keywords: Cardiac magnetic resonance; correlation; left atrial enlargement; left atrium; left atrial volume; obesity; overweight. (Siriraj Med J 2025; 77: 130-136)

Abbreviations

BMI = body mass index
CMR = cardiac magnetic resonance
LA = left atrium
LV = left ventricle
LVEDV = left ventricular end-diastolic volume
LVEF = left ventricular ejection fraction
LVESV = left ventricular end-systolic volume
SD = standard deviation
SSFP = steady-state free precession

INTRODUCTION

According to the World Health Organization (WHO) expert consultation, individuals are generally considered obese if their body mass index (BMI) is 25 kg/m^2 or higher, with a range of 23–24.9 kg/m^2 defined as overweight in Asia-Pacific populations.[1] Patients with obesity are at an increased risk of various health problems.

Obesity is one of the primary preventable causes of death globally, with rising prevalence among both adults and children. The latest WHO report indicates that adult obesity rates worldwide have more than doubled since 1990, while the rates of adolescent obesity have increased fourfold. In 2022, one in eight people worldwide lived with obesity, with 2.5 billion adults (43%) aged 18 and older classified as overweight and 890 million (16%) affected by obesity.[2,3] In Thailand, the incidence of obesity has increased significantly since 1991, when 12.1% of males and 21.9% of females were classified as obese.[4] Data from the Thai National Health Examination Survey indicated that the prevalence of obesity among individuals aged 15 years or older increased from 28.6% in 1991 to 42.2% in 2020. Obesity also increases the risk of several health conditions, including heart disease, type 2 diabetes, obstructive sleep apnea, specific cancers, and osteoarthritis.[5-7]

There is evidence that the presence of left atrial (LA) enlargement is associated with cardiovascular risks, such as heart failure, atrial fibrillation, cerebrovascular disease, and increased mortality.[8-10] It is interesting to note that obesity may be an independent factor associated with LA enlargement. Previous studies have demonstrated a correlation between obesity and LA enlargement. For example, Gerdts et al. found that BMI has a continuous correlation with LA size and is an independent factor, unrelated to left ventricular (LV) hypertrophy or baseline blood pressure.[11] Huang et al. recently showed that

higher BMI and male sex were independently associated with increased LA diameter, LA area, and LA volume. Furthermore, older age and larger LA diameter were independently associated with a higher incidence of atrial fibrillation.[12]

However, previous studies have limitations in identifying LA enlargement by measuring LA diameter in one or two dimensions using transthoracic echocardiography, which has lower accuracy and reliability due to investigator dependence. Cardiac magnetic resonance imaging (CMR) is superior to echocardiography for measuring LA volume, as it provides more reliable, reproducible measurements and serves as a better predictor of cardiovascular risk and mortality.[13,14] Our aim is to investigate the relationship between obesity and LA volume using CMR.

MATERIALS AND METHODS

Study population

We prospectively enrolled consecutive eligible patients aged 18 years or older who underwent clinical CMR at the Division of Cardiology, Department of Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand. Exclusion criteria included patients with contraindications to CMR or those with underlying heart diseases that cause significant LA enlargement, such as moderate or severe mitral regurgitation, mitral stenosis, LV systolic dysfunction (left ventricular ejection fraction [LVEF] less than 40%), restrictive cardiomyopathy, and hypertrophic cardiomyopathy. Blood pressure, BMI, and waist circumference were measured for each patient. Clinical variables and medical history were collected from medical records. The study protocol was approved by the Siriraj Institutional Review Board. Written informed consent was obtained from all patients.

CMR

All patients were examined using a 1.5-T Achieva XR MRI scanner (Phillips, Netherlands). After acquiring a scout image to locate the cardiac axis, steady-state free precession (SSFP) cine imaging was performed in standard axis views, including short axis, and 2-, 3-, and 4-chamber views. CMR parameters included a repetition time/echo time/number of excitations of 3.7/1.8/1, an 8 mm slice thickness with no gap, a field of view measuring 270 x 320 mm for the short axis and 350 x 320 mm for the long axis, a reconstruction pixel size of 1.25 x 1.25 x 8 mm^2, a 60° flip angle, and 25 cardiac phases per minute. Other CMR parameters, including LV end-systolic and end-diastolic dimensions, LVEF, and LV mass index, were measured in accordance with current recommendations.[15]

The LA volume was measured during the end-atrial diastolic phase using the biplane area-length technique, including the LA appendage while excluding the pulmonary veins.[13,16] To identify potential measurement errors, randomization was performed for repeat measurements after 4 weeks by another investigator.

Statistical analysis

Statistical analyses were conducted using SPSS Statistics for Windows, version 20.0 (IBM Corp., Armonk, NY, USA). Continuous variables were reported as mean ± standard deviation, while categorical variables were presented as absolute numbers and percentages. Comparisons of normally distributed continuous variables between the two groups were performed using Student's unpaired t-test. Categorical data were compared using the chi-square test or Fisher's exact test, as appropriate. Correlations were applied to evaluate univariable relationships between clinical variables and LA volume. Stepwise multiple regression analysis was used to determine the independent variable of clinical descriptors. For multiple regression analyses, models were developed using individual predictor variables and by examining all possible two-way interactions. Group comparisons of the distribution of continuous variables were conducted using the chi-square test for homogeneity in contingency tables. All statistical tests were two-tailed, with $p < 0.05$ considered statistically significant.

RESULTS

Patients' characteristics

The study included 140 patients, comprising 58 (41%) men, with an average age of 66.5 ± 10.5 years. The average BMI was 25.7 ± 4.2 kg/m^2, and the mean waist circumference was 92.8 ± 10.8 cm. Patients were categorized into two groups: the obese group (BMI ≥ 25 kg/m^2) and the non-obese group (BMI < 25 kg/m^2). The non-obese group comprised 48.6% (n = 68) of patients, while the obese group comprised 51.4% (n = 72). In the obese group, significant differences were found in waist circumference (99.3 ± 9.4 cm vs. 86.0 ± 7.3 cm, $p < 0.001$), mean body weight (72.8 ± 11.9 kg vs. 56.0 ± 8.1 kg, $p < 0.001$), the prevalence of hypertension (88.9% vs. 67.8%, $p < 0.001$), and hypercholesterolemia (63.2% vs. 84.7%, $p = 0.01$). The other parameters showed no significant differences between the two groups. A comparison of patients' characteristics is presented in Table 1.

CMR parameters

All patients in both groups underwent CMR imaging. Patients in the obese group had significantly greater values

TABLE 1. Baseline demographic, clinical, and anthropometric characteristics of study patients compared between those with a BMI < 25 and those with a BMI ≥ 25 kg/m².

Characteristics	BMI < 25 kg/m² (n = 68)	BMI ≥ 25 kg/m² (n = 72)	P-value
Age (years), mean ± SD	67.9 ± 10.6	65.2 ± 10.3	0.13
Male sex, n (%)	26 (38.0)	32 (44.0)	0.46
Mean waist circumference (cm), mean ± SD	86.0 ± 7.3	99.3 ± 9.4	***<0.001***
Mean body weight (kg), mean ± SD	56.0 ± 8.1	72.8 ± 11.9	***<0.001***
Mean height (cm), mean ± SD	158.7 ± 7.4	158.6 ± 9.3	0.92
Mean SBP (mmHg), mean ± SD	131.3 ± 19.5	135.3 ± 16.8	0.22
Mean DBP (mmHg), mean ± SD	67.8 ± 11.3	68.5 ± 12.2	0.72
Prior coronary heart disease, n (%)	14 (20.6)	23 (31.9)	0.09
Hypertension, n (%)	42 (67.8)	64 (88.9)	***<0.001***
Diabetes mellitus, n (%)	33 (45.8)	22 (32.4)	0.10
Hypercholesterolemia, n (%)	61 (84.7)	43 (63.2)	***0.01***

A p-value < 0.05 indicates statistical significance.
Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; kg/m², kilograms divided by meters squared; mmHg, millimeters of mercury; SBP, systolic blood pressure; SD, standard deviation

than those in the non-obese group for left ventricular end-diastolic volume (124.5 ± 29.1 ml vs. 108.3 ± 30.2 ml, p = 0.002), LVEF (73.2 ± 8.4% vs. 68.9 ± 10.1%, p = 0.007), and LA volume (80.8 ± 21.1 ml vs. 68.4 ± 22.4 ml, p = 0.001) (Table 2). None of the patients had a significant valvular heart disease.

Correlation factors affecting LA volume

Univariable analysis revealed that LA volume, as a continuous variable, was positively correlated with BMI (r = 0.3, p < 0.001), LV mass index (r = 0.3, p = 0.001), waist circumference (r = 0.3, p < 0.001), male sex (p = 0.01), and hypertension (p = 0.01). The correlation between BMI and LA volume is illustrated in Fig 1.

TABLE 2. CMR parameters of study patients compared between those with a BMI < 25 and those with a BMI ≥ 25 kg/m².

CMR parameter (mean ± SD)	BMI < 25 kg/m² (n = 68)	BMI ≥ 25 kg/m² (n = 72)	P-value
LV end-diastolic volume (ml)	108.3 ± 29.1	124.5 ± 30.2	***0.002***
LV end-systolic volume (ml)	35.8 ± 18.7	34.8 ± 17.1	0.75
LVEF (%)	68.9 ± 10.1	73.2 ± 8.4	***0.007***
LV mass index (g/m²)	45.1 ± 15.1	47.0 ± 12.9	0.43
LA volume (ml)	68.4 ± 22.4	80.8 ± 21.1	***0.001***

A p-value<0.05 indicates statistical significance.
Abbreviations: BMI, body mass index; CMR, cardiac magnetic resonance; g/m², grams divided by meters squared; kg/m², kilograms divided by meters squared; LA, left atrial; LV, left ventricular; LVEF, left ventricular ejection fraction; ml, milliliters; SD, standard deviation.

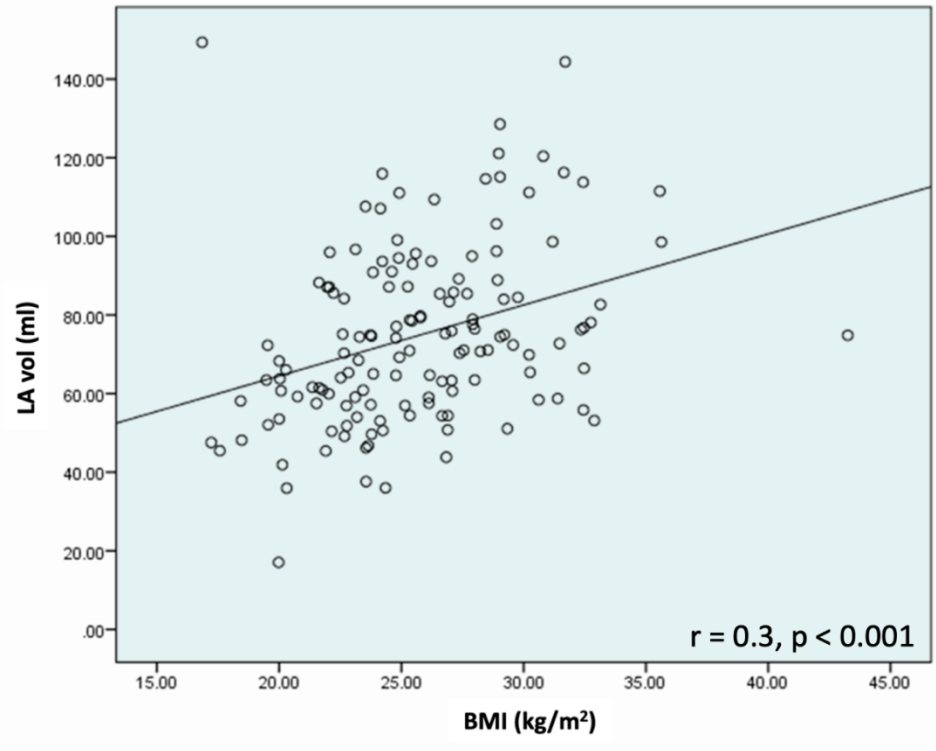


Fig 1. Correlation between body mass index (BMI) and left atrial volume (LA vol)

Stepwise multivariable analysis, using an indicator variable for LA enlargement as the dependent variable, revealed that BMI ($p < 0.001$), LV mass index ($p = 0.02$), and male sex ($p = 0.03$) were independently associated with LA volume.

DISCUSSION

This study revealed that obesity, as measured by BMI, was an independent predictor of LA enlargement. Other independent factors identified included the LV mass index and male sex.

The importance of the LA in cardiac pathophysiology is often underappreciated. However, LA size and/or function has been implicated in various cardiovascular and cerebrovascular disorders, either as a causal factor or as an indirect marker of disease.

Previous studies have demonstrated the significance of LA enlargement in various populations. A study by Bombelli et al. involving 1,785 participants from the general population showed that an LA diameter >2.3 cm/m² was associated with a significantly higher risk of fatal and nonfatal cardiovascular events.[17] LA enlargement and remodeling are also important risk markers for heart failure with preserved ejection fraction.[18] Furthermore, research has shown that a larger LA size is associated with a higher risk of ischemic stroke and all-cause mortality.[19,20] Notably, some studies suggest that this association between LA size and stroke exists independently of atrial fibrillation.[21]

Several studies have indicated that obesity is linked to increased LA size regardless of factors such as hypertension, LV geometry, and age. A study by Garza et al., which included nearly sixty patients who underwent bariatric surgery (gastric bypass) and transthoracic echocardiography before and after the procedure, found a positive correlation between changes in body weight and LA volume ($r = 0.22$, $p = 0.006$).[22] The study concluded that effective weight reduction achieved through bariatric surgery can prevent an increase in LA volume.[22]

Varying methods and values affected the proportion of patients with LA enlargement. Currently, the recommended and most commonly used parameter for normalizing LA volume measurements is body surface area (BSA). However, because BSA is determined using both weight and height, indexing LA volume using BSA may underestimate LA remodeling, particularly in obese individuals. In this study, we used LA volume measured by CMR to determine the correlation between obesity and LA enlargement, providing more accurate, non-operator dependent, and reproducible results.[13,14] In our study, we found a significant association between obesity and LA volume. Patients in the obese group exhibited significantly greater LV end-diastolic volume, LA volume, and LVEF. These findings suggested that obesity is associated with diastolic dysfunction. Next, we identified that BMI, LV mass index, and male sex were significantly correlated with the LA volume. Consistent with previous studies, our findings support the hypothesis that obese individuals

are more prone to diastolic dysfunction leading to LA enlargement.[23]

Diastolic dysfunction is more common among obese patients, especially those with obesity-related conditions such as hypertension. Studies have shown a direct association between obesity and structural cardiac abnormalities. For instance, elevated blood volume in obese individuals promotes eccentric hypertrophy, which reduces LV compliance. When the LV is non-compliant, LA pressure increases to ensure sufficient LV filling, resulting in increased atrial wall tension and subsequent LA enlargement.[23] Other explanations include obesity-induced obstructive sleep apnea, lipotoxicity-induced oxidative stress, and effects of adipokines.[24]

Our study has some limitations. First, we included patients referred for CMR imaging, primarily due to suspected myocardial ischemia, rather than healthy individuals. These factors may have increased the risk of selection bias. Second, patients in the obese group had a significantly higher prevalence of hypertension and hypercholesterolemia than those in the non-obese group did. These conditions may act as confounders; however, we performed adjusted analyses using multivariable methods, which showed that BMI was significantly correlated with LA volume, independent of hypertension or hypercholesterolemia. Finally, we did not use indexed LA volume in this study because it may underestimate LA remodeling, particularly in obese individuals.

CONCLUSION

Obesity, as represented by BMI, was an independent factor associated with LA enlargement. Other independent factors correlated with LA volume included the LV mass index and male sex. Future larger studies, along with prognostic studies, would be beneficial to further support our findings and to enhance the value of LA volume measurement using CMR in this population.

Data Availability Statement

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

None

DECLARATION

Grants and Funding Information

None

Conflict of Interests

No potential conflict of interest relevant to this study was reported.

Author Contributions

S.J. conceived and designed the study, conducted research operations, collected and analyzed data, interpreted results, drafted and revised the manuscript, and approved the final version. Y.K., M.M., and T.B. also contributed to conception and design, data analysis and interpretation, discussion of results, manuscript revision, and final approval

Ethical Statement

Study protocol was reviewed and approved by the Siriraj Institutional Review Board. Written informed consent was obtained from all patients.

REFERENCES

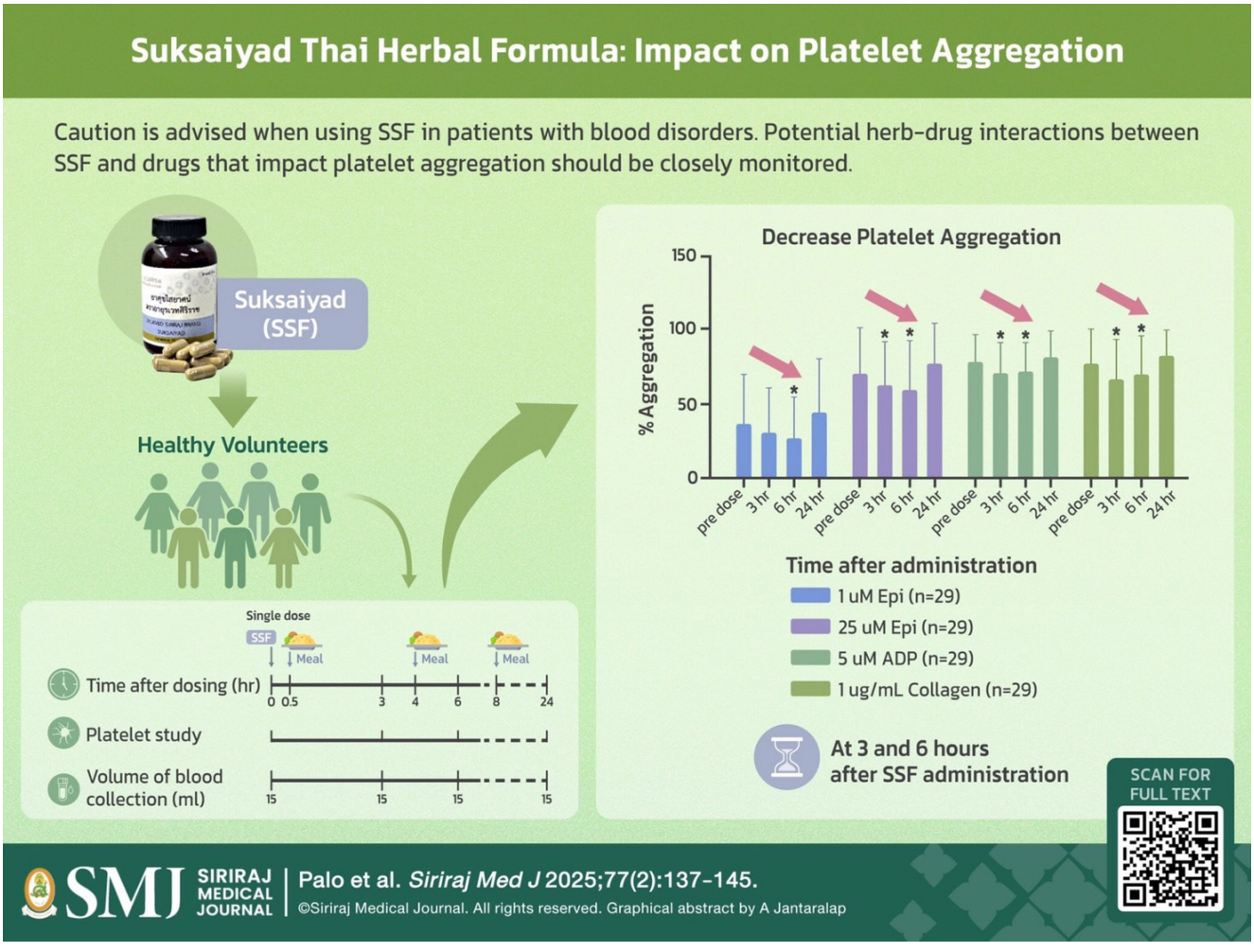
1. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157-63.
2. Murray CJL, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 2020;396(10258):1223-49.
3. Okunogbe A, Nugent R, Spencer G, Powis J, Ralston J, Wilding J. Economic impacts of overweight and obesity: current and future estimates for 161 countries. BMJ Global Health. 2022;7(9):e009773.
4. Aekplakorn W, Mo-Suwan L. Prevalence of obesity in Thailand. Obes Rev. 2009;10(6):589-92.
5. Powell-Wiley TM, Poirier P, Burke LE, Després J-P, Gordon-Larsen P, Lavie CJ, et al. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation. 2021;143(21):e984-e1010.
6. Pati S, Irfan W, Jameel A, Ahmed S, Shahid RK. Obesity and Cancer: A Current Overview of Epidemiology, Pathogenesis, Outcomes, and Management. Cancers (Basel). 2023;15(2):485.
7. Jehan S, Zizi F, Pandi-Perumal SR, Wall S, Auguste E, Myers AK, et al. Obstructive Sleep Apnea and Obesity: Implications for Public Health. Sleep Med Disord. 2017;1(4):00019.
8. Takemoto Y, Barnes ME, Seward JB, Lester SJ, Appleton CA, Gersh BJ, et al. Usefulness of left atrial volume in predicting first congestive heart failure in patients > or = 65 years of age with well-preserved left ventricular systolic function. Am J Cardiol. 2005;96(6):832-6.
9. Tsang TS, Abhayaratna WP, Barnes ME, Miyasaka Y, Gersh BJ, Bailey KR, et al. Prediction of cardiovascular outcomes with left atrial size: is volume superior to area or diameter? J Am Coll Cardiol. 2006;47(5):1018-23.
10. Laukkanen JA, Kurl S, Eränen J, Huttunen M, Salonen JT. Left atrium size and the risk of cardiovascular death in middle-aged men. Arch Intern Med. 2005;165(15):1788-93.
11. Gerdts E, Oikarinen L, Palmieri V, Otterstad JE, Wachtell K, Boman K, et al. Correlates of left atrial size in hypertensive

patients with left ventricular hypertrophy: the Losartan Intervention For Endpoint Reduction in Hypertension (LIFE) Study. Hypertension. 2002;39(3):739-43.
12. Huang G, Parikh PB, Malhotra A, Gruberg L, Kort S. Relation of Body Mass Index and Gender to Left Atrial Size and Atrial Fibrillation. Am J Cardiol. 2017;120(2):218-22.
13. Maceira AM, Cosín-Sales J, Roughton M, Prasad SK, Pennell DJ. Reference left atrial dimensions and volumes by steady state free precession cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance. 2010;12(1):65.
14. Khan MA, Yang EY, Zhan Y, Judd RM, Chan W, Nabi F, et al. Association of left atrial volume index and all-cause mortality in patients referred for routine cardiovascular magnetic resonance: a multicenter study. J Cardiovasc Magn Reson. 2019;21(1):4.
15. Schulz-Menger J, Bluemke DA, Bremerich J, Flamm SD, Fogel MA, Friedrich MG, et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update. Journal of Cardiovascular Magnetic Resonance. 2020;22(1):19.
16. Kawel-Boehm N, Hetzel SJ, Ambale-Venkatesh B, Captur G, Francois CJ, Jerosch-Herold M, et al. Reference ranges ("normal values") for cardiovascular magnetic resonance (CMR) in adults and children: 2020 update. Journal of Cardiovascular Magnetic Resonance. 2020;22(1):87.
17. Bombelli M, Facchetti R, Cuspidi C, Villa P, Dozio D, Brambilla G, et al. Prognostic Significance of Left Atrial Enlargement in a General Population. Hypertension. 2014;64(6):1205-11.
18. Rossi A, Gheorghiade M, Triposkiadis F, Solomon SD, Pieske B, Butler J. Left Atrium in Heart Failure With Preserved Ejection Fraction. Circulation: Heart Failure. 2014;7(6):1042-9.
19. Gardner JD, Skelton WPt, Khouzam RN. Is It Time to Incorporate the Left Atrial Size to the Current Stroke Risk Scoring Systems for Atrial Fibrillation?(☆). Curr Probl Cardiol. 2016;41(9-10):251-9.
20. Biteker M, Kayataş K, Başaran Ö, Dogan V, Özlek E, Özlek B. The Role of Left Atrial Volume Index in Patients with a First-ever Acute Ischemic Stroke. J Stroke Cerebrovasc Dis. 2017;26(2):321-6.
21. Hamatani Y, Ogawa H, Takabayashi K, Yamashita Y, Takagi D, Esato M, et al. Left atrial enlargement is an independent predictor of stroke and systemic embolism in patients with non-valvular atrial fibrillation. Sci Rep. 2016;6:31042.
22. Garza CA, Pellikka PA, Somers VK, Sarr MG, Seward JB, Collazo-Clavell ML, et al. Major weight loss prevents long-term left atrial enlargement in patients with morbid and extreme obesity. Eur J Echocardiogr. 2008;9(5):587-93.
23. Russo C, Jin Z, Homma S, Rundek T, Elkind MS, Sacco RL, et al. Effect of obesity and overweight on left ventricular diastolic function: a community-based study in an elderly cohort. J Am Coll Cardiol. 2011;57(12):1368-74.
24. Lopez-Jimenez F, Almahmeed W, Bays H, Cuevas A, Di Angelantonio E, le Roux CW, et al. Obesity and cardiovascular disease: mechanistic insights and management strategies. A joint position paper by the World Heart Federation and World Obesity Federation. European Journal of Preventive Cardiology. 2022;29(17):2218-37.

The Effect of the Thai Herbal Suksaiyad Formula on Platelet Aggregation in Healthy Volunteers: A Quasi-experimental, Single-dose Study

Titchaphorn Palo, M.Sc.[1], Nalinthika Nathananwanit, Ph.D.[2], Pravit Akarasereenont, Ph.D.[1,2], Suksalin Booranasubkajorn, Ph.D.[2,*]

[1]Department of Pharmacology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]The Center of Applied Thai Traditional Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



*Corresponding author: Suksalin Booranasubkajorn
E-mail: Suksalin.boo@mahidol.ac.th
Received 31 October 2024 Revised 25 November 2024 Accepted 25 November 2024
ORCID ID:http://orcid.org/0000-0002-3192-878X
https://doi.org/10.33192/smj.v77i2.271976

ABSTRACT

Objective: To investigate the effects of the Thai Herbal Suksaiyad formula (SSF) on platelet aggregation in healthy volunteers.

Material and Methods: In a quasi-experimental study, thirty healthy volunteers received a single dose of 2,000 mg SSF. Blood samples were taken at 0, 3, 6, and 24 hours after SSF administration for platelet aggregation analysis using aggregometry. Platelets were induced with epinephrine, adenosine diphosphate, and collagen.

Results: SSF significantly decreased platelet aggregation at 3 and 6 hours post-administration. Sub-analysis revealed no significant differences between males and females. SSF significantly decreased platelet aggregation in normal and hyperaggregation groups but had no effect on the disaggregation group. The effect of SSF was short-lived, reverting to pre-dose values after 24 hours for all agonists. Adverse events included flatulence (1 participant) and diarrhea (1 participant).

Conclusion: Caution is advised when using SSF in patients with blood disorders. Potential herb-drug interactions between SSF and drugs that impact platelet aggregation, such as aspirin, should be closely monitored. These interactions could result in fluctuations that may enhance or diminish the effectiveness of concurrent antiplatelet therapies, potentially increasing the risk of bleeding or decreasing therapeutic outcomes.

Keywords: Thai Herbal Suksaiyad formula; herbal medicine; platelet aggregation; Thai traditional medicine (Siriraj Med J 2025; 77: 137-145)

INTRODUCTION

The re-legalization of marijuana in 2019 has reignited interest in its potential therapeutic applications in Thai medicine.[1,2] To effectively integrate cannabis-based treatments into clinical practice, further research is essential to provide empirical evidence supporting its use in medicinal formulations.

The Suksaiyad formula (SSF) is one of 16 Thai medicinal formulas that contain marijuana (*Cannabis sativa* L.). It has been approved by the Department of Thai Traditional and Alternative Medicine, Ministry of Public Health, in powdered and capsule form for treating diseases or conducting research studies.[3] This formula was first recorded in *Thart Phra Narai Scripture*[4], a palm-leaf (bai-lahn) manuscript composed of various drug formulas. The published version, known as *Tamra Phra Osot Phra Narai* was revised by Prince Damrong Rajanupab in 2466 B.E. (1923) during the Rattanakosin period.[5]

SSF consists of 12 components, including *Cannabis sativa* L. (Leaf), *Piper retrofractum* Vahl. (Flower), *Zingiber officinale* Roscoe (Rhizome), *Piper nigrum* L. (Seed), *Mesua ferrea* L. (Flower), *Myristica fragrans* Houtt. (Nutmeg), *Aucklandia lappa* DC. (Root), *Nigella sativa* L. (Seed), *Cinamomum bejolghota* (Buch. -Ham.) Sweet (Bark), *Micromelum minutum* Wight & Arn. (Stem), *Azadirachta indica* A.Juss. (Leave), and camphor. The SSF has traditionally been used for treating various illnesses, including insomnia and anorexia.[3,6] According to the National List of Essential Medicines, the recommended dosage for treating insomnia is 0.5-2 grams, taken once daily before bedtime. For improving appetite, 0.5-2 grams can be taken once or twice before meals.[7]

Cannabis sativa L. contains key active compounds such as Delta-9-Tetrahydrocannabinol (THC) and Cannabidiol (CBD). Previous studies have shown that these compounds can inhibit human and rabbit platelet aggregation, and chronic THC administration via edibles has been found to desensitize platelet activity and function.[8-10] Previous studies also demonstrate that other components, including *Zingiber officinale* Roscoe[11-13], *Piper nigrum* L.[14], *Myristica fragrans* Houtt.[15], and *Nigella sativa* L.[16,17], exhibit antiplatelet effects through different mechanisms. This suggests that the SSF may influence platelet function. However, balancing its anti-thrombotic potential with the risk of bleeding remains a concern, as evidenced by previous studies on the effects of the traditional herbal formula "Wattana" on platelet aggregation.[18] Accordingly, this study aims to investigate the effect of the Thai Herbal Suksaiyad Formula on platelet aggregation and report any adverse events in healthy volunteers. The findings will provide empirical evidence to support the clinical use of SSF and enhance its safety.

MATERIALS AND METHODS

Study drugs

SSF capsules (500 mg) were manufactured according to GMP PIC/S (Good Manufacturing Practices) by the Manufacturing Unit of Herbal Medicines and Products, Center of Applied Thai Traditional Medicine (CATTM),

Faculty of Medicine, Siriraj Hospital, Mahidol University. All SSF capsules used in the study came from the same batch and underwent quality control testing, including FTIR, UPLC methods, physical properties analysis, and microbial contamination checks. Capsules were stored at room temperature in dry conditions until use.

Patients

Thirty healthy volunteers were recruited for this pilot study. Inclusion criteria were: 1) Thai male or female, 18-45 years old; 2) body mass index (BMI) between 18-29 kg/m^2; 3) good health confirmed by normal or clinically insignificant vital signs, physical examination, and blood chemistry, per the research physician's assessment. Exclusion criteria included: 1) use of drugs affecting platelet function such as aspirin, clopidogrel, and ticlopidine; 2) history of psychiatric disorders such as psychosis or anxiety disorder; 3) use of central nervous system depressants (eg; anti-seizure medications); 4) fever, peptic ulcers, or and gastroesophageal reflux disease; 5) history of major abnormal bleeding within the past 6 months; 6) blood donation or transfusion in the past 2 months; 7) known allergic reactions to herbal medicine; 8) history of impaired kidney and liver function; 9) pregnancy and breastfeeding. Volunteers experiencing adverse events due to SSF or believed to have had such events were withdrawn with physician agreement. All participants were advised to avoid caffeine, alcohol, vitamins, dietary supplements, and foods containing any of the 12 SSF components for at least two weeks prior to and throughout the study.

Study design

A pre-post study was conducted at the Faculty of Medicine, Siriraj Hospital, Bangkok, Thailand. The protocol received approval from the Siriraj Institutional Review Board (COA: Si 269/2020) and registered with the Thai Clinical Trials Registry (TCTR20210405001). All participants provided written informed consent and retained the right to withdraw from the study at any time. Participants were required to fast for 10 hours before taking the SSF. On the day of the experiment, participants underwent a physical examination by a doctor and had blood drawn at pre-dose (baseline) and at 3, 6, and 24 hours after administration single-dose of SSF (2,000 mg), which is the recommended maximum daily dosage. However, since there is no prior research indicating the duration of the drug's effects, the dosage was set as a single dose, with follow-up evaluations conducted at 3, 6, and 24 hours after administration to cover potential effects of SSF that may occur. Blood samples were kept in sodium citrate vacutainers (Greiner Bio-one GmbH, Austria). All participants received standardized food provided by the researchers and were not allowed to eat anything else until the final blood sample was taken. The study physicians evaluated and recorded vital signs, and adverse events were assessed by questioning the participants. (Fig 1).

Platelet aggregation assay

Platelet aggregation was assessed using light transmission aggregometry (LTA) following Born's technique.[19] Platelet aggregation was measured with an aggregometer (AggRAM, Helena, USA) within three hours of blood collection. The platelet agonists used were epinephrine (Epi), adenosine diphosphate (ADP), and collagen (Col), which is crucial because each targets different receptors and mechanisms, providing a comprehensive assessment of platelet function.[20] Platelet-rich plasma (PRP) was prepared by centrifuging citrated whole blood at 250 g for 10 minutes at room temperature. A portion of PRP was further centrifuged at 4500 g for 2 minutes at room temperature to create platelet-free plasma (PFP) as a blank control. PRP was stirred at 600 rpm and incubated at 37°C for about 3 minutes before induction. The reaction

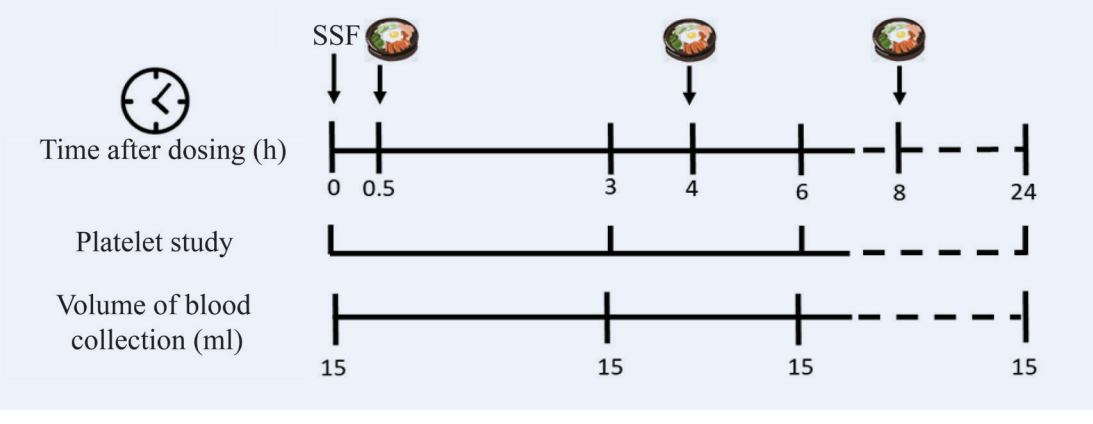


Fig 1. Intervention and outcome measurement screen.

lasted for 5 minutes. The maximal amplitude of platelet aggregation, expressed as a percentage, was determined by computing light transmission between aggregated PRP and PFP. Additionally, platelet count was assessed using a non-metallic hemocytometer (Helena USA).

Platelet status classification

There is no officially accepted standard for evaluating platelet function. A previous method categorized platelet function into three groups based on platelet responses to different epinephrine concentrations.[21,22] The primary phase of platelet aggregation, induced by 25μM epinephrine, is referred to as the "disaggregation pattern". Platelet aggregation triggered by 1μM epinephrine, representing the secondary phase, is referred to as the "hyperaggregation pattern". Consequently, the concentration-dependent response of platelets to both 1μM and 25μM epinephrine is considered the "normal aggregation pattern".[21,22]

Statistical analyses

Data are presented as mean ± standard deviations (SD). Statistical analyses were performed using GraphPad Prism version 10.0.2 (GraphPad Software Inc., San Diego, CA, USA). The normality of platelet aggregation data was assessed using the Kolmogorov–Smirnov and Shapiro-Wilk tests. Time-dependent changes in platelet aggregation data after SSF administration within groups after SSF administration were evaluated using nonparametric Friedman test, followed by Dunn's pairwise post hoc tests. Categorical variables were evaluated using the chi-square test. Statistical significance was defined as a p-value below 0.05.

RESULTS

Demographic characteristics

This study included 30 healthy volunteers, with one participant withdrawing due to difficulty collecting blood at 6 hours after SSF administration. Data from 29 healthy volunteers were analyzed. The average age of the male and female groups was 29.3±5.7 and 30.21±4.6 years, respectively. All participants had normal baseline characteristics (Table 1).

TABLE 1. Demographic data and reference range for laboratory values.

Topic	Mean ± SD		Reference range	
	Male (n=10)	Female (n=19)	Male	Female
Age (years)	29.3±5.7	30.21±4.6	≥18	
Body weight (kg)	64.83±7.4	55.96±7.5	-	
Height (cm)	170±10	161±4.5	-	
Body mass index (kg/m^2)	22.47±2.7	21.56±2.5	18-29	
hemoglobin (g/dL)	14.46±1.2	12.2±1	12.70-16.90	12.0-14.90
hematocrit (%)	44.38±2.5	38.8±2.2	40.30-51.90	37.0-45.70
WBC (x 10^3/uL)	6.03±1	6±1.2	4.50-11.30	4.40-10.30
Platelet count (x 10^3/uL)	264.2±28.3	294.5±54.8	160-356	179-435
MCV (fl)	84.65±7.6	84.53±5.9	80.4-95.9	
MCHC (g/dL)	32.57±1.2	31.37±1.1	30.2-34.2	
MCH (pg)	27.56±2.5	26.55±2.4	25.0-31.2	
FBS (mg/dL)	85.7±8.2	83±7	74-99	
BUN (mg/dL)	13.05±2.9	10.4±1.5	6-20	
Creatinine (mg/dL)	0.90±0.1	0.72±0.1	0.67-1.17	0.51-0.95
AST (U/L)	19.8±4	17.5±5.1	0-40	0-32
ALT (U/L)	18.4±5.1	14.4±12.3	0-41	0-33

Platelet status pattern before dosing

Before SSF treatment, 35% of the 29 subjects exhibited a hyperaggregation pattern, while 41% had a normal pattern, and 24% showed a disaggregation pattern, respectively (Fig 2a). The Chi-square test results showed no significant association between gender and platelet status pattern before SSF dosing (Fig 2b).

Effect of SSF on platelet aggregation

Among the 19 female participants, 24 hours after SSF administration, 26% experienced an increase in platelet aggregation, 68% showed no change, and 5% a decrease in platelet aggregation (Fig 3). Among the 10 male participants, 30% showed an increase in platelet aggregation, 60% showed no change, and 10% experienced a decrease in platelet aggregation (Fig 3). Figs 3 and 4 demonstrate the individual response variations observed in the study.

The results indicated a significant decrease in the percentage of maximum aggregation for all agonists. Platelet aggregation induced by 1μM epinephrine significantly decreased at 6 hours after SSF administration. Aggregation induced by other agonists significantly decreased at 3 and 6 hours post-administration. By 24 hours, there was no significant difference in aggregation compared to pre-dose levels for all agonists (Fig 5a).

Both male and female groups showed a significant decrease in the percentage of maximum platelet aggregation at 3 and 6 hours after SSF administration. By 24 hours, there were no significant differences compared to pre-dose levels for any agonists in either group (Fig 5b, 5c).

In the normal aggregation group, a significant decrease in maximum aggregation was observed after induction with 25μM Epi at 6 hours and with 5μM ADP at 3 and 6 hours post-SSF administration. The hyperaggregation group also showed a significant decrease in maximum aggregation with all agonists at 3 and 6 hours post-SSF administration. By 24 hours, no significant differences were noted compared to pre-dose levels in either the normal or hyperaggregation groups. However, the

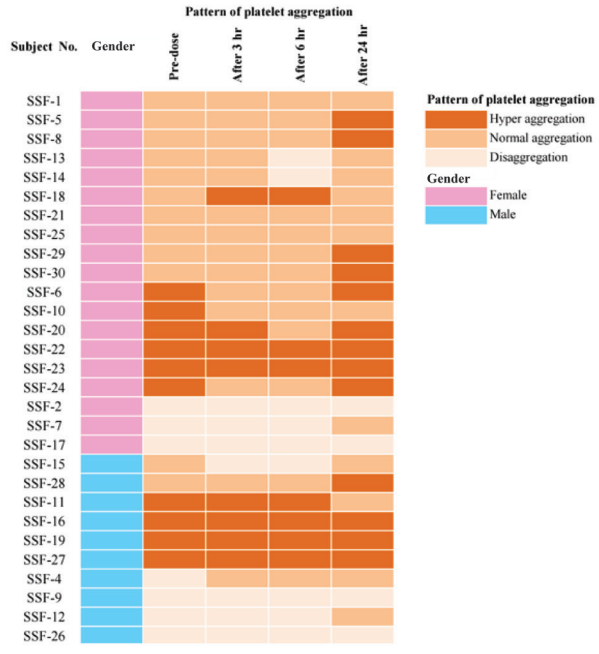


Fig 3. Effect of SSF on platelet aggregation by gender and platelet aggregation characteristics at pre-dose and at 3, 6, and 24 hours after administration of SSF.

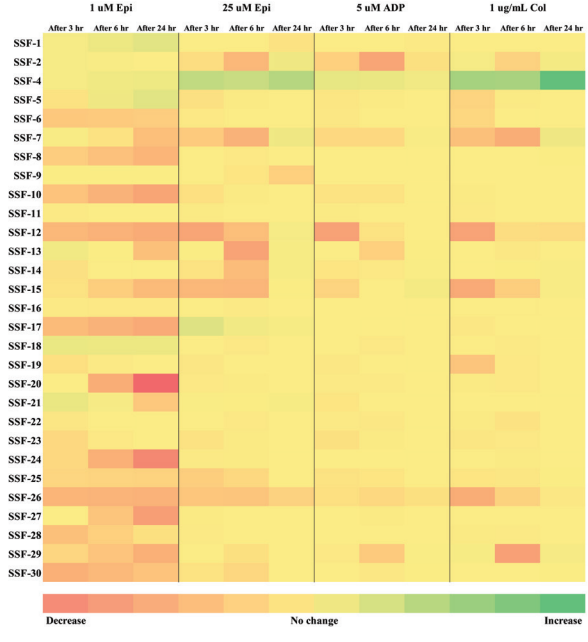


Fig 4. Heatmap illustrating the percentage of platelet aggregation in each participant for each agonist after 3, 6, and 24 hours of SSF administration.

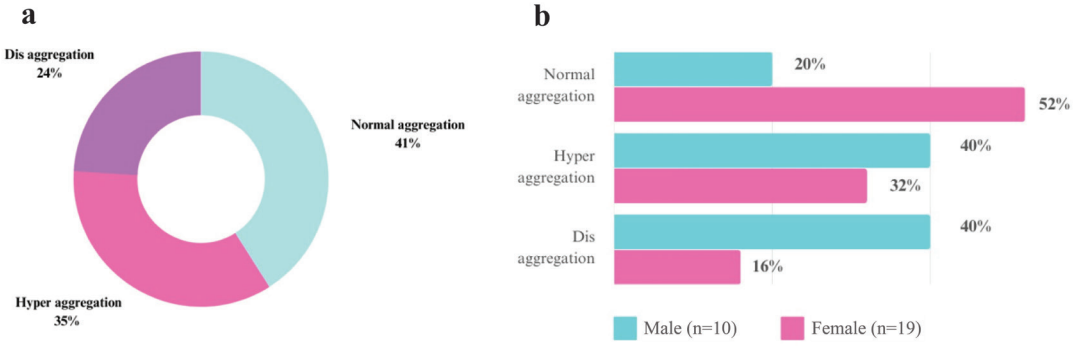


Fig 2. Percent platelet status pattern before SSF dosing: a) all 29 participants b) 10 males and 19 females.

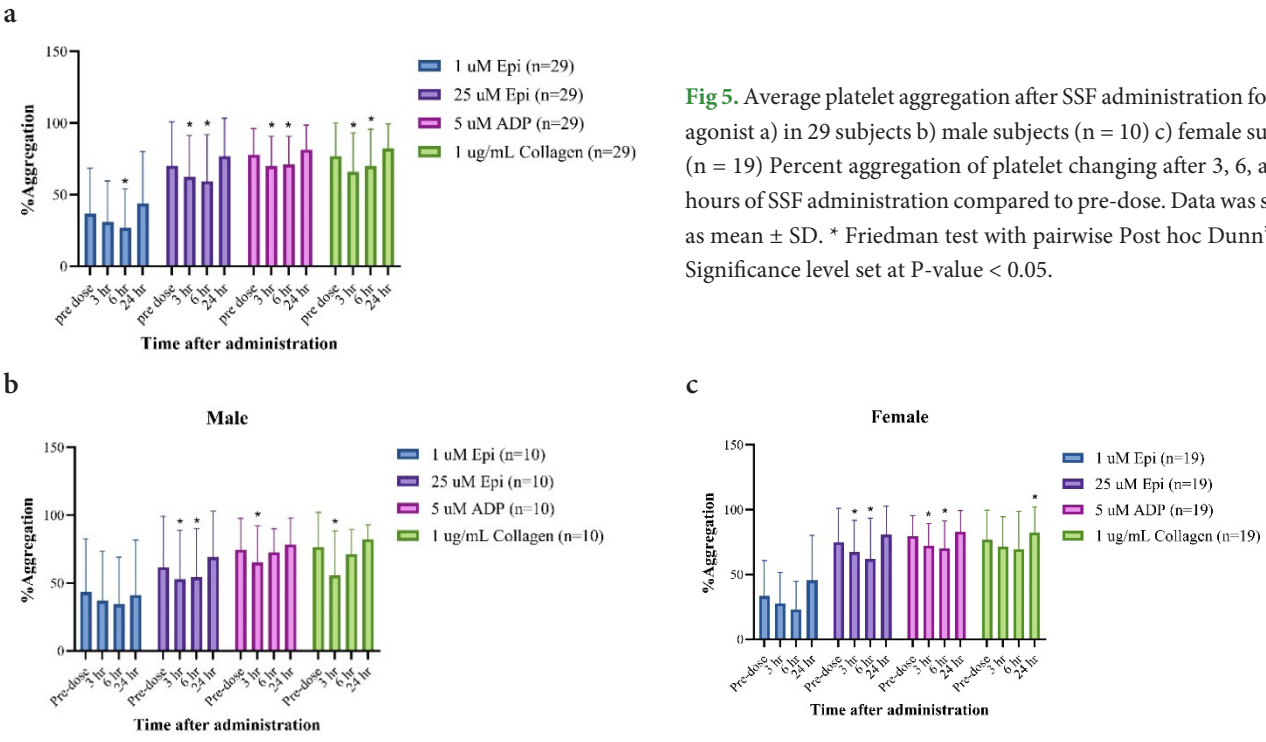


Fig 5. Average platelet aggregation after SSF administration for each agonist a) in 29 subjects b) male subjects (n = 10) c) female subjects (n = 19) Percent aggregation of platelet changing after 3, 6, and 24 hours of SSF administration compared to pre-dose. Data was shown as mean ± SD. * Friedman test with pairwise Post hoc Dunn's test; Significance level set at P-value < 0.05.

disaggregation group showed no significant changes in maximum aggregation after induction with any agonists following SSF administration (Fig 6). Adverse events reported by the 30 participants included flatulence (3.33%, n=1) and diarrhea (3.33%, n=1).

DISCUSSION

The novel findings of this study reveal changes in platelet aggregation after a single administration of 2000 mg of SSF. This herbal formula can reduce platelet aggregation induced by ADP, collagen, and epinephrine. However, this suppressive effect is short-lived, typically resolving within 24 hours post-SSF ingestion. The process of platelet aggregation is intricate, involving numerous receptors and signaling pathways. Collagen-induced platelet aggregation plays a key role in platelet adhesion and activation at sites of vascular injury, serving as one of the initial triggers for clotting.[23] ADP-induced platelet aggregation is vital for the early stages of clot formation, initiating platelet recruitment and activation.[24] Epinephrine-induced aggregation, less specific to vascular injury, responds primarily to circulating catecholamines like

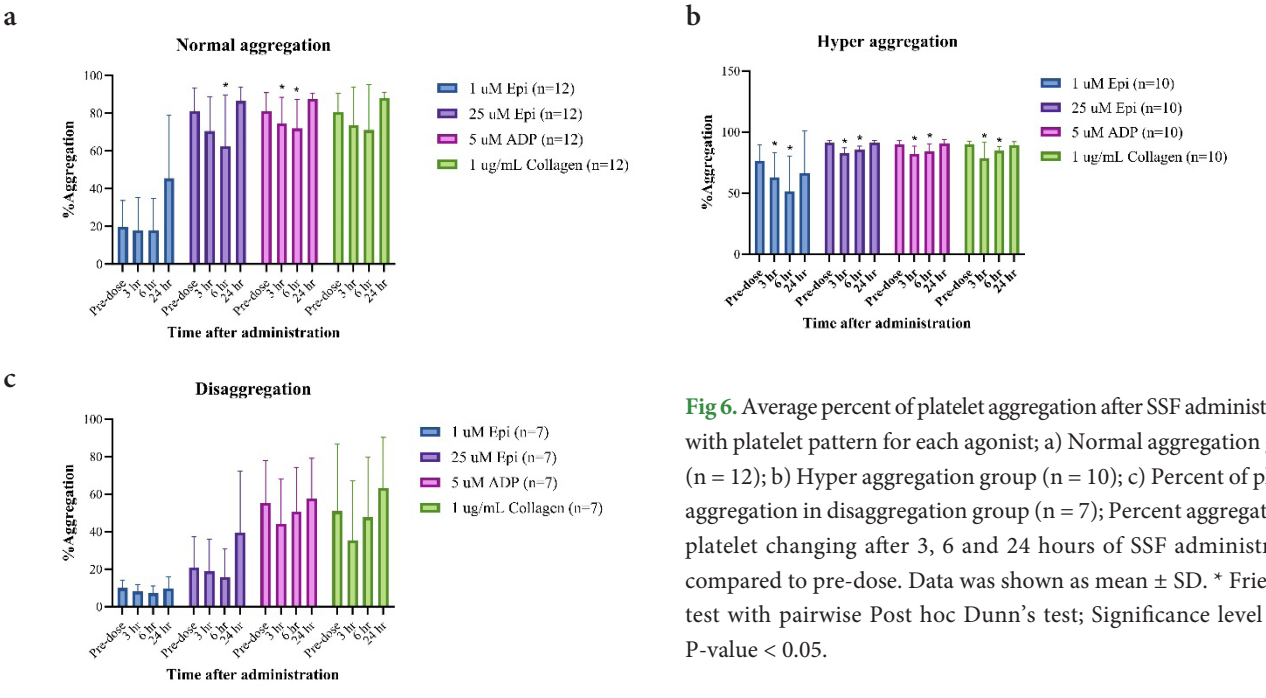


Fig 6. Average percent of platelet aggregation after SSF administration with platelet pattern for each agonist; a) Normal aggregation group (n = 12); b) Hyper aggregation group (n = 10); c) Percent of platelet aggregation in disaggregation group (n = 7); Percent aggregation of platelet changing after 3, 6 and 24 hours of SSF administration compared to pre-dose. Data was shown as mean ± SD. * Friedman test with pairwise Post hoc Dunn's test; Significance level set at P-value < 0.05.

adrenaline and is part of the body's overall hemostatic response but is not as site-specific as collagen or ADP-induced aggregation.[25-27]

The mechanisms of platelet aggregation induced by ADP, collagen, and epinephrine involve distinct pathways. The results indicate that SSF broadly inhibits platelet aggregation, likely due to its multiple components affecting these pathways. Previous studies have demonstrated that many SSF components inhibit platelet aggregation. For example, cannabinoids from *Cannabis sativa* L. inhibited epinephrine-induced human platelet aggregation in a dose-dependent manner, partially inhibiting primary aggregation and completely inhibiting secondary aggregation.[10] *Zingiber officinale* Roscoe extracts suppressed platelet aggregation induced by ADP, epinephrine, collagen, and arachidonate in a dose-dependent manner.[11-13] *Piper nigrum* L. extract displayed moderate antiplatelet effects induced by thrombin and collagen.[14] A neolignan extracted from *Myristica fragrans* Houtt. demonstrated inhibition of thrombin- and platelet-activating factor (PAF)-induced platelet aggregation in a concentration-dependent manner, without causing any damage to the platelets.[15] *Nigella sativa* L. extract also showed inhibitory effects on arachidonic acid (AA) and ADP-induced platelet aggregation.[16,17]

The individual variation in response to SSF's inhibitory effects on platelet aggregation, as shown in the heatmap, may be influenced by several factors. While the study protocol aimed to control for lifestyle and environmental factors such as age, diet, exercise, smoking habits, and overall health status, which can influence response to drug administration, platelet function and overall response to treatments[28], other physiological parameters, such as differences in baseline platelet activity, platelet count can affect the degree of platelet aggregation inhibition.[29,30] Additionally, genetic polymorphisms among individuals can influence the expression and function of platelet receptors and enzymes involved in aggregation, leading to diverse treatment responses.[31] Metabolic variability, such as variations in drug metabolism and clearance rates, can also impact the bioavailability and effective concentration of the drug. Factors like liver function, enzyme activity, and other metabolic processes can influence how the drug is processed and eliminated, affecting the degree and duration of platelet aggregation inhibition.[32]

The potential interaction of SSF with antiplatelet drugs could increase bleeding risks when co-administered. For patients already on such therapies, the use of SSF may require close monitoring or dosage adjustments. Recognizing and accounting for these individual variations is essential in assessing and managing SSF's effects on platelet aggregation. A personalized medicine approach, considering individual variability, can lead to more effective, tailored treatments that optimize therapeutic outcomes while minimizing potential adverse effects.

This study's small sample size may limit the generalizability of the results to broader populations and reduce the ability to detect individual response variations. Furthermore, the use of a single dose and the brief observation period may not adequately capture the prolonged or cumulative effects of repeated SSF administration. Future research should aim to increase the number of participants, utilize a randomized controlled trial design, and explore the long-term effects of repeated dosing to enhance the robustness and applicability of the findings.

CONCLUSIONS

A single administration of 2,000 mg of SSF can inhibit platelet aggregation induced by ADP, collagen, and epinephrine, but this effect is short-lived, with recovery occurring within 24 hours. SSF use may pose certain contraindications, particularly for individuals with health conditions such as bleeding disorders or a history of hemorrhagic stroke. Its use should be approached with caution, especially among individuals taking antiplatelet medications, as SSF may amplify their effects and increase the risk of bleeding complications. Patients on such therapies or with clotting disorders are advised to consult a healthcare provider before using SSF.

Data Availability Statement

The authors confirm that the data supporting the findings of this study are available within the article.

ACKNOWLEDGEMENTS

The authors would like to thank Natchaya Ziangchin and the research team for their collaboration during data collection. Furthermore, the authors would like to thank the Manufacturing Unit of Herbal Medicines and Products, the Center of Applied Thai Traditional Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, for authenticating the plant material and providing the study facilities.

DECLARATION

Grants and Funding Information

The study was financially supported by the Siriraj Routine to Research (R2R) (No. 21TT00023/032/21) Management Fund.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Author Contributions

T.P., P.A., and S.B. conceived and designed the experiments. T.P., N.N., and S.B. conducted the experiments. T.P. and S.B. contributed to data interpretation. S.B. led the manuscript writing. All authors provided critical feedback and contributed to refining the research, analysis, and manuscript.

REFERENCES

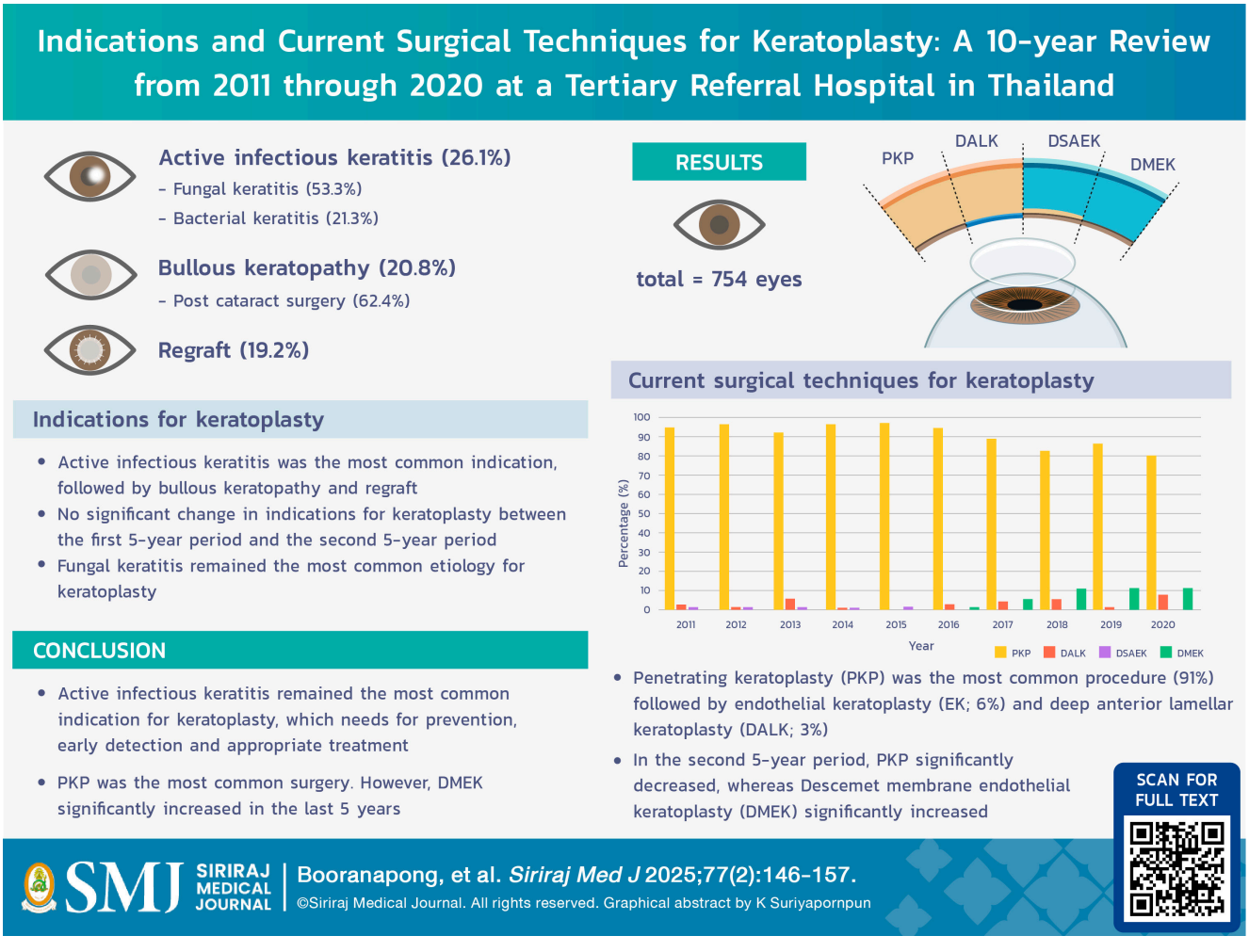
1. Narcotic Drugs Act B.E. 2522 (Includes additional amendments) [Internet]. [cited 22 Oct 2024]. Available from: https://catalog.fda.moph.go.th/gl/dataset/law-narcotic/resource/88e1084c-4cb3-4ecf-b537-bade8507aa26
2. Narcotic Drugs Act (Issue 7) B.E. 2562 [Internet]. [cited 22 Oct 2024]. Available from: https://catalog.fda.moph.go.th/gl/dataset/law-narcotic/resource/91905ef7-033b-4558-b053-2f734191a2c9
3. Vaibhav K, Shrivastava P, Javed H, Khan A, Ahmed ME, Tabassum R, et al. Piperine suppresses cerebral ischemia-reperfusion-induced inflammation through the repression of COX-2, NOS-2, and NF-kappaB in middle cerebral artery occlusion rat model. Mol Cell Biochem. 2012;367(1-2):73-84.
4. The National Library. Thart Phra Narai Scripture. palm-leaf book. Thai alphabet. Thai language. stylus-inscribed lines. raw palm-leaf version, no. 1143. medicine section
5. Prince Damrong Rajanupab. Tamra Phra Osot Phra Narai, Third printing for distribution at the funeral of Mr. Pan Chaisuwan1923, (B.E.2466).
6. Chayan Pichiansuntorn MC, Wichian Jeerawong. Explanation of King Narai's drug formulas (Tamra Phra Osot Phra Narai) The edition published in commemoration of His Majesty the King's 72nd Birthday Anniversary, 5 December 1999. Bangkok: Amarin Printing and Publishing; 2001.
7. Thailand. National Drug System Development Committee. The National List of Essential Medicines 2023: Specific medicine list. Government Gazette no. 140, section 130. June 2, 2023.
8. Reitsma SE, Lakshmanan HHS, Johnson J, Pang J, Parra-Izquierdo I, Melrose AR, et al. Chronic edible dosing of Delta9-tetrahydrocannabinol (THC) in nonhuman primates reduces systemic platelet activity and function. Am J Physiol Cell Physiol. 2022;322(3):C370-C81.
9. De Angelis V, Koekman AC, Weeterings C, Roest M, de Groot PG, Herczenik E, et al. Endocannabinoids control platelet activation and limit aggregate formation under flow. PLoS One. 2014;9(9):e108282.
10. Formukong EA, Evans AT, Evans FJ. The inhibitory effects of cannabinoids, the active constituents of Cannabis sativa L. on human and rabbit platelet aggregation. J Pharm Pharmacol. 1989;41(10):705-9.
11. Bordia A, Verma S K, Srivastava K C. Effects of ginger (Zingiber officinale Rosc.) and fenugreek (Trigonella foenumgraecum L.) on blood lipids, blood sugar and platelet aggregation in patients with coronary artery disease. Prostaglandins Leukot Essent Fatty Acids. 1997;56(5):379-84.
12. Srivas KC. Effects of aqueous extracts of onion, garlic and ginger on platelet aggregation and metabolism of arachidonic acid in the blood vascular system: in vitro study. Prostaglandins Leukot Med. 1984;13(2):227-35.
13. Liao YR, Leu YL, Chan YY, Kuo PC, Wu TS. Anti-platelet aggregation and vasorelaxing effects of the constituents of the rhizomes of Zingiber officinale. Molecules. 2012;17(8):8928-37.
14. Yen H-F, Wang S-Y, Wu C-C, Lin W-Y, Wu T-Y, Chang F-R, et al. Cytotoxicity, anti-platelet aggregation assay and chemical components analysis of thirty-eight kinds of essential oils. J Food Drug Anal. 2012;20(2):478-83.
15. Kang JW, Min BS, Lee JH. Anti-platelet activity of erythro-(7S,8R)-7-acetoxy-3,4,3',5'-tetramethoxy-8-O-4'-neolignan from Myristica fragrans. Phytother Res. 2013;27(11):1694-9.
16. Enomoto S, Asano R, Iwahori Y, Narui T, Okada Y, Singab AN, et al. Hematological studies on black cumin oil from the seeds of Nigella sativa L. Biol Pharm Bull. 2001;24(3):307-10.
17. Waqar MA, Mahmood Y, Saleem A, Saeed SA. An Investigation of Platelet Anti-aggregation Activity in Inigenous Medicinal Herbs. Journal of the Chemical Society of Pakistan. 2009;31(2):324-28.
18. Limsuvan S, Palo T, Mamaethong D, Yuthanarat N, Booranasubkajorn S, Lumlerdkij N, et al. The Effect of The Thai Herbal Wattana Formula on Platelet Aggregation and The Relationship with Innate Dhatu Chao Ruean. Siriraj Med J. 2023;75(5):321-9.
19. Born GV. Aggregation of blood platelets by adenosine diphosphate and its reversal. Nature. 1962;194:927-9.
20. Cattaneo M. Light Transmission Aggregometry and ATP Release for the Diagnostic Assessment of Platelet Function. Semin Thromb Hemost. 2009;35:158-67.
21. Ketsa-Ard K, Poungvarin N, Juengchareon M, Jarerat S, Kittigul L. Clinical study on antithrombotic effects of ticlopidine in ischemic stroke. J Med Assoc Thai. 1991;74(6):331-9.
22. Akarasereenont P, Tripatara P, Chotewuttakorn S, Palo T, Thaworn A. The effects of estrone, estradiol and estriol on platelet aggregation induced by adrenaline and adenosine diphosphate. Platelets. 2006;17(7):441-7.
23. Nieswandt B, Watson SP. Platelet-collagen interaction: is GPVI the central receptor? Blood. 2003;102(2):449-61.
24. Aslam M, Sedding D, Koshty A, Santoso S, Schulz R, Hamm C, et al. Nucleoside triphosphates inhibit ADP, collagen, and epinephrine-induced platelet aggregation: role of P2Y(1) and P2Y(1)(2) receptors. Thromb Res. 2013;132(5):548-57.
25. Evangelou AM, Malamas MP, Vezyraki P, Karkabounas SC. Is epinephrine-induced platelet aggregation autoregulated by its metabolic degradation products in vivo? In Vivo. 1998;12(3):321-5.
26. Lanza F, Beretz A, Stierle A, Hanau D, Kubina M, Cazenave JP. Epinephrine potentiates human platelet activation but is not an aggregating agent. Am J Physiol. 1988;255(6 Pt 2):H1276-88.
27. Lindkvist M, Fernberg U, Ljungberg LU, Falker K, Fernstrom M, Hurtig-Wennlof A, et al. Individual variations in platelet reactivity towards ADP, epinephrine, collagen and nitric oxide, and the association to arterial function in young, healthy adults. Thromb Res. 2019;174:5-12.
28. Peace AJ, Egan K, Kavanagh GF, Tedesco AF, Foley DP, Dicker P, et al. Reducing intra-individual variation in platelet aggregation: implications for platelet function testing. J Thromb

Haemost. 2009;7(11):1941-3.
29. Jones CI. Platelet function and aging. Mamm Genome. 2016; 27(7-8):358-66.
30. Carazo A, Hrubsa M, Konecny L, Skorepa P, Paclikova M, Musil F, et al. Sex-Related Differences in Platelet Aggregation: A Literature Review Supplemented with Local Data from a Group of Generally Healthy Individuals. Semin Thromb Hemost. 2023;49(5):488-506.
31. Sokol J, Skerenova M, Ivankova J, Simurda T, Stasko J. Association of Genetic Variability in Selected Genes in Patients With Deep Vein Thrombosis and Platelet Hyperaggregability. Clin Appl Thromb Hemost. 2018;24(7):1027-32.
32. Bhalerao S, Deshpande T, Thatte U. Prakriti (Ayurvedic concept of constitution) and variations in platelet aggregation. BMC Complement Altern Med. 2012;12:248.

Indications and Current Surgical Techniques for Keratoplasty: A 10-year Review from 2011 through 2020 at a Tertiary Referral Hospital in Thailand

Wipawee Booranapong, M.D.*, Chayanit Junsangsri, M.D., Chareenun Chirapapaisan, M.D., Panotsom Ngowyutagon, M.D., Sathiya Kengpunpanich, M.D.

Department of Ophthalmology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



*Corresponding author: Wipawee Booranapong
E-mail: wipawee.boo@mahidol.edu
Received 26 April 2024 Revised 14 November 2024 Accepted 24 November 2024
ORCID ID:http://orcid.org/0000-0002-8812-0319
https://doi.org/10.33192/smj.v77i2.268932

ABSTRACT

Objective: To report the indications and current trends in surgical techniques for keratoplasty over 10 years at Siriraj Hospital.

Materials and Methods: We retrospectively reviewed the medical records of hospitalized patients who underwent keratoplasty from 2011 to 2020. The collected data comprised demographics, indications for keratoplasty, and surgical techniques used.

Results: A total of 754 eyes were included. Active infectious keratitis was the most common indication (26.1%), followed by bullous keratopathy (20.8%) and regraft (19.2%). There was no significant change in indications for keratoplasty between the first 5-year period and the second 5-year period. Fungal keratitis remained the most common etiology for keratoplasty. Penetrating keratoplasty (PKP) was the most common procedure overall (90.7%), followed by endothelial keratoplasty (EK; 5.7%) and deep anterior lamellar keratoplasty (DALK; 3.6%). However, surgical techniques used in the second 5-year period differed significantly from those in the first 5-year period ($P < 0.05$). The use of PKP decreased from 96.0% to 86.7%; however, EK increased from 1.5% to 8.9%, while DALK rose slightly from 2.5% to 4.4%.

Conclusion: Active infectious keratitis was the most common indication for keratoplasty, followed by bullous keratopathy and regraft. These indications may be reduced through collaborative efforts among government leaders, public health officers, and ophthalmologists. PKP remained the most common surgical technique due to the high incidence of infectious keratitis. However, the use of lamellar keratoplasty, including EK and DALK, significantly increased for other indications. Descemet membrane endothelial keratoplasty was performed mostly in eyes with Fuchs endothelial dystrophy.

Keywords: Indications for keratoplasty; corneal transplantation; penetrating keratoplasty; deep anterior lamellar keratoplasty; Descemet membrane endothelial keratoplasty (Siriraj Med J 2025; 77: 146-157)

INTRODUCTION

According to the World Health Organization, blindness is defined as a corrected visual acuity of 3/60 or less in the better eye using Snellen's chart.[1] In 2015, the global prevalence of blindness was estimated at 36 million people, with 1.3 million cases attributed to corneal opacity.[2] The most common causes of blindness worldwide include cataract (51%), glaucoma (8%), age-related macular degeneration (5%), and corneal opacities (4%).[3,4]

The cornea is a transparent and avascular connective tissue that plays a critical role in visual function by providing protective and refractive components. Corneal blindness results from diseases that alter corneal transparency, leading to opacities and visual loss.[5] The etiology of corneal blindness varies significantly across geographical regions and is influenced by local ocular diseases, environmental factors, and socioeconomic status.[4,6-13] For example, trachoma is the leading cause of bilateral blindness in Africa, whereas bacterial and fungal infectious keratitis are the most common etiologies in East, South, and Southeast Asia. In contrast, pseudophakic bullous keratopathy—corneal edema following cataract surgery—is the leading cause of corneal blindness in the United States, while keratoconus is the primary cause in Europe.[6-13]

In Thailand, the first Rapid Assessment of Avoidable Blindness survey reported a blindness prevalence of 0.6%, with corneal opacities accounting for 2% of total blindness cases.[14] A multicenter retrospective review by Prabhasawat et al found that the prevalence of corneal blindness was 1.7%. The leading causes were corneal infection (35.6%), surgical bullous keratopathy (27.8%), and trauma (14.0%).[6]

The management of corneal blindness depends on the underlying disease causing the vision loss. Corneal transplantation, or keratoplasty, is a preferred clinical modality and is the most common surgery performed for visual rehabilitation. Keratoplasty is also the most successful form of allogeneic organ transplantation because the avascular nature of the cornea minimizes graft rejection risk.[15] The procedure has evolved from replacing the full-thickness cornea—penetrating keratoplasty (PKP)—to replacing only the diseased layer—lamellar keratoplasty (LK).[16-19] The outcomes of keratoplasty vary based on the etiologies of corneal diseases in different regions, donor cornea quality, availability of skilled surgeons and medical teams, and effective long-term postoperative care.[20]

Currently, in many parts of the world, patients requiring donor corneas may wait 4 to 5 years due to a

shortage of donor tissue. The number of patients with corneal blindness has increased substantially in recent years, leading to a demand for corneal transplants that exceeds the availability of suitable donor corneas.[7,12,21] Relying solely on surgical intervention, even with increased donor corneas and intensive keratoplasty training for surgeons, may not suffice to address the growing waiting lists. Therefore, reviewing the leading indications for keratoplasty over the past decade is critical. Such analysis helps determine trends in preventable etiologies and assists in planning healthcare systems to mitigate these causes.

Meanwhile, keratoplasty techniques have evolved from PKP to LK. LK offers several advantages, including reduced astigmatism, increased postoperative structural stability, shorter rehabilitation time, decreased risk of immunological graft rejection, and the effective utilization of one donor cornea for two patients with different corneal layer diseases. This approach potentially reduces the number of patients waiting for donor corneas by half.[22-24]

This study aimed to explore the indications for keratoplasty from 2011 to 2020 to determine trends in preventable etiologies. Additionally, we sought to report the current keratoplasty techniques employed at our tertiary-care hospital.

MATERIALS AND METHODS

Study design and patient selection

We conducted a retrospective review of medical records of hospitalized patients who underwent keratoplasty at the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand, between January 2011 and December 2020. Data were retrieved from the Medical Information Department's database using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes:[25] 11.60 ("Corneal transplant, not otherwise specified"), 11.61 ("Lamellar keratoplasty with autograft"), 11.62 ("Other lamellar keratoplasty"), 11.63 ("Penetrating keratoplasty with autograft"), and 11.64 ("Other penetrating keratoplasty with homograft"). All keratoplasty procedures were performed by six cornea specialists and three cornea fellows at our hospital.

Ethical approval

This study was approved by the Siriraj Institutional Review Board (IRB number 452/2565 [IRB2]; approval number Si 599/2022).

Data collection

We collected data on patients' age, sex, indications for keratoplasty, and surgical techniques employed.

Exclusion criteria

Patients who underwent keratoplasty using tissues other than fresh donor corneas—such as sclera, amniotic membrane, or corneoscleral rim—were excluded from the study.

Classification of indications

Indications for keratoplasty were defined based on preoperative clinical diagnoses. These indications were categorized into 11 diagnostic groups, as detailed in Table 1 and illustrated in Fig 1.

Active infectious keratitis cases were further classified according to the pathogenic microorganisms involved: bacteria, fungus, Acanthamoeba, herpes virus (herpes simplex and herpes zoster), Mycobacterium species, and others.

Bullous keratopathy cases were subclassified based on etiology, including post-infection, post-cataract surgery, post-glaucoma surgery, post-vitreoretinal surgery, post-laser iridotomy or iridoplasty, and others, as shown in Table 1.

Regraft was defined as a repeat corneal transplantation due to irreversible loss of graft clarity. These cases were further subclassified based on the original or primary diagnosis that led to the first corneal transplantation (Table 1). The number of repeat keratoplasties, including the current operation, was recorded.

Other indications for keratoplasty were subclassified as presented in Table 1.

Classification of surgical techniques

The surgical techniques for keratoplasty were classified as PKP and LK. LK was subcategorized as deep anterior lamellar keratoplasty (DALK), Descemet stripping automated endothelial keratoplasty (DSAEK), and Descemet membrane endothelial keratoplasty (DMEK). The indications for each type of keratoplasty were recorded.

Statistical analysis

All statistical analyses were carried out using IBM SPSS Statistics for Windows, version 27.0. (IBM Corp., Armonk, NY, USA).

The indications for keratoplasty and the surgical techniques used over the 10-year study period are reported as numbers and percentages. For descriptive and analytical purposes, the study period was divided into two 5-year periods: from January 2011 to December 2015, and from January 2016 to December 2020.

We used Pearson's chi-square test and Fisher's exact test to assess differences in the types of indications for keratoplasty and the surgical techniques between the two

TABLE 1. Indications for keratoplasty over the ten-year study period (2011–2020).

Indications	N (%) (Total 754 eyes)	
Active infectious keratitis	197 (26.1)	
Fungus	105	(53.3)
Bacteria	42	(21.3)
Acanthamoeba	7	(3.6)
Mixed bacteria and fungus	6	(3.0)
Herpes	4	(2.0)
Mycobacteria	2	(1.0)
Unidentified organism	31	(15.7)
Bullous keratopathy	157 (20.8)	
Post cataract surgery	98	(62.4)
Post infection	18	(11.5)
Post glaucoma surgery	10	(6.4)
Post vitreoretinal surgery	9	(5.7)
Post laser iridotomy/iridoplasty	6	(3.8)
Post cataract and glaucoma surgery	5	(3.2)
Post cataract and retinal surgery	2	(1.3)
Unknown	9	(5.7)
Regraft	145 (19.2)	
Original diagnosis		
Bullous keratopathy	59	(40.7)
Active infectious keratitis	38	(26.2)
Corneal dystrophy	20	(13.8)
Corneal scar	10	(6.9)
Congenital opacity	10	(6.9)
Mechanical trauma	3	(2.0)
Corneal thinning/melting (non-infectious)	2	(1.4)
Burn	1	(0.7)
Unknown	2	(1.4)
Corneal dystrophy	86 (11.4)	
Fuchs endothelial dystrophy	64	(74.4)
Avellino dystrophy	10	(11.6)
Lattice dystrophy	8	(9.3)
Macular dystrophy	2	(2.3)
Unknown	2	(2.3)
Corneal scar	73 (9.7)	
Post infection	44	(60.3)
Post trauma	11	(15.1)
Post pterygium excision	2	(2.7)
Post ocular surface surgery	1	(1.4)
Unknown	15	(20.5)

TABLE 1. Indications for keratoplasty over the ten-year study period (2011–2020). (Continue)

Indications	N (%) (Total 754 eyes)	
Congenital opacity and anterior segment dysgenesis	23 (3.1)	
Iridocorneal endothelial syndrome	10	(43.5)
Peters anomaly	9	(39.2)
Sclerocornea	2	(8.7)
Corneal dermoid	1	(4.3)
Mucopolysaccharidosis	1	(4.3)
Corneal thinning/melting (non-infectious)	18 (2.4)	
SJS	9	(50.0)
Exposure keratopathy	3	(16.7)
Terrien marginal degeneration	3	(16.7)
Neurotrophic keratopathy	1	(5.5)
Unknown	2	(11.1)
Burn	17 (2.3)	
Chemical	15	(88.2)
Thermal	2	(11.8)
Corneal ectasia	16 (2.1)	
Keratoconus	15	(93.7)
Pellucid marginal degeneration	1	(6.3)
Mechanical trauma	12 (1.6)	
Immune keratitis	10 (1.3)	
Peripheral ulcerative keratitis	8	(80.0)
Mucous membrane pemphigoid	1	(10.0)
Chronic anterior uveitis	1	(10.0)

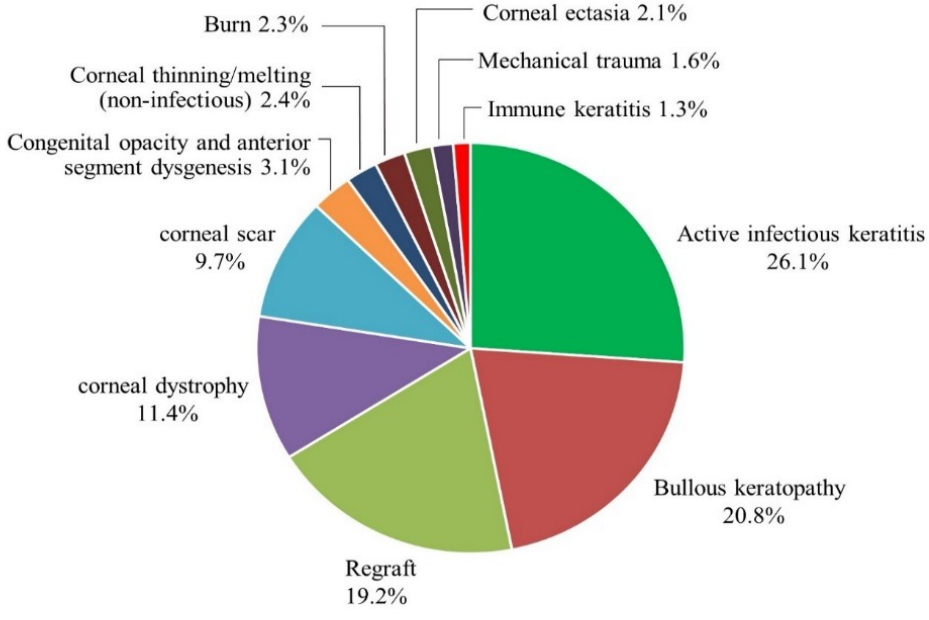


Fig 1. Indications for keratoplasty between 2011 and 2020

time periods. A *P* value less than 0.05 was considered statistically significant.

RESULTS

Demographic characteristics

A total of 754 eyes were included in the study. The mean age of the patients was 57.4 ± 18.0 years. The sex distribution was 434 male eyes (57.6%) and 320 female eyes (42.4%).

Overall distribution of keratoplasty indications

Over the 10-year study period, the most common indication for keratoplasty was active infectious keratitis (26.1%). This was followed by bullous keratopathy (20.8%), regraft (19.2%), corneal dystrophy (11.4%), and corneal scar (9.7%; Table 1 and Fig 1).

Infectious Keratitis: microbial analysis

The predominant etiology of active infectious keratitis was fungal infection, accounting for 53.3% of total microorganisms identified. The most frequently isolated fungus was *Fusarium* species (32.4% of fungal cases), followed by *Aspergillus* (13.3%), *Pythium* (8.6%), *Candida* (5.7%), and other fungi (40%).

Bacterial infections were the second most common cause of active infectious keratitis, representing 21.3% of total microorganisms. The most common causative bacterium was *Pseudomonas aeruginosa* (45.2% of bacterial cases). This was followed by other Gram-negative bacteria excluding *P. aeruginosa* (23.8%), Gram-positive bacteria (23.8%), and mixed bacterial infections (7.1%).

Bullous Keratopathy: etiological factors

Bullous keratopathy was the second most common indication for keratoplasty. It most frequently occurred after cataract surgery (62.4%), followed by post-infection cases (11.5%), post-glaucoma surgery (6.4%), and post-vitreoretinal surgery (5.7%; Table 1).

Regraft analysis

Regraft was the third most common indication for keratoplasty. The original diagnoses leading to regraft were bullous keratopathy (40.7%), active infectious keratitis (26.2%), corneal dystrophy (13.8%), and other conditions (19.3%; Table 1). The mean interval between the initial keratoplasty and the regraft was 7.0 ± 6.8 years.

Multiple keratoplasty procedures

Multiple keratoplasty procedures were performed on many patients. The distribution of repeat keratoplasties, including the current operation, was as follows: second keratoplasty (80.6% of regraft cases), third keratoplasty (16.6%), fourth keratoplasty (2.1%), and fifth keratoplasty (0.7%). The most frequent original diagnosis for the second, third, fourth, and fifth repeat keratoplasties was bullous keratopathy (38.5%, 45.8%, 66.7%, and 100%, respectively). The second most frequent original diagnosis for the second, third, and fourth repeat keratoplasties was active infectious keratitis (24.8%, 33.3%, and 33.3%, respectively).

Corneal dystrophy and scarring

Corneal dystrophy was the fourth most common indication for keratoplasty. Among these cases, Fuchs endothelial corneal dystrophy (FECD) was the most prevalent, accounting for 74.4% of corneal dystrophy instances (Table 1).

Corneal scar was the fifth most common indication. The majority of corneal scars resulted from post-infection (60.3%) and post-trauma (15.1%) events (Table 1).

TABLE 2. Comparison of the indications for keratoplasty between the first and the second five-year periods.

Indications	Number (%)			*P* value
	2011–2015 (n=326)	2016–2020 (n=428)	Total (n=754)	
Active infectious keratitis	85 (26.1)	112 (26.2)	197 (26.1)	0.115
Bullous keratopathy	67 (20.6)	90 (21.0)	157 (20.8)	
Regraft	58 (17.8)	87 (20.3)	145 (19.2)	
Corneal dystrophy	33 (10.1)	53 (12.4)	86 (11.4)	
Corneal scar	43 (13.2)	30 (7.0)	73 (9.7)	
Others	40 (12.3)	56 (13.1)	96 (12.8)	

TABLE 3. Comparison of the surgical techniques of keratoplasty between the first and the second five-year periods.

Surgical technique	Number (%)				*P* value
	2011–2015 (n=326)		2016–2020 (n=428)		
PKP	313	(96.0)	371	(86.7)	<0.001
DALK	8	(2.5)	19	(4.4)	0.58
DSAEK	5	(1.5)	0		0.06
DMEK	0		38	(8.9)	<0.001

Temporal trends in keratoplasty indications

The top five indications for keratoplasty in the first 5-year period (2011–2015) were similar to those in the second 5-year period (2016–2020). The percentages of keratoplasties performed for each indication did not differ significantly between the two periods ($P > 0.05$; Table 2).

Surgical techniques and their indications

Regarding surgical techniques over the entire study period, PKP was the most common procedure (90.7%). This was followed by DMEK (5.0%), DALK (3.6%), and DSAEK (0.7%).

Indications for PKP included active infectious keratitis (28.5%), bullous keratopathy (22.1%), regraft (20.2%), corneal scar (9.5%), corneal dystrophy (6.7%), congenital corneal opacity (3.2%), non-infectious corneal thinning or melting (2.6%), and other conditions (7.2%).

Indications for DALK were stromal corneal dystrophy (29.6%), corneal scar (25.9%), corneal ectasia such as keratoconus (22.2%), regraft (11.1%), active infectious keratitis (7.4%), and other causes (3.8%).

Indications for endothelial keratoplasty (EK), including DSAEK and DMEK, were FECD (74.4%), bullous keratopathy (14.0%), regraft (9.3%), and corneal scar (2.3%).

Changes in surgical techniques over time

The surgical techniques employed in the second 5-year period differed significantly from those in the initial 5-year period ($P < 0.01$; Table 3). The use of PKP decreased from 96.0% in the initial period to 86.7% in the second period. Conversely, the utilization of DMEK and DSAEK increased from 1.5% to 8.9%, and DALK slightly increased from 2.5% to 4.4%. DSAEK was performed only in the initial 5-year period. In the subsequent period, all EK procedures transitioned to DMEK, which was predominantly performed in cases of FECD.

DISCUSSION

According to the Thai Red Cross Eye Bank—the only eye bank in Thailand—the number of keratoplasties performed using corneas from eye bank donors increased from 759 in 2022 to 1546 in 2023. This rise resulted from proactive efforts in sourcing corneal tissues and providing intensive keratoplasty training courses. Nevertheless, as of June 2023, the waiting list for donor corneas comprised 7931 eyes, with patients potentially waiting 4 to 5 years for transplantation. Moreover, the number of new patients with corneal diseases requiring keratoplasty has increased rapidly, adding to the existing waiting lists.

Given this context, our study aimed to review the indications for keratoplasty over the decade from January 2011 to December 2020. We divided the data into two 5-year periods to identify trends in preventable or avoidable causes of corneal blindness. We also sought to determine whether the indications for keratoplasty had changed between the two periods. These findings should aid in strategic planning at the national level to prevent such corneal diseases.

Our study found that the three most common indications for keratoplasty during 2011–2020 were active infectious keratitis (26.1%), bullous keratopathy (20.8%), and regraft procedures (19.2%). This ranking mirrors the leading indications for keratoplasty in Asia reported by a systematic review of 34 years of global data.[10] In contrast, in the United States, the top three indications for corneal transplantation in 2014 were pseudophakic bullous keratopathy, regraft, and keratoconus, respectively.[10] Conversely, keratoconus was the primary indication for

corneal transplantation in Europe, Australia, the Middle East, Africa, and South America.[8,10]

Interestingly, the overall indications for keratoplasty in the second 5-year period did not significantly differ from those in the first 5-year period ($P > 0.05$). Active infectious keratitis remained the leading indication for keratoplasty throughout the decade. This trend aligns with findings from other studies in Asia,[7,10,26] including China,[26] India,[10] Nepal,[27] and Vietnam.[11] Previous reports in Thailand also indicated that corneal ulcers were the leading cause for keratoplasty.[6,28,29] Additionally, other studies showed that corneal ulcers were among the top three indications for keratoplasty in central and other regions of Thailand.[21,30]

Furthermore, the percentage of total infectious keratitis cases would actually increase if we included corneal scar cases resulting from inactive or healed keratitis (5.9%) along with active infectious keratitis (26.1%).

Infectious keratitis has both local and systemic risk factors. Local risk factors include contact lens wear, ocular trauma, ocular surface diseases, post-ocular surgery, and topical steroid administration. Systemic risk factors encompass conditions such as diabetes mellitus and autoimmune diseases.[31] Understanding these major risk factors in different countries is crucial for developing effective public health policies aimed at reducing the incidence of infectious keratitis.

Ting et al[31] reported that in the United Kingdom, Europe, North America, and Australia, the most common risk factor for infectious keratitis was contact lens wear. The predominant infectious microorganisms in these regions were bacteria, including *Staphylococcus spp*. and *Pseudomonas spp*. In contrast, in Asia, Africa, and the Middle East, the majority of infectious keratitis cases were caused by trauma, particularly exposure to vegetative materials, organic matter, and animal products in agricultural communities.[31-35] Consequently, the most common pathogens were fungi, such as *Fusarium spp*. and *Aspergillus spp*.

Similarly, a prospective multicenter study by the Asia Cornea Society reported that trauma was the most common risk factor for infectious keratitis in Asia (34.7%), followed by contact lens wear (10.7%).[36] The most frequently isolated microorganisms were *Fusarium* species (18.3%), *Pseudomonas aeruginosa* (10.7%), and *Aspergillus flavus* (8.3%).[36] Our results align with these findings, as fungi—including *Fusarium*, *Aspergillus*, *Pythium*, and *Candida*—remained the most common causative organisms.

The high prevalence of fungal keratitis in agricultural countries may be attributed to factors such as corneal trauma from plants or vegetative matter, hot climates, and socioeconomic status. Fungal keratitis is particularly challenging to treat due to delayed diagnosis, limited availability of effective antifungal agents, difficult access to new antifungal medications, and poor corneal penetration of existing treatments. These challenges result in a poor therapeutic response to current antifungal medications.

Bacterial pathogens were the second most common cause of infectious keratitis in our study, with *Pseudomonas aeruginosa* being the most prevalent, followed by other Gram-negative bacteria, Gram-positive bacteria, and mixed bacterial infections. Generally, bacterial keratitis responds well to antibacterial medications. However, some patients require keratoplasty due to large perforated corneal ulcers or severe bacterial keratitis at the time of referral, especially in diabetic or immunocompromised patients.

Infectious keratitis—particularly fungal keratitis—remains the leading indication for keratoplasty and has persisted as a significant problem for decades. Addressing this issue requires serious recognition and collaborative efforts among government leaders, public health officials, and ophthalmologists. Strategies should include increasing public education through social media, promoting early diagnosis and prompt, appropriate treatment of corneal ulcers by general ophthalmologists, providing research funding for the development of new effective antifungal agents, and implementing universal health coverage schemes that permit free access to antifungal medications such as liposomal amphotericin B, voriconazole, posaconazole, and future antifungal agents.

Bullous keratopathy emerged as the second leading indication for keratoplasty in our study, aligning with reports from Thailand[6,28] and other regions in Asia,[7,10] as well as Europe, Australia, the Middle East, and South America.[10] Notably, bullous keratopathy was reported as the most prevalent indication for keratoplasty in Japan,[7] Singapore,[7,37] and the United States.[10] While the primary indication for keratoplasty may vary based on geographical location, environmental factors, and socioeconomic status, bullous keratopathy consistently ranks as the first or second most common indication worldwide.[6-12,21]

The most frequent cause of bullous keratopathy is corneal endothelial injury following cataract surgery. This was evident in our study, where 62.4% of bullous keratopathy cases were attributed to post-cataract surgery, and our results align with findings in the literature.[6-12] Several factors contribute to this trend, including greater corneal endothelial damage resulting from both increasing

cataract surgery rates and a growth in the number of trainees, anterior chamber intraocular lens implantations,[38] and phacoemulsification surgery in hard nucleus or brunescent cataracts. Patients at high risk for corneal endothelial decompensation—such as those with elderly corneas, FECD, shallow anterior chambers, or histories of multiple intraocular surgeries—are particularly susceptible.

To mitigate this issue, preventative strategies could include the implementation of surgical simulators in cataract surgery training programs for residents and discontinuing the use of anterior chamber intraocular lens implantation. Additionally, converting phacoemulsification procedures to extracapsular cataract extraction or manual small incision cataract surgery in high-risk patients may also minimize endothelial damage.[39,40] Preoperative evaluations should assess central corneal thickness and corneal endothelial cell count and morphology in these patients to determine the most appropriate surgical approach—whether phacoemulsification, extracapsular cataract extraction, or manual small incision cataract surgery. Furthermore, utilizing advanced phacoemulsification machines with enhanced safety features, when possible, may decrease the risk of endothelial injury.

Regraft procedures were the third most common indication for keratoplasty in our study, mirroring findings from previous reports in Thailand[28,30] and regions such as Asia, Australia, and South America.[10] Interestingly, in New Zealand, regraft or repeat corneal transplantation was the most prevalent indication.[41] In our cohort, the primary diagnoses preceding the first keratoplasty were bullous keratopathy (40.7%), active infectious keratitis (26.2%), corneal dystrophy (13.8%), and other conditions (Table 1). This suggests that active infectious keratitis not only remains the first leading indication but also contributes significantly to the need for regrafts.

Bullous keratopathy was the foremost indication for regraft in our study, consistent with findings from Mezaine et al in Saudi Arabia.[42] In contrast, infectious keratitis was the primary indication for single and multiple repeat keratoplasties in China,[43] and it was the second most common indication for regrafts in our study. Several factors may contribute to corneal graft failures in bullous keratopathy cases, including immunological rejection, corneal vascularization, secondary glaucoma, donor corneal endothelial cell function, surgical techniques, and postoperative follow-up and management.[44,45]

In cases where keratoplasty was performed due to active infectious keratitis, graft failure or rejection was often associated with recurrent infections on the corneal graft, persistent inflammation at the time of surgery, vascularization in recipient corneas, and secondary elevated intraocular pressure—more common in fungal keratitis. Additional factors include the use of lower-quality donor corneas (as high-quality corneas are typically reserved for patients with better prognoses) and the omission of corticosteroids or immunosuppressive drugs in early postoperative treatment for fungal keratitis patients. Therefore, active infectious keratitis significantly influences both the primary and tertiary indications for keratoplasty, underscoring the need to address it as a national public health concern to reduce the demand for corneal transplants. Notably, multiple repeat keratoplasties were also frequently necessitated by infectious keratitis.

Addressing these challenges may reduce regraft rates. Prevention of infectious keratitis—particularly fungal keratitis—is crucial. Improvements can be achieved by providing high-quality donor corneas, enhancing surgical techniques and instruments to prevent recurrent infections and minimize endothelial damage, and ensuring robust long-term postoperative care and follow-up by medical teams. Patient education on the importance of regular follow-up, medication adherence, and recognition of symptoms indicative of graft rejection is essential. Additionally, implementing universal health coverage that grants free access to immunosuppressive medications—such as cyclosporine, mycophenolate mofetil, and tacrolimus—could aid in preventing and treating graft rejection.

Patients with corneal blindness awaiting keratoplasty suffer from long-term vision loss, reduced quality of life, loss of income, and a need for assistance. This condition represents not only a social burden but also an economic one. Therefore, preventive measures are crucial public health strategies to reduce corneal blindness and diminish the necessity for corneal transplants.

During the study period, PKP accounted for 90.7% of all keratoplasties performed. It remains the most common form of keratoplasty in Thailand[21] and Asia.[7,10] LK constituted 9.3% of procedures, including DMEK at 5.0%, DALK at 3.6%, and DSAEK at 0.7%.

LK was already being performed in 2011, at the start of our study period, with DALK and DSAEK as the available options. Later, DMEK was introduced in 2016. The surgical techniques employed in the second 5-year study period differed significantly from those in the initial 5-year period ($P < 0.05$). PKP decreased from 96.0% to 86.7%, while EK—including DMEK and DSAEK—increased from 1.5% to 8.9%. DALK slightly increased from 2.5% to 4.4%. Notably, DSAEK was performed only during the initial 5-year period (2011–2015). After 2015, all EK procedures transitioned to DMEK in the second 5-year period.

Surgeons preferred DMEK over DSAEK, considering it superior in several aspects. DMEK involves fewer surgical steps and instruments, resulting in less time-consuming operations. Surgeons at our hospital have found that DMEK achieves faster visual recovery, better visual outcomes, reduced interface irregularity, less refractive error, and lower rates of immune rejection, similar to a report by the American Academy of Ophthalmology.[46] Currently, DMEK is performed mostly in eyes with FECD. These LK techniques can reduce regraft rates, lessen the use of immunosuppressive drugs, and allow one donor cornea to benefit more than one patient on the waiting list.

This study has several limitations. First, as a retrospective study, it may have incomplete data due to inadequate documentation in medical records. Second, it was conducted at a single tertiary referral hospital in central Thailand, which may not reflect the national situation of corneal blindness among patients awaiting keratoplasty. However, reviews of multiple reports in Thailand,[6,28-30] including a report from the Thai Red Cross Eye Bank,[21] allowed us to summarize that the primary indication for keratoplasty was infectious keratitis (corneal ulcer) or corneal scar, depending on the research institute. Infectious keratitis was the most common indication in tertiary referral hospitals, whereas corneal scar was the most prevalent indication in non-referral hospitals and the Thai Red Cross Eye Bank. Bullous keratopathy and regraft (graft failure) consistently ranked within the top three indications for keratoplasty in all reports.

Lastly, due to the evolution of EK, corneal dystrophy—mainly FECD—has significantly increased the rates of keratoplasty, especially DMEK. In the future, FECD may become a more common indication for keratoplasty, although perhaps not as high as for active infectious keratitis.

CONCLUSIONS

Active infectious keratitis was the most common indication for keratoplasty, followed by bullous keratopathy and regraft respectively. This study highlights the critical need for prevention, early detection, and appropriate management of infectious keratitis, especially fungal keratitis. Achieving this requires collaborative efforts among government leaders, public health officials, and ophthalmologists. Promoting public education through social media and facilitating free access to antimicrobial agents and immunosuppressive drugs for post-keratoplasty management are also essential.

PKP remained the most common surgical technique due to severe infectious keratitis. However, the use of LK—including EK and DALK—significantly increased for other indications during the second 5-year study period. Moreover, DMEK was primarily performed in eyes with FECD.

Data Availability Statement

The participants of this study did not give written consent for their data to be shared publicly, so due to the sensitive nature of the research supporting data is not available.

ACKNOWLEDGMENTS

The authors are grateful to Assistant Professor Chulaluk Komoltri from the Office for Research and Development for her assistance with the statistical analysis.

DECLARATION

Grants and Funding Information

No funding was received for this study.

Conflicts of Interest

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

Author Contributions

Conceptualization and methodology, W.B., C.J., C.C., P.N., and S.K.; data collection, C.J.; data analysis and interpretation, W.B., C. J.; writing—original draft, W.B., and C.J.; writing—review and editing, W.B., C.C., and P.N. All authors read and agreed to the published version of the manuscript.

Use of Artificial Intelligence

Not applicable

REFERENCES

1. World Health Organization International Statistical Classification of Diseases and Related Health Problems 10th revision Current version Version for 2003 Chapter VII H54 Blindness and low vision.
2. Bourne RRA, Flaxman SR, Braithwaite T, Cicinelli MV, Das A, Jonas JB, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Health. 2017;5(9):888-97.
3. Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. Lancet Glob Health. 2017;5:de1221-34.
4. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. Br J Ophthalmol. 2012;96(5):614-8.
5. Tidke SC, Tidake P. A Review of Corneal Blindness: Causes

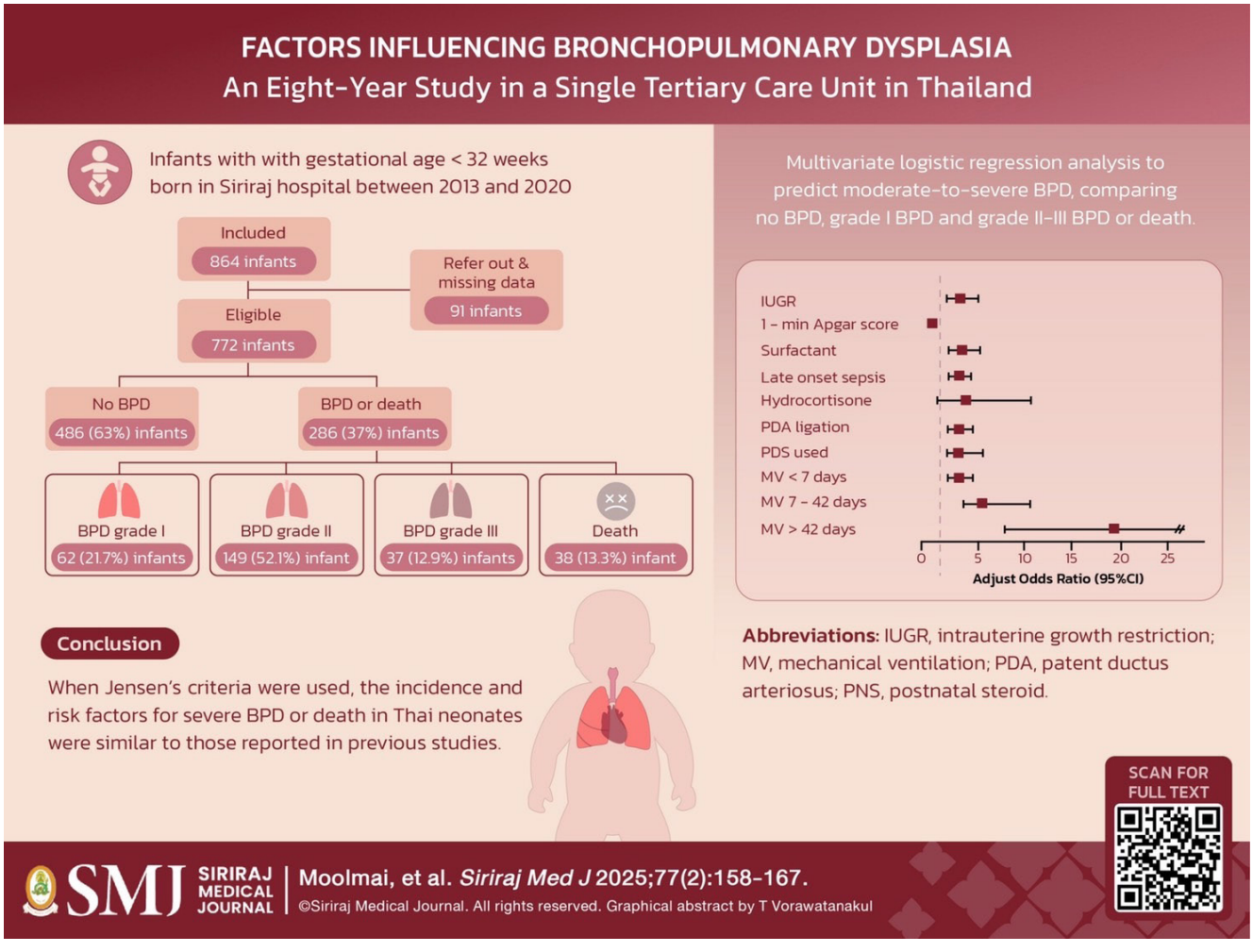
and Management. Cureus 2022;14(10):e30097.
6. Prabhasawat P, Trethipwanit K, Prakairungthong N, Narenpitak S, Jaruroteskulchai S, Anantachai J. Causes of Corneal Blindness: A Multi-Center Retrospective Review. J Med Assoc Thai. 2007; 90(12):2651-7.
7. Tran TM, Duong H, Bonnet C, Kashanchi A, Buckshey A, Aldave AJ. Corneal Blindness in Asia: A Systematic Review and Meta-Analysis to Identify Challenges and Opportunities. Cornea. 2020; 39(9):1196-205.
8. Mathews PM., Lindsley K, Aldave AJ, Akpek EK. Etiology of Global Corneal Blindness and Current Practices of Corneal Transplantation: A Focused Review. Cornea. 2018;37(9):1198-203.
9. Seidu MA, Olusanya BA, Ogundipe AO. Prevalence and determinants of corneal blindness in a Semi-Urban population of southwest Nigeria. Niger J Ophthalmol. 2017;25:18-27.
10. Matthaei M, Sandhaeger H, Hermel M, Adler W, Jun AS, Cursiefen C, et al. Changing indications in penetrating keratoplasty: a systematic reviews of 34 years of global reporting. Transplantation. 2017;101:1387-99.
11. Dong PN, Han TN, Aldave AJ, Chau HT. Indications for and techniques of keratoplasty At Vietnam National Institute of Ophthalmology. Int J Ophthalmol. 2016;9:379-83.
12. Park CY, Lee JK, Gore PK, Lim CY, Chuck RS. Keratoplasty in the United States: A 10-year Review from 2005 through 2014. Ophthalmology. 2015;122:2432-42.
13. Oliva MS, Schottman T, Gulati M. Turning the tide of corneal blindness. Indian J Ophthalmol. 2012;60:423-7.
14. Isipradit S, Sirimaharaj M, Charukamnoetkanok P, Thonginnetra O, Wongsawad W, Sathornsumetee B, et al. The first rapid assessment of avoidable blindness (RAAB) in Thailand. PLoS One. 2014;9(12):e114245.
15. Gain P, Jullienne R, He Z, Aldossary M, Acquart S, Cognasse F, et al. Global Survey of Corneal Transplantation and Eye Banking. JAMA Ophthalmol. 2016;134(2):167-73.
16. Anwar M, Teichmann K.D. Deep lamellar keratoplasty: surgical techniques for anterior lamellar keratoplasty with and without baring of Descemet's membrane. Cornea. 2002;21:374-83.
17. Price FW Jr, Price M.O. Descemet's stripping with endothelial keratoplasty in 50 eyes: a refractive neutral corneal transplant. J Refract Surg. 2005;21:339-45.
18. Gorovoy M.S. Descemet-stripping automated endothelial keratoplasty. Cornea. 2006;25:886-9.
19. Melles GR, Ong TS, Ververs B, van der Wees J. Descemet membrane endothelial keratoplasty (DMEK). Cornea. 2006;25: 987-90.
20. Singh R, Gupta N, Vanathi M, Tandon R. Corneal transplantation in the modern era. Indian J Med Res. 2019;150(1):7-22.
21. Pariyakanok L, Erjongmanee S, Saonanon P. Indications for corneal transplantation in Thailand between 1996 and 2008. Asian Biomedicine. 2011;5:843-8.
22. Price MO, Giebel AW, Fairchild KM, Price FW Jr. Descemet's membrane endothelial keratoplasty: prospective multi-center study of visual and refractive outcomes and endothelial survival. Ophthalmology. 2009;116:2361-8.
23. Anshu A, Price MO, Price FW Jr. Risk of corneal transplant rejection significantly reduced with Descemet's membrane endothelial keratoplasty. Ophthalmology. 2012;119:536-40.
24. Sharma N, Agarwal P, Titiyal JS, Kumar C, Sinha R, Vajpayee RB. Optimal use of donor corneal tissue: one cornea for two recipients. Cornea. 2011;30:1140-4.
25. "ICD-9-CM Diagnosis Codes—International Classisfication of Diseases—Medical Diagnosis Codes." Available from: www.findacode.com. Retrieved 2015-12-30.
26. Tan H, Lin M, Gou Q, Li A, Gu F, Liu Q, et al. Trends in Corneal Transplantation and Characteristics of Donors in the Chongqing Eye Bank, China: A Retrospective Study, 1999-2018. Front Med (Lausanne). 2021;8:750898.
27. Bajracharya L, Gurung R, Demarchis EH, Oliva M, Ruit S, Tabin G. Indications for keratoplasty in Nepal: 2005-2010. Nepal J Ophthalmol. 2013;5(2):207-14.
28. Poonyathalang A, Simaroj P. Penetrating Keratoplasty in Ramathibodi Hospital (1981-1992). Thai J Ophthalmol. 1993;7(1):13-20.
29. Ausayakhun S, Juntaramanee J. Clinical Indications for Penetrating Keratoplasty in Maharaj Nakorn Chiang Mai Hospital, 1990-1995. Thai J Ophthalmol. 1997;11(1):17-23.
30. Saranyakoob W, Horatanaruang O. Indications and Outcomes of Penetrating Keratoplasty in Songkhlanagarind Hospital from 1996-2003. Thai J Ophthalmol. 2006;20(1):43-52.
31. Ting DSJ, Ho CS, Deshmukh R, Said DG, Dua HS. Infectious keratitis: an update on epidemiology, causative microorganisms, risk factors, and antimicrobial resistance. Eye. 2021;35(4):1084-101.
32. Xie L, Dong X, Shi W. Treatment of fungal keratitis by penetrating keratoplasty. Br J Ophthalmol. 2001;85(9):1070-4.
33. Xie L, Zhong W, Shi W, Sun S. Spectrum of fungal keratitis in north China. Ophthalmology. 2006;113(11):1943-8.
34. Karsten E, Watson SL, Foster LJR. Diversity of microbial species implicated in keratitis: a review. Open Ophthalmol J. 2012;6:110.
35. Kengpunpanich S, Prabhasawat P, Gem C, Chirapapaisan C, Booranapong W, Ngowyutagon P. Therapeutic Penetrating Keratoplasty for Severe Fungal Keratitis in a Thai Tertiary Care Center. Siriraj Med J. 2023;75:575-83.
36. Khor WB, Prajna VN, Garg P, Mehta JS, Xie L, Liu Z, et al. The Asia Cornea Society Infectious Keratitis Study: A Prospective Multicenter Study of Infectious Keratitis in Asia. Am J Ophthalmol. 2018;195:161-70.
37. Tan DT, Janardhanan P, Zhou H, Chan YH, Htoon HM, Ang LP, et al. Penetrating keratoplasty in Asian eyes: The Singapore Corneal Transplant Study. Ophthalmology. 2008;115(6):975-82.
38. Auffarth GU, Wesendahl TA, Brown SJ, Apple DJ. Complications after implantation of anterior chamber lenses. An analysis of 4,100 explanted intraocular lenses. Ophthalmologe. 1994;91(4): 512-7.
39. Bourne RR, Minassian DC, Dart JK, Rosen P, Kaushal S, Wingate N. Effect of cataract surgery on the corneal endothelium: modern phacoemulsification compared with extracapsular cataract surgery. Ophthalmology. 2004;111(4):679-85.
40. Tabin G, Chen M, Espandar L. Cataract surgery for the developing world. Curr Opin Ophthalmol. 2008;19(1):55-9.
41. Chilibeck CM, Brookes NH, Gokul A, Kim BZ, Twohill HC, Moffatt SL, et al. Changing Trends in Corneal Transplantation in Aotearoa/New Zealand, 1991 to 2020: Effects of Population Growth, Cataract Surgery, Endothelial Keratoplasty, and Corneal Cross-Linking for Keratoconus. Cornea. 2022;41(6): 680-7.
42. Al-Mezaine H, Wagoner MD; King Khaled Eye Specialist Hospital Cornea Transplant Study Group. Repeat penetrating keratoplasty:

indications, graft survival, and visual outcome. Br J Ophthalmol. 2006;90(3):324-7.
43. Wan X, Yao W, Zhao S, Xu J, Le Q. Indications and Surgical Techniques for Repeat Corneal Transplantation in Eastern China: A Twelve-Year Study. J Ophthalmol. 2021;2021:5514004.
44. Patel SV, Diehl NN, Hodge DO, Bourne WM. Donor risk factors for graft failure in a 20-year study of penetrating keratoplasty. Arch Ophthalmol. 2010;128(4):418-25.
45. Chotikavanich S, Prabhasawat P, Satjapakasit O. Ten-Year Survival of Optical Penetrating Keratoplasty and Risk Factors for Graft Failure in Thai Patients. J Med Assoc Thai. 2020;103(9):883-90.
46. Deng SX, Lee WB, Hammersmith KM, Kuo AN, Li JY, Shen JF, et al. Descemet Membrane Endothelial Keratoplasty: Safety and Outcomes: A Report by the American Academy of Ophthalmology. Ophthalmology. 2018;125(2):29.

Factors Influencing Bronchopulmonary Dysplasia: An Eight-Year Study in a Single Tertiary Care Unit in Thailand

Pallapa Moolmai, M.D.[1], Prattana Rattanachamnongk[2], Buranee Yangthara, M.D., Ph.D.[3], Punnanee Wutthigate, M.D.[3,*]

[1]Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [2]Registered Nurse and Neonatal Nurse Practitioner, Siriraj Hospital, Bangkok, Thailand, [3]Division of Neonatology, Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.



**Corresponding Author: Punnanee Wutthigate*
E-mail: punnanee.wut@mahidol.edu
Received 23 September 2024 Revised 5 November 2024 Accepted 5 November 2024
ORCID ID:http://orcid.org/0000-0003-2990-5412
https://doi.org/10.33192/smj.v77i2.271246

ABSTRACT

Objective: To assess the incidence trends of severe BPD or death, identify associated risk factors, and develop a predictive model using Jensen's BPD grading system.

Materials and Methods: This retrospective study included infants with a gestational age (GA) < 32 weeks born between 2013 and 2020. Infants were classified into no BPD and BPD (all grades) or death categories. Risk factors and a predictive model for outcomes were identified using multivariable logistic regression and evaluated with a receiver operating characteristic (ROC) curve.

Results: Among the 772 infants, 286 (37%) were classified into the BPD group. The incidence of BPD continuously increased from 16.3% in 2013 to 49.4% in 2020. Multivariate analyses revealed factors associated with severe BPD, including IUGR, the 1-minute Apgar score, surfactant administration, late-onset sepsis, hydrocortisone, PDA ligation, postnatal steroid and mechanical ventilation (MV) days. The highest adjusted odds ratio (aOR) was for MV > 42 days at 19.29 (95% CI; 7.22–51.55; p<0.001). The area under the curve (AUC) was 0.898, with 80.68% sensitivity and 84.83% specificity.

Conclusion: When Jensen's criteria were used, the incidence and risk factors for severe BPD or death in Thai neonates were similar to those reported in previous studies. The regression model exhibited good predictive value, potentially assisting clinicians in targeted interventions.

Keywords: Bronchopulmonary dysplasia (BPD); respiratory outcomes; mortality; Thailand (Siriraj Med J 2025; 77: 158-167)

INTRODUCTION

Bronchopulmonary dysplasia (BPD) is the most important respiratory complication in preterm infants. Over time, advancements in technologies as well as neonatal care treatment strategies, including surfactant replacement therapy, new models of mechanical ventilation such as targeted tidal volume, high-frequency oscillatory ventilation (HFOV), noninvasive ventilation (NIV) and postnatal steroids, have increased the survival rate of preterm infants, especially extremely preterm infants born at 22–24 weeks.[1] However, the incidence rate of BPD remains unchanged. For infants born at a gestational age (GA) of less than 29 weeks, it is estimated to be between 48% and 68%, with the incidence inversely proportional to GA.[2] In the post-surfactant era, with the widespread adoption of NIV for implementing protective lung strategies in very preterm infants, the new definition of BPD aligns with the criteria set forth by the National Institutes of Health (NIH) in 2018[3] and Jensen in 2019.[4] However, the NIH 2018 definition has various limitations, with a primary concern being its ability to predict long-term outcomes.[5] In 2019, Jensen addressed and resolved this particular issue.[6,7]

Although numerous studies have explored risk factors associated with BPD outcomes in the past, few studies have focused on the Asian population, particularly in Thailand. This gap can be attributed to limited resources, variations in equipment, differing treatment guidelines across countries, and the use of different BPD definitions to assess outcomes. Geetha et al. reported that the prevalence and risk factors for BPD among extremely low gestation age neonates (ELGANs) with a composite outcome of moderate to severe BPD/death was 67%; however, they used the NIH 2001 definition, which may not accurately reflect current clinical practices.[8] Siriraj Hospital is a university hospital with a level IV NICU in Thailand and has an annual delivery rate of approximately 150 infants born premature with a GA less than 32 weeks.

Our study investigated the incidence of moderate to severe BPD and/or mortality and explored the associated risk factors among preterm infants, thereby providing comprehensive insights that reflect the broader landscape of the neonatal population in Thailand.

MATERIALS AND METHODS

We conducted a retrospective cohort data collection from the Siriraj Informatics and Data Innovation Center (SiData+) to include neonates who were born at a GA < 32 weeks between January 2013 and December 2020. Infants with major congenital anomalies (trisomy 13, 18, or 21 or hydrops fetalis), congenital heart disease or congenital lung and airway malformation were excluded. The definition of BPD was defined by Jensen in 2019[4], categorized BPD severity based on the mode of respiratory support administered at 36 weeks postmentrual age (PMA). Infants on nasal cannula 2 L/min were classified as grade

I BPD, while those on nasal cannula greater than 2 L/min or non-invasive respiratory support were classified as grade II BPD. Infants on mechanical ventilation at 36 weeks PMA were identified as grade III BPD.

Outcome measurements

The maternal demographic data included maternal age, antenatal steroids, maternal complications that lead to preterm labor and the risk for BPD in infants, such as diabetes mellitus (DM), hypertensive disorders, and chorioamnionitis, which was defined as a clinical diagnosis of intra-amniotic infection[9], intrauterine growth restriction (IUGR), and multiple pregnancies. The demographic data of the neonates included GA, birth weight (BW), sex, and Apgar scores at 1 and 5 minutes.

Neonatal outcomes

The neonatal outcomes included intubation in the delivery room (DR), respiratory comorbidities such as respiratory distress syndrome (RDS), surfactant administration, pulmonary hemorrhage, persistent pulmonary hypertension of the newborn (PPHN), and inotropic drug use during the first week of life. Late-onset sepsis (LOS) was defined as positive hemoculture occurring after 72 hours of life or negative blood culture, but clinicians decided to administer antibiotics for more than 5 days on the basis of clinical sepsis. Necrotizing enterocolitis (NEC) stage ≥ 2 was diagnosed using modified Bell's criteria.[10] The hemodynamic significance of the patent ductus arteriosus (HsPDA) was defined as a ductal diameter exceeding 1.5 mm with a left-to-right shunt and a left atrium to aorta diameter ratio (LA:Ao) greater than 1.5.

Respiratory outcome measurements

The modes of assisted ventilatory support were collected and specified as invasive mechanical ventilation (IMV), HFOV, and NIV. The postmenstrual age (PMA) at the time of the last extubation and the total number of days of mechanical ventilation (MV) were recorded. Postnatal steroids (PNS) were used for prolonged mechanical ventilation after 2 weeks of life. The respiratory severity index (RSI; mean airway pressure (MAP) multiplied by FiO_2 ($MAP \times FiO_2$)) was calculated on Days 7, 14, 21 and 28 of life. The BPD definition reported by Jensen in 2019[4] was chosen to assess the severity of BPD in this cohort. During the time of the study, pulmonary hypertension (PHT) resulting from BPD was not routinely addressed. Nevertheless, infants underwent echocardiography when facing clinical instability and when an acute pulmonary hypertensive crisis was suspected.

Statistical analysis

SPSS version 22.0 (SPSS Inc., Chicago, IL, USA) was used for all the statistical analyses. Continuous variables are expressed as the means ± SD or medians (interquartile ranges), and categorical variables are expressed as counts and percentages. The chi-square test and Fisher's exact test were used for categorical variables, and Student's t test was used for continuous variables for the appropriate comparisons of interest. A p-value < 0.05 was considered statistically significant. Significant variables with p-values < 0.2 were included in the model for univariate analyses, and the associations between potential risk factors for BPD were calculated. A multivariate logistic regression model was developed to identify independent significant risk factors for BPD of any grade or death and BPD Grades II and III or death, and odds ratios (ORs) with 95% confidence intervals (CIs) are reported.

RESULTS

Maternal and neonatal characteristics

A total of 941 infants were born at GA < 32 weeks. In addition to 77 infants with congenital anomalies, 92 infants were excluded from our cohort: 89 were transferred to other hospitals before BPD diagnosis, and three infants had incomplete data. The final analysis included a total of 772 infants (Fig 1). Among these, 248 (32.1%) infants developed BPD, and 38 (13.3%) infants died before 36 weeks PMA. The overall incidence rates of BPD Grades I, II, and III were 21.7% (62/286), 52.1% (149/286) and 12.9% (37/286), respectively. The cohort experienced a mortality rate of 6.5% (50/772), and the majority of infants died from respiratory complications.

We categorized the infants into two groups: those without BPD and those with BPD of any grade or who died. No statistically significant difference in maternal characteristics was noted between the two groups, except for a greater prevalence of IUGR in the BPD group (39 (8%) *vs.* 36 (12.6%), $p = 0.04$) (Table 1). Compared with those without BPD, infants in the BPD groups had lower GA, lower birth weights, more male infants and lower Apgar scores at 1 minute and 5 minutes (Table 1).

Neonatal comorbidities

Infants diagnosed with BPD were more likely to be critically ill than those without BPD. This was evident in significantly increased neonatal complications, such as higher rates of intubation at the delivery room, surfactant administration, pulmonary hemorrhage, PPHN, and hemodynamic instability requiring inotropic drugs, as well as hydrocortisone treatment. Additionally, there was a higher incidence of late-onset neonatal sepsis and NEC

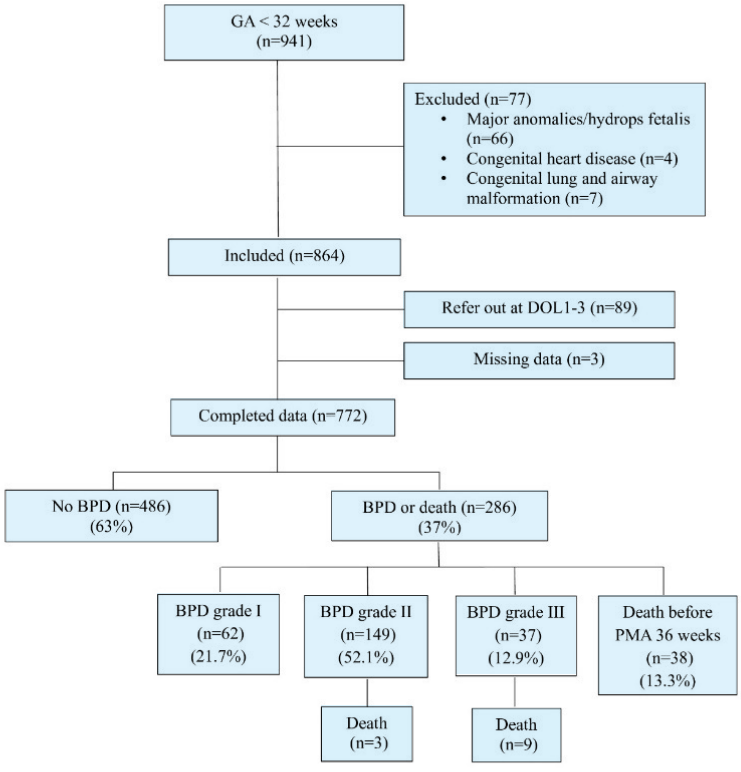


Fig 1. Summary flow chart
Abbreviations: BPD, bronchopulmonary dysplasia; DOL, day of life; GA, gestational age; PMA, postmenstrual age.

TABLE 1. Demographic data for patients with no BPD and those with BPD of all grades or death.

	No BPD (n=486)	BPD or death (n=286)	*p*-value
Maternal			
Maternal age (year)	29.41 (± 6.9)	29.89 (± 7.2)	0.36
Antenatal steroid	451 (92.8)	260 (90.9)	0.35
Completed dose	287 (59.1)	152 (53.1)	0.11
Diabetes mellitus	55 (11.3)	31 (10.8)	0.84
Hypertensive disorder	110 (22.6)	62 (21.7)	0.76
Chorioamnionitis	33 (6.8)	27 (9.4)	0.18
IUGR	39 (8.0)	36 (12.6)	0.04
Cesarean section	288 (59.3)	182 (63.6)	0.23
Multiple pregnancy	121 (24.9)	68 (23.8)	0.73
Neonatal			
GA (weeks)	30.1 (± 1.5)	27.9 (± 2.0)	< 0.001
$24^{0/7}$-$26^{6/7}$ weeks	23 (4.7)	103 (36.0)	< 0.001
$27^{0/7}$-$31^{6/7}$ weeks	463 (95.3)	183 (64.0)	< 0.001
Birth weight (grams)	1371.5 (± 348.8)	985.3 (± 314.6)	< 0.001
Male	256 (52.7)	173 (60.5)	0.04
1-minute Apgar score	7 [5-8]	4 [1-6]	< 0.001
5-minute Apgar score	9 [8-9]	7 [5-8]	< 0.001

Expressed as the count (%), mean± SD, median [IQR].
Abbreviations: GA, gestational age; IUGR, intrauterine growth restriction.

stage ≥ 2 (Table 2). Among infants without BPD, 22.8% received PDA treatment, whereas 59.8% in the other groups did (p<0.001). For infants receiving medications to close the PDA, 90.1% of infants without BPD were successfully closed before discharge. However, it was closed in only 59.6% of infants with BPD (p<0.001). There was no difference in the age at PDA ligation between the groups (31.5 (±1.2) *vs*. 30.5 (±3.4) weeks; p=0.36) (Table 2). Nevertheless, when infants with each BPD grade were compared, infants with BPD Grades II and III experienced more complications, with the exception of pulmonary complications, hydrocortisone use and stage ≥ 2 NEC (Supplemental Tables 1-2).

Respiratory variables

The total incidence of BPD or death was 37%. Among infants born at < 29 weeks, the incidence of BPD was 61.3% (152/248), and the incidence of Grade II–III BPD was 49% (122/248). Compared with those without BPD, infants with BPD had an extended duration of mechanical ventilation (0 [0–3] *vs*. 15 [3–44]; p<0.001), and their length of hospital stay was longer. The PMA at discharge was 42.9 (± 12.3) *vs*. 37.4 (2.9) weeks; p<0.001 (Table 2). We investigated the type of respiratory support and RSI on Days 7, 14, 21, and 28. The percentages of infants with BPD who required MV on Days 7, 14, 21, and 28 were 56% (139), 50% (124), 44.4% (110), and 40.3%

TABLE 2. Neonatal comorbidities between the no BPD group and the all-grade BPD or death group.

	No BPD (n=486)	BPD or death (n=286)	*p*-value
Intubation at DR	110 (22.6)	194 (67.8)	< 0.001
Intubation at NICU during admission	192 (39.5%)	258 (90.2%)	< 0.001
Surfactant administration	39 (8.0)	116 (40.6)	< 0.001
Repeated dose of surfactant	0	13 (4.5)	< 0.001
Pulmonary hemorrhage	10 (2.1)	32 (11.2)	< 0.001
PPHN	6 (1.2)	36 (12.6)	< 0.001
iNO initiation	0	9 (3.1)	< 0.001
Use of inotropic drugs	73 (15.0)	152 (53.1)	< 0.001
Use of hydrocortisone	5 (1.0)	36 (12.8)	< 0.001
Late onset sepsis	91 (18.7)	141 (51.5)	< 0.001
NEC stage ≥ 2	69 (14.2)	58 (20.9)	0.02
Surgical treatment	6 (8.7)	18 (31)	0.001
Treatment of PDA	111 (22.8)	171 (59.8)	< 0.001
Medical closure	100 (90.1)	102 (59.6)	< 0.001
PDA ligation	11 (9.9)	69 (42.9)	< 0.001
Age at PDA ligation (days)	23 [16-25]	19 [15-27]	0.85
PMA at PDA ligation (weeks)	31.5 (± 1.2)	30.5 (±3.4)	0.36
Postnatal steroid	19 (3.9)	73 (26.2)	< 0.001
PHT at 34-36 weeks PMA	7 (7.1)	19 (6.6)	< 0.001
Total MV days	0 [0-3]	15 [3-44]	< 0.001
PMA at discharge (weeks)	37.4 (± 2.9)	42.9 (± 12.3)	< 0.001

Expressed as the count (%), mean ± SD, median [IQR].

Abbreviations: DR, delivery room; iNO, inhaled nitric oxide; MV, mechanical ventilation; NEC, necrotizing enterocolitis; PDA, patent ductus arteriosus; PHT, pulmonary hypertension; PMA, postmenstrual age; PPHN, persistent pulmonary hypertension of the newborn.

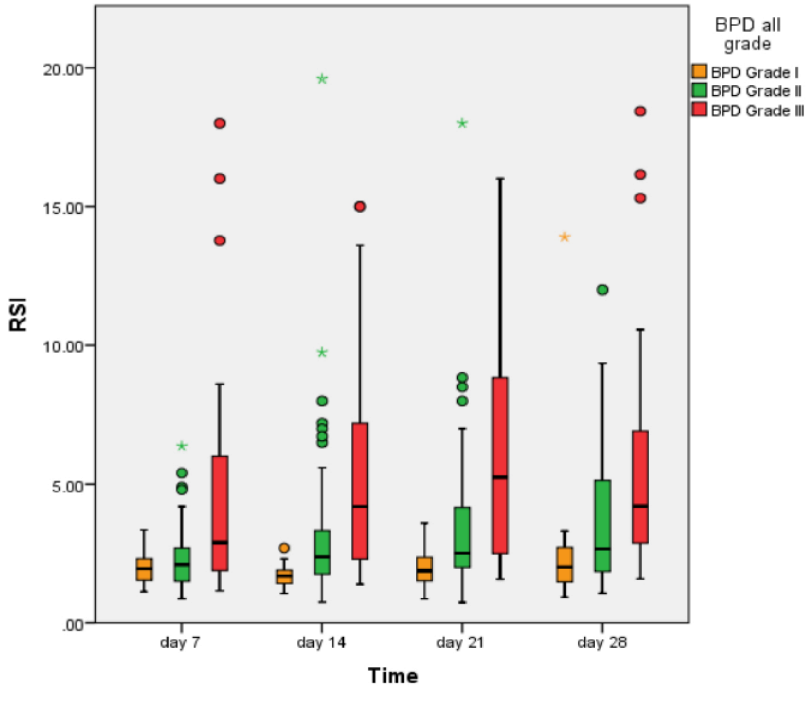
(100), respectively. In contrast, a smaller proportion of infants without BPD were on MV on these days (10.3% (50/486), 5.3% (26/486), 4.5% (22/486), and 3.7% (18/486), respectively). Infants diagnosed with Grade III BPD presented a significantly elevated RSI, which increased over time at Days 7, 14, and 21 (2.89 [1.84–6.13], 4.20 [2.24–7.50], and 5.25 [2.40–9.10], respectively). However, there was a slight decrease in the RSI from Day 28 to 4.21 [2.85-7.41] (Fig 2). A total of 18 (2.3%) of the 772 infants with BPD required home oxygen, and two of them underwent tracheostomy before discharge with home ventilation.

Model for the prediction BPD Grades II-III or death

Univariate analyses were conducted to predict BPD outcomes by comparing the no BPD group with the all-grade BPD or death group (Supplemental Table 3). Multivariate logistic regression analysis identified a subset of independent risk factors for BPD of any grade or death (Table 3) and BPD Grades II-III or death (Table 4). The receiver operating characteristic (ROC) curve was constructed utilizing the final model, and the cutoff value was 0.3 according to the Youden index, resulting in a sensitivity of 80.68%, specificity of 84.83%, positive predictive value (PPV) of 66.80%, and negative predictive value (NPV) of 92.06% (Supplemental Fig 1). The final equation is shown in the Appendix.

DISCUSSION

We demonstrated the incidence of BPD in preterm infants born at < 32 weeks of gestation using the new severity-based diagnostic criteria proposed in 2019 by the NICHD Neonatal Research Network.[4] The overall incidence of BPD or death was 37%, with 24% of those

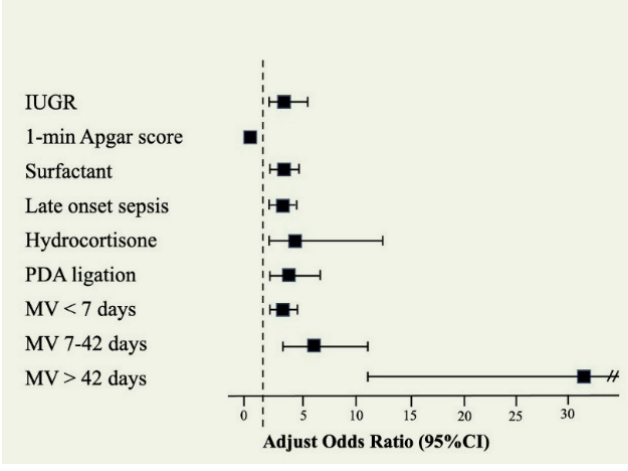


Expressed as median [IQR].

Postnatal day	N	BPD Grade I	BPD Grade II	BPD Grade III	*p*-value
7	139	1.99 [1.50-2.36]	2.10 [1.49-2.70]	2.89 [1.84-6.13]	0.003
14	123	1.69 [1.37-1.99]	2.38 [1.76-3.42]	4.20 [2.24-7.50]	0.004
21	110	1.88 [1.51-2.38]	2.52 [2.00-4.20]	5.25 [2.40-9.10]	0.001
28	117	2.01 [1.44-2.80]	2.67 [1.85-5.14]	4.21 [2.85-7.41]	0.027

Fig 2. Respiratory severity index (RSI) at 7, 14, 21 and 28 DOL in inats with BPD.

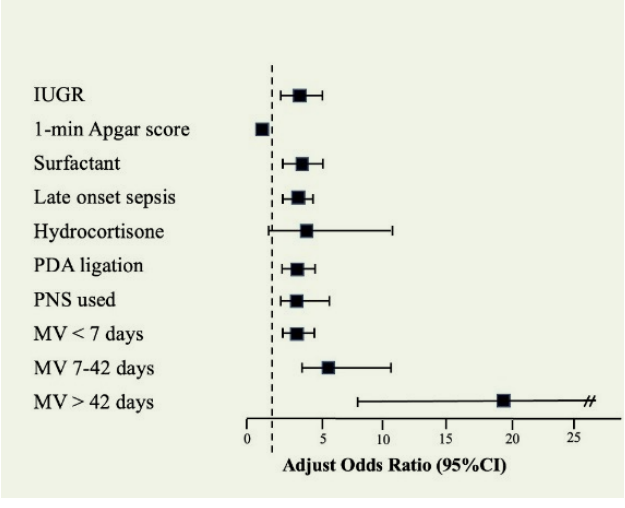
TABLE 3. Multivariate logistic regression analysis of risk factors for BPD compared between the no BPD group and the all-grade BPD or death group.



Total n=772	BPD/ death N=286	aOR	95%CI	*p* -value
75	36	2.57	1.33 - 4.97	0.005
771	286	0.85	0.78 - 0.92	< 0.001
155	116	2.58	1.52 - 4.38	< 0.001
232	141	2.12	1.36 – 3.31	0.001
41	36	4.21	1.39 - 12.75	0.011
80	69	2.89	1.34 - 6.23	0.007
202	72	2.24	1.28 - 3.89	0.004
167	110	6.24	3.43 - 11.33	< 0.001
81	76	32.38	11.01 - 95.19	< 0.001

Abbreviations: IUGR, intrauterine growth restriction; MV, mechanical ventilation; PDA, patent ductus arteriosus.

TABLE 4. Multivariate logistic regression analysis to predict moderate-to-severe BPD, comparing no BPD, grade I BPD and grade II-III BPD or death.



Total n=772	BPD II-III/ death N=224	aOR	95%CI	*p* -value
75	30	2.075	1.06 - 4.08	0.03
771	224	0.850	0.78 - 0.93	< 0.001
155	101	2.491	1.47 - 4.21	0.001
232	119	2.242	1.40 - 3.59	0.001
41	35	3.144	0.94 - 10.52	0.06
80	59	2.160	1.10 - 4.25	0.03
92	72	2.588	1.16 - 5.76	0.02
202	54	2.261	1.13 - 4.53	< 0.001
167	84	5.602	2.75 - 11.41	< 0.001
81	71	19.294	7.22 - 51.55	< 0.001

Abbreviations: IUGR, intrauterine growth restriction; MV, mechanical ventilation; PDA, patent ductus arteriosus; PNS, postnatal steroid.

who survived until 36 weeks PMA being diagnosed with Grade II-III BPD or death. Notably, for ELGANs, the incidence increased to 61.3%, and 49% of the surviving infants exhibited Grade II-III BPD. Compared with international data, specifically the 2022 figures from the Canadian Neonatal Network (CNN)[11], where the incidence of BPD in infants younger than 33 weeks was reported as 32.6%, and considering Yang et al.'s findings[12] of a 61% BPD rate in infants born before 32 weeks in their study (utilizing the NIH 2001 definition), our results fall within an acceptable range. Nevertheless, within the ELGANs group, our observed rate of BPD surpassed that reported by Jensen et al.[13] In their cohort, the incidence of BPD, which was determined using a definition similar to ours, was 45.5%. In global cohorts, the overall BPD incidence varies widely, with values of 25–56% in Asia, 17–73% in Europe and 18–75% in North America.[14] In addition, a notable rate of 67% was reported in a Singaporean cohort.[8] However, they used different BPD definitions for grading severity,

which might not accurately reflect recent practices and outcomes. Previously, at our center, we did not have a unit protocol for respiratory management in preterm infants < 32 weeks; however, all neonatologists followed standard guidelines for RDS treatment, such as early CPAP and rescue surfactant therapy, using conventional methods with slow weaning. The less invasive surfactant administration (LISA) method, oxygen reduction test (ORT) to determine BPD at 36 weeks PMA, and respiratory management protocol for preterm infants < 32 weeks were introduced in 2018. A comparable trend was observed in the ELGANs population. However, infants exhibited greater critical illness during this period than during the earlier period, with an increase in IUGR cases, lower birth weight, lower 1-minute Apgar scores, an elevated need for surfactant treatment, an increased incidence of LOS and a greater requirement for postnatal steroids.

BPD is a chronic lung disease that results from multiple factors, including genetic, prenatal, or postnatal influences. Antenatal corticosteroids are considered the standard of care for accelerating fetal lung maturity to prevent RDS but not BPD. Owing to the widespread use of antenatal steroids following the WHO guidelines15 to improve preterm birth outcomes, there were no differences in antenatal steroid usage between the non-BPD and BPD groups in our study, which is consistent with recent previous studies.[13,16] The use of antenatal steroids in our unit was consistent with the broader cohort. Our findings identify the factors associated with a composite of Grades II-III or death, including eight factors (Table 4), which aligns with known factors reported in other studies.[12,16-18]

Timing of surfactant replacement therapy (SRT) is one of the crucial factors that determine the outcome of BPD. A previous systematic review demonstrated that intubated infants who received early selective surfactant administration had a decreased risk of BPD or death at 28 days compared to those who received late SRT (after more than 2 hours).[19] However, data from a recent cohort indicated that early SRT was associated with a longer duration of mechanical ventilation, with no significant difference in the rate of BPD.[20] In our cohort, we do not have recorded data on the timing of surfactant administration due to the retrospective nature of the study. European Consensus guidelines on the management of respiratory distress syndrome recommended that intubated infants less than 30 weeks' gestation should receive SRT and rescue surfactant should be administered early in the course of the disease using a thin catheter when infants required CPAP pressure greater than 6 cmH_2O or FiO_2 > 0.3.[21]

We opted to exclude PDA treatment from the model, as numerous randomized controlled trials (RCTs) comparing early treatment with conservative approaches have failed to demonstrate a significant impact on the increased rate of BPD or death outcomes.[22,23] Given the absence of a standardized definition for HsPDA, infants in our units underwent echocardiography based on the presence of a murmur. The decision to pursue medical closure was exclusively determined using criteria involving the ductal diameter and the LA:Ao ratio. We considered that this could be a confounding factor leading to excessive PDA treatment, making the establishment of a direct association with the outcome of BPD challenging. Therefore, we chose to include only PDA ligation in the model, which is supported by prior research indicating an association between PDA ligation and moderate to severe BPD.[24,25]

One commonly recognized factor contributing to BPD is chorioamnionitis. However, in our cohort, no significant difference was observed between the two groups. We hypothesize that the clinical diagnosis of chorioamnionitis may lead to over recognition. Only half of the clinically diagnosed cases of chorioamnionitis had confirmed placental pathology indicating acute inflammation, as per unpublished local data. The longer the duration of MV is, the greater the increase in the incidence and severity of BPD as well as the likelihood of requiring home oxygen at discharge and experiencing neurodevelopmental impairment.[26] We observed a similar trend in our cohort, where infants with MV days had a significantly increased OR.

The strengths of this study include the use of our own cohort to represent our practices and to develop a care bundle for improving BPD outcomes. We also plan to create an BPD calculator for the Thai population. Another strength is the adoption of the Jensen 2019 BPD definition, which aligns more closely with contemporary practices and allows for better prediction of short-term and long-term outcomes. This study has several limitations. First, it involves a small cohort and is representative of a single center. Second, the data retrieval spans from 2013, and treatment strategies may have evolved. Thus, recent populations were potentially not captured. Third, our study does not incorporate long-term outcomes, such as neurodevelopmental impairment.

CONCLUSION

The overall incidence of BPD in our study correlated with the global incidence. Factors such as IUGR, the 1-minute Apgar score, surfactant administration, late-onset sepsis, hydrocortisone use for hemodynamic instability, PDA ligation, and total MV days independently contribute

to the risk of all grades of BPD. Our findings suggest that prospective research could utilize our developed model for external validation to predict the probability of Grade II–III BPD in Asian preterm infants.

ACKNOWLEDGEMENT

None

DECLARATION

Grants and Funding Information

None

Conflicts of Interest

None

Author Contributions

Conceptualization and methodology, P.M., P.W., P.R.; Investigation, P.M., P.R., P.W.; Formal analysis, P.M., B.Y., P.W.; Visualization and writing - original draft, P.M. and P.W.; Writing - review and editing, B.Y. and P.W.; Supervision, P.W. All authors have read and agreed to the final version of the manuscript.

Data Availability

Derived data generated will be shared upon reasonable request to the corresponding author.

Ethics Statement

The study was approved by the Siriraj Institutional Review Board of the Faculty of Medicine Siriraj Hospital, Mahidol University (CoA no. Si 298/2022).

REFERENCES

1. Bell EF, Hintz SR, Hansen NI, Bann CM, Wyckoff MH, DeMauro SB, et al. Mortality, In-Hospital Morbidity, Care Practices, and 2-Year Outcomes for Extremely Preterm Infants in the US, 2013-2018. JAMA. 2022;327(3):248-63.
2. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. Pediatrics. 2010;126(3):443-56.
3. Higgins RD, Jobe AH, Koso-Thomas M, Bancalari E, Viscardi RM, Hartert TV, et al. Bronchopulmonary Dysplasia: Executive Summary of a Workshop. J Pediatr. 2018;197:300-8.
4. Jensen EA, Dysart K, Gantz MG, McDonald S, Bamat NA, Keszler M, et al. The Diagnosis of Bronchopulmonary Dysplasia in Very Preterm Infants. An Evidence-based Approach. Am J Respir Crit Care Med. 2019;200(6):751-9.
5. Stoecklin B, Simpson SJ, Pillow JJ. Bronchopulmonary dysplasia: Rationale for a pathophysiological rather than treatment based approach to diagnosis. Paediatr Respir Rev. 2019;32:91-7.
6. Li J, Xu H. Comparisons of two definitions of bronchopulmonary dysplasia for the premature infants. Pediatr Pulmonol. 2022;57(1):217-23.
7. Katz TA, van Kaam AH, Schuit E, Mugie SM, Aarnoudse-Moens CSH, Weber EH, et al. Comparison of New Bronchopulmonary Dysplasia Definitions on Long-Term Outcomes in Preterm Infants. J Pediatr. 2023;253:86-93.e4.
8. Geetha O, Rajadurai VS, Anand AJ, Dela Puerta R, Huey Quek B, Khoo PC, et al. New BPD-prevalence and risk factors for bronchopulmonary dysplasia/mortality in extremely low gestational age infants ≤28 weeks. J Perinatol. 2021;41(8):1943-50.
9. Higgins RD, Saade G, Polin RA, Grobman WA, Buhimschi IA, Watterberg K, et al. Evaluation and Management of Women and Newborns With a Maternal Diagnosis of Chorioamnionitis: Summary of a Workshop. Obstet Gynecol. 2016;127(3):426-36.
10. Walsh MC, Kliegman RM. Necrotizing enterocolitis: treatment based on staging criteria. Pediatr Clin North Am. 1986;33(1):179-201.
11. Chronic lung disease: Gestational age < 33 weeks [Internet]. Canadian Neoantal Network. 2022. Available from: https://www.canadianneonatalnetwork.org/portal/Portals/0/Annual%20Reports/2022%20CNN%20Annual%20Report.pdf.
12. Yang T, Shen Q, Wang S, Dong T, Liang L, Xu F, et al. Risk factors that affect the degree of bronchopulmonary dysplasia in very preterm infants: a 5-year retrospective study. BMC Pediatr. 2022;22(1):200.
13. Jensen EA, Edwards EM, Greenberg LT, Soll RF, Ehret DEY, Horbar JD. Severity of Bronchopulmonary Dysplasia Among Very Preterm Infants in the United States. Pediatrics. 2021;148(1):e2020030007.
14. Siffel C, Kistler KD, Lewis JFM, Sarda SP. Global incidence of bronchopulmonary dysplasia among extremely preterm infants: a systematic literature review. J Matern Fetal Neonatal Med. 2021;34(11):1721-31.
15. WHO recommendations on Antenatal corticosteroids for improving preterm birth outcomes [Internet]. World Health Organization. 2021. Available from: https://www.who.int/publications/i/item/9789240057296.
16. Shen X, Patel N, Zhu W, Chen X, Lu K, Cheng R, et al. A nomogram for predicting the risk of Bronchopulmonary dysplasia in premature infants. Heliyon. 2023;9(8):e18964.
17. Jassem-Bobowicz JM, Klasa-Mazurkiewicz D, Żawrocki A, Stefańska K, Domżalska-Popadiuk I, Kwiatkowski S, et al. Prediction Model for Bronchopulmonary Dysplasia in Preterm Newborns. Children (Basel). 2021;8(10):886.
18. Yazici A, Buyuktiryaki M, Simsek GK, Kanmaz Kutman HG, Canpolat FE. Factors associated with neurodevelopmental impairment in preterm infants with bronchopulmonary dysplasia. Eur Rev Med Pharmacol Sci. 2022;26(5):1579-85.
19. Bahadue FL, Soll R. Early versus delayed selective surfactant treatment for neonatal respiratory distress syndrome. Cochrane Database Syst Rev. 2012;11(11):CD001456.
20. Wongkhonkaen C, Paopongsawan P, Jirapradittha J, Kiatchoosakun P. Outcomes Comparison of Early versus Late Surfactant Replacement Therapy in Neonates with Respiratory Distress Syndrome. Siriraj Med J. 2023;75(5):330-42.
21. Sweet DG, Carnielli VP, Greisen G, Hallman M, Klebermass-Schrehof K, Ozek E, et al. European Consensus Guidelines on the Management of Respiratory Distress Syndrome: 2022 Update. Neonatology. 2023;120(1):3-23.
22. Clyman RI, Liebowitz M, Kaempf J, Erdeve O, Bulbul A, Håkansson S, et al. PDA-TOLERATE Trial: An Exploratory

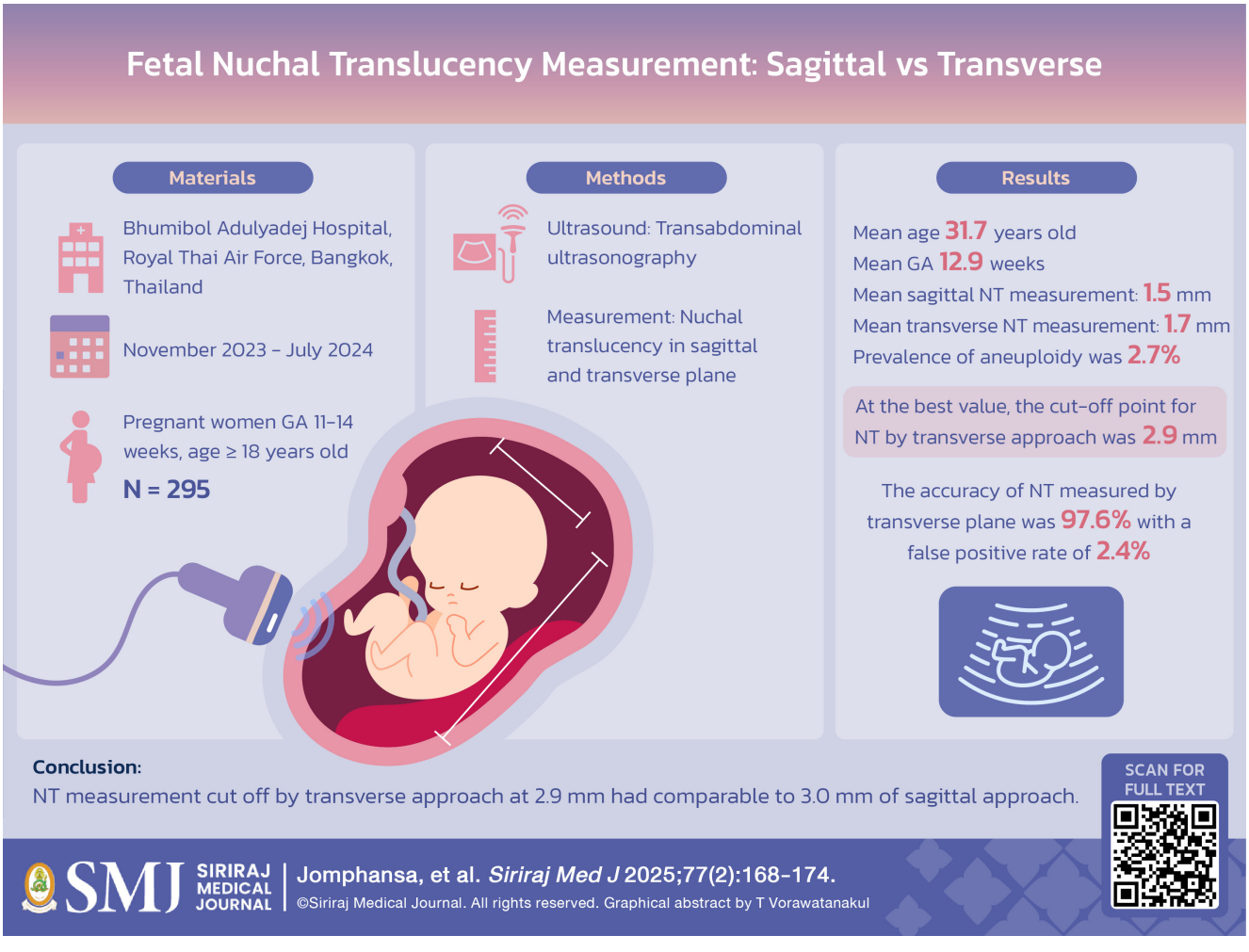
Randomized Controlled Trial of Treatment of Moderate-to-Large Patent Ductus Arteriosus at 1 Week of Age. J Pediatr. 2019;205: 41-8.e6.

23. Hundscheid T, Onland W, Kooi EMW, Vijlbrief DC, de Vries WB, Dijkman KP, et al. Expectant Management or Early Ibuprofen for Patent Ductus Arteriosus. N Engl J Med. 2023;388(11): 980-90.
24. Garg A, Athiraman N, Jenkins E, Crossland D, Loganathan PK. Is early patent ductus arteriosus ligation helpful in premature neonates? A 10-year retrospective study. World J Pediatr. 2022; 18(5):350-7.
25. Wang H, Jain A, Weisz DE, Moraes TJ. Trends in patent ductus arteriosus ligation in neonates and changes in outcomes: A 10-year multicenter experience. Pediatr Pulmonol. 2021;56(10): 3250-7.
26. Menshykova AO, Dobryanskyy DO. Duration of mechanical ventilation and clinical outcomes in very low birth weight infants: A single center 10-years cohort study. J Neonatal Perinatal Med. 2023;16(4):673-80.

Correlation Between Sagittal and Transverse Plane Fetal Nuchal Translucency Measurement

Pananya Jomphansa, M.D.[1], Suthasinee Mataneedol, M.D.[1,*], Sinart Prommas, M.D.[1], Buppa Smanchat, M.D.[1], Kornkarn Bhamarapravatana, Ph.D.[2], Komsun Suwannarurk, M.D.[3]

[1]Department of Obstetrics and Gynecology, Bhumibol Adulyadej Hospital, Royal Thai Air Force, Thailand. [2]Department of Preclinical Science, Faculty of Medicine, Thammasat University, Pathumthani, Thailand. [3]Department of Obstetrics and Gynecology, Faculty of Medicine, Thammasat University, Pathumthani, Thailand.



**Corresponding author: Suthasinee Mataneedol*
E-mail: sutha_koi@hotmail.com
Received 3 October 2024 Revised 3 November 2024 Accepted 3 November 2024
ORCID ID:http://orcid.org/0009-0001-2356-6803
https://doi.org/10.33192/smj.v77i2.271485

ABSTRACT

Objective: This study aimed to evaluate and compare nuchal translucency (NT) measurement between sagittal and transverse planes for aneuploidy screening.

Materials and Methods: This prospective study was conducted at the antenatal clinic, department of obstetrics and gynecology, Bhumibol Adulyadej Hospital, Royal Thai Air Force, Bangkok, Thailand between November 2023 and July 2024. Subjects were singleton pregnant women who underwent fetal chromosome anomaly screening by NT measurement of both sagittal and transverse planes by transabdominal ultrasonography between 11 and 14 weeks of gestational age (GA). Demographic and clinical data were collected and evaluated. Relationship between sagittal and transverse planes measurement was analyzed.

Results: A total of 295 pregnant women were recruited. The mean age and GA of participants were 31.7 years old and 12.9 weeks, respectively. Prevalence of aneuploidy was 2.7% (8/295). The mean NT measurement in the sagittal plane was significantly lower than in the transverse plane (1.5 and 1.7 mm). Transverse and sagittal plane measurement exhibited a linear correlation. The cut-off point of NT by transverse and sagittal approach were 2.9 mm and 3.0 mm, respectively that giving comparable diagnostic value.

Conclusion: NT measurement cut-off by transverse approach at 2.9 mm had comparable to 3.0 mm of sagittal approach.

Keywords: Nuchal translucency; aneuploidy; pregnancy; transverse plane; ultrasound (Siriraj Med J 2025; 77: 168-174)

INTRODUCTION

Fetal malformation or genetic abnormalities should be diagnosed as early as possible.[1] Ultrasound screening in the first trimester is mainly performed by fetal nuchal translucency (NT), with serum analytes, namely pregnancy associated plasma protein-A (PAPP-A), human chorionic gonadotropin (hCG) and alpha-fetoprotein (AFP) between 10 and 14 weeks of gestation were introduced by American college of Obstetricians and Gynecologists (ACOG) recommendation in year 2020.[1] NT is a measurement of subcutaneous development behind the fetus' neck during the first trimester of pregnancy.[1,2] Thick NT signifies fetal heart failure, extracellular matrix abnormality, or abnormality of the lymph system.[1] Meanwhile, ultrasound screening in the second and third trimester is mainly performed to predict of neonatal outcomes.[3-6]

NT is the most effective tool for detecting Down Syndrome during the first trimester of pregnancy around 11-14 weeks of gestation at nearly 90%, with only a 5% false positive rate.[7,8] Thick NT (of more than 3.0 mm) measured by median sagittal view was associated with fetal anomalies or cardiac anomalies, this allows for early management to prevent high risk of spontaneous fetal loss.[1,7,9-11]

In current practice NT is measured using the sagittal plane. The sagittal plane necessitates an experienced sonographer and is successful only if the fetus holds a specific position.[12,13] A transverse approach was an easier operation to perform and offered a shorter learning curve.[13] If the measurement could be performed using a transverse plane, the result would be less dependent on the position of the fetus. This investigation aimed to explore sagittal and transverse plane NT imaging as a tool for aneuploidy risk assessment.

MATERIALS AND METHODS

A prospective study was conducted in the Obstetrics and Gynecology Department at Bhumibol Adulyadej Hospital (BAH), Royal Thai Air Force, Bangkok, Thailand between November 2023 and July 2024. This study was approved by the BAH Institutional Review Board in 2023 (IRB No.84/66).

Participants were singleton pregnant women between 11 and 14 weeks of gestation (GA). Inclusion criteria was pregnant women with intrauterine pregnancy, age 18 and up, who attended antenatal care at BAH. Exclusion criteria included inability to communicate in Thai. All participants were counseled about fetal chromosome screening (FCS). All participants received briefings about the investigation and all gave written informed consents. According to BAH protocol, subjects were screened either using maternal serum for cell-free DNA (cfDNA) screening detection or quadruple test. The cfDNA was performed by taking maternal blood samples. Maternal serum was sent for fetal microDNA labs. Quadruple tests consisted of maternal serum AFP, hCG, estriol and inhibin. Standard amniocentesis for fetal chromosome study was performed on subjects who had abnormal FCS either via cfDNA or quadruple test methods. Amniocentesis was performed during 15 and 20 weeks of GA.

During FCS routine protocol, all participants underwent transabdominal ultrasound (Voluson E10 model: GE healthcare, Zipf, Austria) for NT measurement. The measurement of NT in sagittal plane was performed three times. In addition, the same sonographer measured on the transverse plane using transabdominal ultrasound three times. Sonographer was member of the maternal-fetal medicine (MFM) fellowship under the supervision of experienced MFM staff.

According to International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) recommendation, NT measurement was performed via sagittal plane.[14] During NT measurement, fetuses were aligned in neutral position. Fetal head, neck, tip of nose, upper thorax and NT edge margin were demonstrated in midsagittal plane during NT measurement.[14]

During NT measurement via transverse plane, transverse plane of fetal head was aligned similarly to the transcerebellar view.[15,16] Frontal horns, cerebral peduncles and optic thalami were the landmarks for NT measurement. Distance from the external contour of the occipital bone to the outer contour of the skin by positioning ultrasound calipers was the NT value.[13]

Data from sagittal and transverse measurement was collected and further analyzed. Demographic and clinical data were also collected from electronic medical records including maternal age, parity and body mass index.

The data was processed using the Statistic Package for the Social Sciences (SPSS version 27.0). A p-value less than 0.05 was considered statistically significant. The mean values of each measurement plane were used for calculation. Transverse and sagittal NT were compared and evaluated. The accuracy of transverse measurement was assessed and compared to the NT value of more than 3.0 mm. Scatter plots and linear correlations during regression analysis were used for strengthening the interpretation of the results.

RESULTS

A total of 295 cases were included in the study as presented in Fig 1. Half (164/295) of participants were of nulliparity. Mean age and GA of participants were 31.7 years old and 12.9 weeks, respectively. The total of amniocentesis was 30 cases. Only 8 cases had a report of abnormal chromosome study, namely trisomy 21 (6 cases) and trisomy 18 (2 cases). Prevalence of aneuploidy was 2.7% (8/295). NT measurement by sagittal approach and transverse approach averaged 1.5 and 1.7 mm, respectively as shown in Table 1.

Comparison of NT measurement between sagittal and transverse approaches by scatter plot were presented in Fig 2. Correlation between sagittal and transverse approaches followed a linear pattern as shown in Fig 3. Standard cut-off point of NT by sagittal approach was chosen at 3.0 mm.[1] Area under curve (AUC) was at 0.980. At the best value, the cut-off point for NT by transverse approach was 2.9 mm, giving the same diagnostic value as the cut-off point of NT by sagittal approach at 3.0 mm. We proposed that subjects who had NT transverse measurement of 2.9 mm or greater represented a high risk of aneuploidy at the same level as those who had 3.0 mm measurement using a sagittal approach. The accuracy of NT measured by transverse approach was 97.6% with a false positive rate of 2.4% (7/291).

A transformation model of NT by transverse approach to sagittal approach was created. The formula was 0.597 + [0.528*transverse(mm)] as depicted in Fig 2.

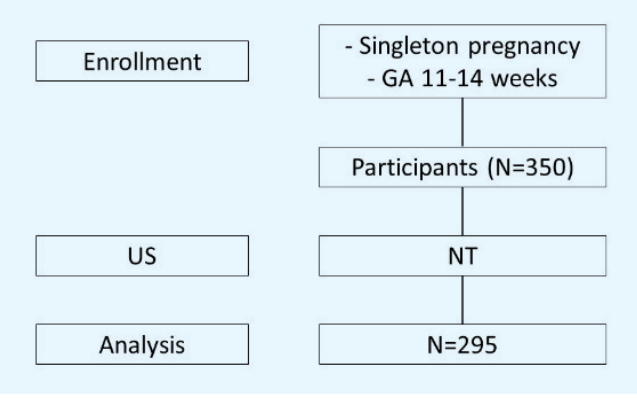


Fig 1. Flow chart of study
Abbreviations: US: ultrasound, NT: nuchal translucency, GA: gestational age

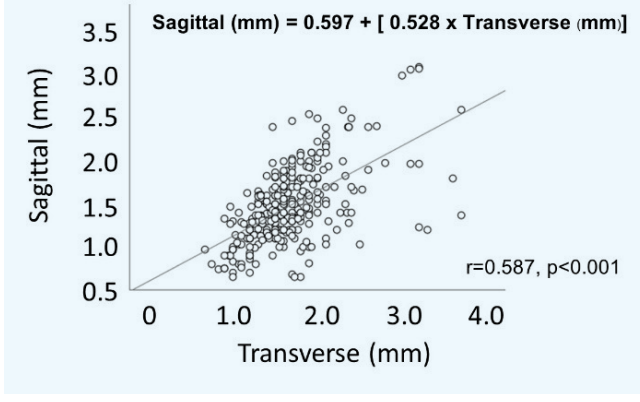


Fig 2. Comparison of nuchal translucency (NT) measurement between sagittal and transverse approach by scatter plot
Abbreviations: Sag: sagittal plane, Trans: transverse plane, MD: mean different

TABLE 1. Demographic character of participants who had positive and negative test result for aneuploidy screening

	Total (n=295)	Range	Positive (n=30)	Negative (n=265)	*p*-value
Age (year)	31.7 ± 5.5	18.0 - 44.0	35.9 ± 3.7	31.2 ± 5.5	<0.001*
BMI (kg/m^2)	25.7 ± 4.6	15.8 - 39.4	26.5 ± 5.6	25.6 ± 4.5	0.368
GA (days)	90.6 ± 4.7	77.0 - 98.0	89.4 ± 5.9	90.8 ± 4.6	0.144
CRL (mm)	66.9 ± 9.1	36.7- 87.0	65.6 ± 10.1	67.1 ± 9.0	0.421
Nulliparity	164 (55.6)		11 (36.7)	153 (57.7)	0.037*
Sagittal (mm)	1.5 ± 0.4	0.65 - 3.1	1.8 ± 0.6	1.4 ± 0.4	0.001*
Transverse (mm)	1.7 ± 0.5	0.67 - 3.2	1.9 ± 0.5	1.7 ± 0.5	0.030*

Abbreviations: BMI: body mass index, GA: gestational age, CRL: crown-rump length

DISCUSSION

In the current study, the aneuploidy rate was 2.7% from chromosome study via amniocentesis. Thai study in year 2024, prevalence of fetal aneuploidy was 0.33% (13/3928) .[17] Previous studies reported NT measurement conducted by transabdominal ultrasonography (TAS) in pregnant women with average age between 28.9 and 33 years old.[10,11,13,18,19] In 2022 Montaguti from Italy a prevalence of aneuploidy from chorionic villous sampling of 0.6% (6/1023).[13] Prevalence of aneuploidy from a Vietnamese study utilizing chromosome study from amniocentesis in 2011 was 1.5%.[18] Yin's study in China in 2022 performed NT measurement and amniocentesis for karyotype study in all cases and reported the prevalence of aneuploidy from at 3.4%.[19] Montaguti performed NT measurement and genetic screening by combined test or noninvasive prenatal testing (NIPT). Subjects who had high risk of trisomy 13,18, or 21 (45 cases) from combined testing and NIPT underwent chorionic villous sampling reporting only 6 cases.[13] Tomai performed NT measurement and genetic screening by combined test. Subjects who had high risk of trisomy 13,18, or 21 from combined test underwent amniocentesis.[18] Low aneuploidy prevalence of the current study might interfere the diagnostic efficacy of screening tool.

In the current study, subjects underwent NT measurement and genetic screening either using quadruple test or cfDNA testing. Subjects from the current study who had high risk of aneuploidy screening underwent amniocentesis. There were no aneuploidy cases with NT less than 3.0 mm and low risk of genetic screening. There was no indication for amniocentesis for karyotype in every pregnancy. However, high prevalence of karyotype was reported from the center that performed all karyotype study or high-risk subjects. According to Yin's 2022 study, pregnant women with mean age of 29 and NT between 2.5 and 2.9 mm underwent amniocentesis for chromosome study by DNA sequencing. Prevalence of aneuploidy in Yin's study reported was 3.4%.[19] They did not state the reason for amniocentesis in subjects with NT under 3.0 mm with an average age of 29 years old. The hidden reason for high-risk aneuploidy was not reported.

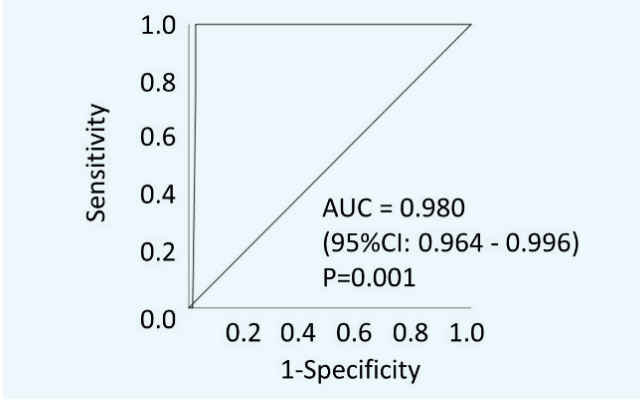


Fig 3. Correlation of nuchal translucency (NT) measurement between sagittal and transverse approach to predict positive aneuploidy screening.
Accuracy of NT by transverse approach were 97.6% with false positive rate 2.4% (7/291).

According to ACOG recommendation in 2020, pregnant women with NT greater than 3.0 mm were recommended for genetic study.[1] Turkish study by Dinc in 2021 found healthy fetus with NT greater than 3.0 and 3.5 mm at 4.1% and 2.1%, respectively.[7] The cut-off point at 3.0 mm might include 6.2% of normal healthy fetuses. The karyotype study was needed to confirm and rule out these healthy fetuses. NT measurement was widely

performed for aneuploidy screening in early pregnancy by two-dimension ultrasonography (2D US).[13,19,20] Khalifeh reported NT measurements by 2D US and three-dimension ultrasonography (3D US) were with comparable accuracy.[20] Standard cut-off point was widely accepted at 3.0 mm or more as the high-risk indicator for aneuploidy.[1] Sagittal plane approach during ultrasonography was recommended as following ISUOG guideline.[14] Sagittal plane measurement is a very time-consuming task for obtaining high accuracy of measurement. Transverse planar approach by ultrasonography was introduced as an alternative approach.[13]

From the current study, NT measurement by sagittal and transverse plane were both performed by a single operator. The mean value of NT by transverse plane was significantly higher than those obtained from sagittal plane measurement. The result of NT measurement at 3.5 mm via sagittal plane was comparable to that of 3.2 mm measurement via transverse approach.[13] There was a correlation between the sagittal and transverse line following a linear pattern. The suggestion cut-off value by transverse plane at 2.9 mm was equal to the cut- off value of 3.0 mm by sagittal plane. Previous studies reported the cut-off point ranged from 2.4 to 3.5 mm by 2D US via sagittal plane.[18,19,21,22] Montaguti reported the same result as Khalifeh using 2D US. The current study supports the findings of Montaguti and Khalifeh with different cut-off point values. The comparison of NT measurement from the previous studies were summarized and presented in Table 2. When the situation requires massive NT screening it usually faces a lack of personnel with trained expertise in ultrasonography.[13] Screening by 2D US and transverse approach was an alternative diagnostic methodology.

Strengths of the current study include data from NT measurement by sagittal and transverse plane that operated by a single experienced ultrasound operator to minimize inter-operator variability. Small sample size and low prevalence of aneuploidy affecting the robustness of study might be the limitation of the study.

In conclusion, when massive screening of aneuploidy was needed, the institution usually faced a shortage of experienced ultrasonography personnel. NT measurement by transverse approach with cut point at 2.9 mm gave similar accuracy as sagittal approach with a cut-off threshold of 3.0 mm as recommended by ACOG. In resource-limited setting, NT measurement by transverse approach might be the first step of screening tool for referring to experience MFM person. Validation of these findings in the larger and more diverse populations might be the future research plan.

CONCLUSION

NT measurement cut-off by transverse approach at 2.9 mm was comparable to 3.0 mm of sagittal approach.

TABLE 2. Comparison of NT measurement during GA 11-14 weeks between present and previous studies.

	Present	Pinnington	Montaguti	Yin	Dinç	Grossman	Khalifeh	Karki	Tomai
Year	2024	2024	2022	2022	2021	2020	2015	2013	2011
Country	Thailand	Thailand	Italy	China	Turkey	US	US	Nepal	Vietnam
Cases (n)	295	3,928	1,023	617	1,541	110	366	211	2,500
Age (years)	31.7	31.0	33	32.2	30.2		31.2	25.03	28.9
Nulliparity (%)	55.6		61.3						43.9
Prevalence (%)	2.7	0.33	0.6	3.4	6.7/4.1/2.1	0.7			1.5
Mode	2D	2D	2D	2D	2D	2D	3D	2D	2D
View	Tx	Sag	Tx	Sag	Sag	Sag	Tx	Sag	Sag
NT (mm)									
Sag	3.0	3.0	3.5	2.5	2.5/3.0/3.5	3.5	1.7	2.5	2.4
Tx	2.9		3.2				1.66		

Abbreviations: GA: gestational age, NT: nuchal translucency, 2D: two-dimensional sonography, 3D: three-dimensional sonography, Sag: sagittal view, Tx: transverse or axial view, Mode: ultrasonography mode, View: plane of study

What is already known on this topic?

Genetic abnormalities or fetal malformation should be diagnosed as early as possible. Combination of fetal nuchal translucency (NT) and maternal serum analytes between 10 and 14 weeks of GA were recommended in year 2020 by ACOG. NT is the most effective tool for detecting Down Syndrome. Thick NT (more than 3.0 mm) measured by median sagittal view was associated with fetal anomalies or cardiac anomalies. Currently NT has been measured in a sagittal plane that necessitates an experienced sonographer. The transverse approach required less skill from staff and merited investigation.

What does this study add?

When universal aneuploidy screening was needed in situations with a shortage of experienced ultra sonographers, NT measurement via transverse approach with a cut-off point at 2.9 mm gave comparable accuracy to a sagittal approach.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

The authors would like to thank all subjects who participated in this study. Special thanks to Sawanya Benchahong, M.D., Nawabhorn Orungrote, Gp.Capt., Monyada Pleankong M.D. and Wiyada Luangdansakul M.D. for expert consultation.

DECLARATION

Grants and Funding Information

This project is not funded by any external sources.

Conflict of Interests

The author declares no conflict of interest.

Author Contributions

Conceptualization and methodology, P.J., S.M., S.P. and B.S. ; Investigation, P.J. and S.M. ; Formal analysis, P.J.,K.S.,S.P. and B.S. ; Visualization and writing – original draft, P.J., S.M. and K.S. ; Writing – review and editing, P.J., S.M., K.S. and K.B. ; Supervision, S.M. ; All authors have read and agreed to the final version of the manuscript.

Institutional Review Broad Statement

This prospective study had been approved by the local Institutional Review Broad (IRB No.84/66).

REFERENCES

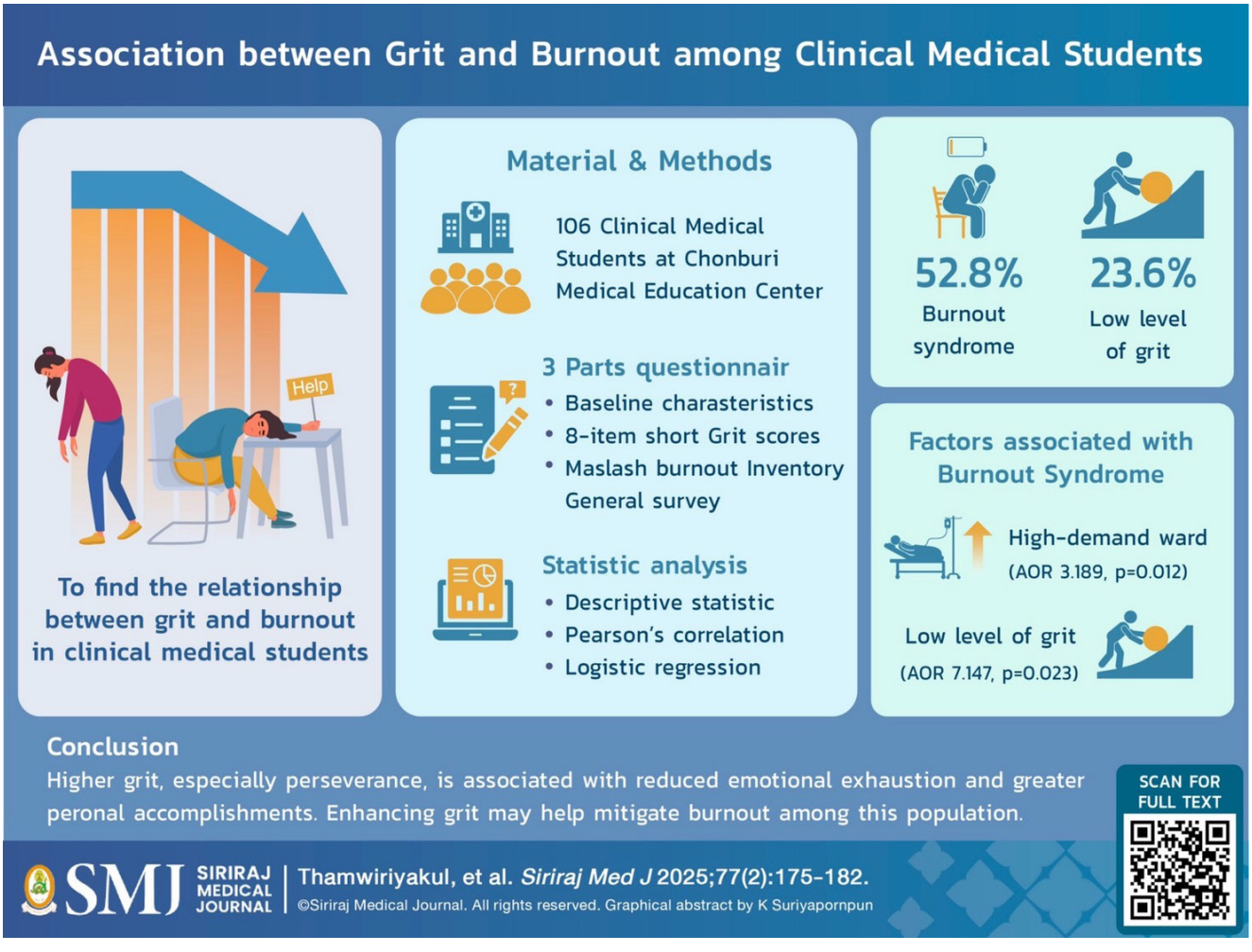
1. Screening for Fetal Chromosomal Abnormalities: ACOG Practice Bulletin, Number 226. Obstet Gynecol 2020;136:48-69.
2. Kagan KO, Avgidou K, Molina FS, Gajewska K, Nicolaides KH. Relation between increased fetal nuchal translucency thickness and chromosomal defects. Obstet Gynecol. 2006;107: 6-10.
3. Tubtimkhiao P, Orungrote N, Mungkorntongsakul S, Bhamarapravatana K, Suwannarurk K. The Correlation between Small Placental Volume Measured by Three-dimensional Ultrasonography and Adverse Pregnancy Outcomes: An Observational Study. Siriraj Med J. 2024;76(9):595-603.
4. Ounphamornlah N, Luangdansakul W, Orungrote N, Pleankong M, Petcharat K, metaneedol S, et al. Prevalence of Abnormal Cerebroplacental Ratio in Uncomplicated Full-term Pregnancy and Correlation with Adverse Perinatal Outcomes. Siriraj Med J. 2024;76(8):514-21.
5. Ruangvutilert P, Rekhawasin T, Phatihattakorn C, Boriboonhirunsarn D. Accuracy of Third Trimester Ultrasound for Predicting Large-for-Gestational Age Newborn in Women with Gestational Diabetes Mellitus. Siriraj Med J. 2021;73(5):322-9.
6. Jaisaby A, Phaliwong P, Prommart S, Smanchat B, Bhamarapravatana K, Suwannarurk K. The Accuracy of Cervical Length for Prediction of Delivery in Term Pregnancy Patients Presenting with Labor Pain. Siriraj Med J. 2019;71(4):278-83.
7. Dinç G, Eyüboğlu İ. Distribution of nuchal translucency thickness at 11 to 14 weeks of gestation in a normal Turkish population. Turk J Med Sci. 2021;51:90-4.
8. Cunningham FG, Leveno KJ. Obstetrical imaging. In: Cunningham FG, Leveno KJ, editors. Williams obstetrics. 26th ed. New York: McGraw Hill Education; 2026.p.248-50.
9. Liao Y, Wen H, Ouyang S, Yuan Y, Bi J, Guan Y, et al. Routine first-trimester ultrasound screening using a standardized anatomical protocol. Am J Obstet Gynecol. 2021;224:396.e1-15
10. Kor-Anantakul O, Suntharasaj T, Suwanrath C, Chanprapaph P, Sirichotiyakul S, Ratanasiri T, et al. Distribution of normal nuchal translucency thickness: a multicenter study in Thailand. Gynecol Obstet Invest. 2011;71(2):124-8.
11. Suntharasaj T, Ratanasiri T, Chanprapaph P, Kengpol C, Kor-anantakul O, Leetanaporn R, et al. Variability of nuchal translucency measurement: a multicenter study in Thailand. Gynecol Obstet Invest. 2005;60:201-5.
12. Montaguti E, Diglio J, Petrachi B, Arosio V, Fiorentini M, Cavalera M, et al. Identification of Fetuses at Increased Risk of Trisomies in the First Trimester Using Axial Planes. Fetal Diagn Ther. 2024;51:1-6.
13. Montaguti E, Rizzo R, Diglio J, Di Donna G, Brunelli E, Cofano M, et al. Increased nuchal translucency can be ascertained using transverse planes. Am J Obstet Gynecol. 2022;227(5):750. e1-6.
14. Bilardo CM, Chaoui R, Hyett JA, Kagan KO, Karim JN, Papageorghiou AT, et al. ISUOG Practice Guidelines (updated): performance of 11-14-week ultrasound scan. Ultrasound Obstet Gynecol. 2023;61:127-43.
15. Malinger G, Paladini D, Haratz KK, Monteagudo A, Pilu GL, Timor-Tritsch IE. ISUOG Practice Guidelines (updated): sonographic examination of the fetal central nervous system. Part 1: performance of screening examination and indications for targeted neurosonography. Ultrasound Obstet Gynecol. 2020; 56:476-84.

16. Paladini D, Malinger G, Birnbaum R, Monteagudo A, Pilu G, Salomon LJ, et al. ISUOG Practice Guidelines (updated): sonographic examination of the fetal central nervous system. Part 2: performance of targeted neurosonography. Ultrasound Obstet Gynecol. 2021;57:661-71.
17. Pinnington TR, Arthan J, Komoltri C, Chanprapaph P. Accuracy of the Combined First Trimester Down Syndrome Screening Test and the Optimum Range of the Cut-off Point for Intermediate-Risk Identification: Twelve years' Experience. Thai J Obstet Gynaecol. 2024;32:269-77.
18. Tomai XH, Schaaps JP, Foidart JM. Fetal nuchal translucency thickness in different cut-off points for aneuploidy screening in the south of Vietnam. J Obstet Gynaecol. 2011;37:1327-34.
19. Yin D, Chen L, Wang L, Zeng Y, Tang F, Wang J. Does isolated nuchal translucency from 2.5 to 2.9 mm increase the risk of fetal chromosome disease. Mol Genet Genomics. 2022;297:1643-8.
20. Khalifeh A, Weiner S, Berghella V, Scott S, Gerson A. Comparative Analysis of Two- Versus Three-Dimensional Sonography for Nuchal Translucency Measurement. Am J Perinatol. 2016;33:486-9.
21. Karki S, Joshi KS, Tamrakar SR, Regmi S, Khanal K. Nuchal translucency in normal fetus and its variation with increasing crown rump length (CRL) and gestational age. Kathmandu Univ Med J (KUMJ). 2013;11(44):282-6.
22. Grossman TB, Bodenlos KL, Chasen ST. Abnormal nuchal translucency: residual risk with normal cell-free DNA screening. J Matern Fetal Neonatal Med. 2020;33:3062-7.

Association between Grit and Burnout among Clinical Medical Students

Nirucha Thamwiriyakul, M.D.[1], Supitcha Thamissarakul, M.D.[1], Prakasit Wannapaschaiyong, M.D.[2,*]

[1]Chonburi Hospital Medical Education Center, Chonburi 20000, Thailand, [2]Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.



*Corresponding author: Prakasit Wannapaschaiyong
E-mail: prakasit.wan@mahidol.ac.th
Received 12 November 2024 Revised 22 December 2024 Accepted 22 December 2024
ORCID ID:http://orcid.org/0000-0001-7099-0183
https://doi.org/10.33192/smj.v77i2.272188

ABSTRACT

Objective: The purpose of this study was to determine the relationship between grit and burnout among clinical medical students.

Material and Methods: A cross-sectional descriptive study was conducted between December 2023 and January 2024 at the Chonburi Medical Education Center. Data from 106 clinical medical students were gathered using the 8-item Short Grit scale and the Maslach Burnout Inventory General Survey. Descriptive statistics, Pearson's correlation, and logistic regression were used to analyze the data.

Results: Among the 106 clinical medical students, 52.8% met the criteria for burnout syndrome. 44.3% had high levels of emotional exhaustion and 32.1% had high levels of depersonalization. However, most of these medical students had moderate levels of total grit, passion, and perseverance. Pearson's correlation analysis showed that total grit and perseverance scores were negatively correlated with emotional exhaustion and positively correlated with personal accomplishments. Further analysis using multivariate logistic regression revealed that burnout syndrome in clinical medical students was significantly associated with high-demand ward responsibility (AOR 3.189, p = 0.012) and low levels of total grit (AOR 7.147, p = 0.023).

Conclusion: Burnout is prevalent among clinical medical students, particularly among those with high-demand ward responsibilities and low grit levels. Higher grit, especially perseverance, is associated with reduced emotional exhaustion and greater personal accomplishments. Enhancing grit may help mitigate burnout among this population.

Keywords: Grit; burnout; clinical medical students (Siriraj Med J 2025; 77: 175-182)

INTRODUCTION

Burnout syndrome, a psychological condition resulting from chronic work-related stress, affects individuals across various professions, with medical students experiencing particularly high prevalence rates.[1,2] A global meta-analysis reported that the burnout rate among medical students was 44.2%, ranging from 33.4% to 55.0%.[3] More recent evidence suggests that these high rates persist. A systematic review by Di Vincenzo et al. (2024) reported that burnout prevalence among medical students can be as high as 88%[4], underscoring the severity and persistence of the problem. This high prevalence is especially concerning for clinical medical students who undergo practical training in hospital wards, where stressors are more intense compared to preclinical years.[5]

The clinical years of medical training, in particular, are a period of heightened vulnerability to burnout. During these years, students transition from predominantly classroom-based learning to high-pressure, patient-focused clinical environments. They must rapidly assimilate knowledge, refine their clinical skills, and navigate complex professional hierarchies, often with limited decision-making authority. Chronic exposure to demanding work schedules, challenging interpersonal interactions, high academic expectations, and insufficient rest contribute substantially to psychological distress.[6,7] Many previous studies have highlighted that sustained burnout during these formative clinical years not only undermines students' mental health and academic engagement but may also erode their clinical competencies and empathy, subsequently affecting patient care quality and their future professional satisfaction.[1,4,8]

Researchers have explored various factors that might mitigate or regulate burnout levels among medical students. One such factor is "grit", which is defined as sustained passion and perseverance for long-term goals.[9] Grit has been recognized as a critical trait for learners, predicting high academic performance and lower attrition rates in multiple disciplines.[10] Medical students with high levels of grit tend to persevere in the face of challenges, potentially counteracting burnout.[11] Grit also appears to provide mental health benefits, such as lower levels of depression and anxiety. Although multiple determinants, such as workload, interpersonal conflicts, and institutional policies,[1-4] have been implicated in the development of burnout, we focused on grit in this study because it is an intrinsic, potentially modifiable personal characteristic.[9,10] By understanding grit's role, we may identify strategies to enhance students' resilience and their capacity to cope with the inevitable stressors of clinical training.

Existing studies on the relationship between grit and burnout have notable limitations. Many studies have focused on postgraduate trainees[12-16] or general medical students[17,18], there is a marked absence of studies specifically targeting clinical medical students. This represents a

critical gap in understanding this subgroup's grit and burnout dynamics. Furthermore, no prior studies have explored this relationship among medical students in Thailand, underscoring the need to address geographic and cultural nuances in the literature.

Given these gaps, this study aims to explore the association between grit and burnout among clinical medical students at the Chonburi Medical Education Center in Thailand. By addressing these limitations, the findings of this study could inform the development of targeted interventions to support mental health and academic resilience in medical trainees.

MATERIALS AND METHODS

Study design and population

This cross-sectional study was conducted with clinical medical students at the Chonburi Medical Education Center between December 2023 and January 2024. A total of 106 clinical medical students completed an online questionnaire using Google Forms. Because all 106 enrolled clinical medical students at the time of the study agreed to participate, we obtained complete participation rather than a traditional response rate calculation. No formal sample size calculation was performed, as we included the entire population of clinical medical students as a convenience sample. This study was approved by the Chonburi Hospital Research Ethics Committee (code 145/66/S/q).

Data collection

The authors invited students to participate in various student activities. The anonymous electronic questionnaire included information regarding the purpose of the research, consent forms, and study questionnaires. The participants had the right to choose to participate in the research project or decline.

The questionnaire consisted of three parts: demographic information, the 8-item Short Grit Scale (Thai version), and the Maslach Burnout Inventory-General Survey (Thai version). Demographic information collected included sex, age, academic year, current ward responsibilities, cumulative grade point average (GPAX), and intention to pursue medicine.

The 8-item Short Grit Scale (Thai version)[19] was used to assess grit. This instrument comprises two dimensions, consistency of interest (passion) and perseverance of effort (perseverance), each consisting of four items. Participants' responses were recorded on a 5-point Likert scale ranging from 'very much like me' to 'not like me at all'. Higher scores in passion, perseverance, and total grit indicated a stronger drive to improve, persistence in goals, and long-term commitment, respectively. The scores were classified as low (below the 25th percentile), moderate (25-75th percentile), or high (above the 75th percentile). The scale exhibited acceptable internal consistency with a Cronbach's alpha coefficient of 0.69, as reported in the Thai version.

The Maslach Burnout Inventory-General Survey; MBI-GS (Thai version)[20] was utilized to assess burnout among medical students in this study. This 22-item questionnaire is designed to evaluate three dimensions of burnout: emotional exhaustion (nine items), depersonalization (five items), and personal accomplishment (eight items). Responses were measured on a 7-point frequency scale ranging from 'never' to 'every day,' with higher scores in emotional exhaustion and depersonalization, and lower scores in personal accomplishment, indicating more severe symptoms of burnout. Each dimension's results were categorized into three severity levels: low, moderate, and high. A higher severity level reflects more pronounced symptoms within that dimension. For this study, participants were considered to have burnout syndrome if they scored at the high level for emotional exhaustion and/or depersonalization.

The MBI-GS was selected for its ability to comprehensively measure burnout in contexts that involve high work-related stress. While the tool is typically applied to general populations, its structure and scope are well-suited for medical students, who often experience stressors resembling those faced by medical personnel, such as patient care responsibilities, academic pressures, high expectations from peers and supervisors.

The reliability of the MBI-GS in this study was satisfactory, with Cronbach's alpha coefficients of 0.86 for emotional exhaustion, 0.86 for depersonalization, and 0.77 for personal accomplishment, indicating good internal consistency.

Statistical analyses

Data analysis was performed using SPSS software (version 23). Descriptive statistics were used to summarize demographic data, presenting results as means with standard deviations or medians with interquartile ranges, as appropriate. The relationship between grit and burnout was examined using Pearson's correlation coefficient. A logistic regression analysis was conducted to identify factors associated with burnout. Statistical significance was set at $p < 0.05$ for all analyses.

RESULTS

A total of 106 clinical students responded to the survey, representing 100% of the sample (Table 1).

Demographic data indicated that 57.5% of the respondents were female medical students. The distribution of students was uniform for each academic year. 70.8% achieved a grade point average (GPAX) of 3.00. 86.8% reported that they had autonomously chosen to pursue medical studies. Among the participants, 67.9% were currently working in high-demand wards (obstetrics and gynecology, surgery, internal medicine, and pediatrics), wards where clinical medical students carry heavier responsibilities, collaborate closely with multiple levels of healthcare professionals, and undertake overnight on-call duties. In contrast, 32.1% were placed in low-demand wards (Eye, ENT, and Psychiatry), with fewer responsibilities, reduced workplace stress, and no overnight on-call shifts. More than half of the medical students (52.8%) met the criteria for burnout syndrome.

Table 2 shows that the average total grit score was 25.35±4.05. Most of the students (57.6%) scored in the moderate range, 23.6% in the low range, and 18.9% in the high range. When examining each domain, more than half of the clinical medical students exhibited moderate levels of passion and perseverance.

TABLE 1. Baseline characteristics of clinical medical students (N=106).

Demographic Characteristics	Descriptive Results
Gender	
Male	45(42.5)
Female	61(57.5)
Year-level of clinical medical students	
4th-year medical students	36(34.0)
5th-year medical students	35(33.0)
6th-year medical students	35(33.0)
GPAX	
<3.00	31(29.2)
≥3.00	75(70.8)
Intention to pursue medicine	
No	14(13.2)
Yes	92(86.8)
Current ward responsibilities	
High-demand wards	72(67.9)
Low-demand wards	34(32.1)
Burnout syndrome	
No	50(47.2)
Yes	56(52.8)

GPAX=cumulative grade point average
Data presented as numbers (percentages)

TABLE 2. Grit of clinical medical students (N= 106).

Grit domains	Low level	Moderate level	High level	Average score[a]
Passion	23(21.7)	58(54.7)	25(23.6)	12.51±2.69
Perseverance	16(15.1)	68(64.2)	22(20.8)	12.81±2.58
Total Grit	25(23.6)	61(57.6)	20(18.9)	25.35±4.05

Data are presented as numbers (percentages), aData are presented as mean ± standard deviation.

Table 3 presents the burnout levels of the clinical medical students at Chonburi Hospital. The results indicate high levels of negative aspects, with emotional exhaustion being the most prevalent in 44.3% of students who reported high levels. Depersonalization was also significant, with 32.1% of the students reporting high levels. On a positive note, most of the students (96.2%) reported high levels of personal accomplishment.

Table 4 illustrates the relationship between grit and burnout among clinical medical students. The total grit score showed a negative correlation with emotional exhaustion (r = -0.225, p = 0.021) and a positive correlation with personal accomplishment (r = 0.247, p = 0.011). Similarly, the perseverance domain of grit had a negative correlation with emotional exhaustion (r = -0.276, p = 0.004) and a positive correlation with personal accomplishment (r = 0.287, p = 0.003).

Table 5 presents the factors influencing academic burnout using regression analysis. The results showed that students currently working on the main patient wards were 3.2 times more likely to experience burnout than those working on the low-demand wards, with an adjusted odds ratio (AOR) of 3.189 (95% confidence interval [CI]: 1.288, 7.897; p=0.012). Additionally, medical students with low levels of total grit are 7.1 times more likely to experience burnout compared to those with moderate and high levels of perseverance, with an AOR of 7.147 (95% CI: 1.311, 38.964; p=0.023). Both findings were statistically significant.

DISCUSSION

The findings revealed a high prevalence of burnout, with 52.8% of the students meeting the criteria for burnout syndrome. This rate is consistent with previous research that showing similar levels of burnout among medical students worldwide.[3] Emotional exhaustion and depersonalization were common, affecting 44.3% and 32.1% of students, respectively. These findings align with the Thamissarakul study (2024), which indicates that many clinical medical students are likely to experience higher levels of emotional exhaustion and depersonalization due to the increased workload compared to the preclinical

TABLE 3. Burnout symptoms of clinical medical students (N=106).

Burnout symptoms	Low level	Moderate level	High level	Average score[a]
Emotional exhaustion	22(20.6)	37(34.9)	47(44.3)	25(18, 31)
Depersonalization	37(34.9)	35(33.0)	34(32.1)	9(5, 14)
Personal accomplishment	1(0.9)	3(2.8)	102(96.2)	17(13, 24)

Data presented as numbers (percentages), Data presented as median ± IQR

TABLE 4. Pearson's correlations between Grit and Burnout symptoms.

	Total Grit	PS	PV	EE	DP	PA
Total Grit	1.000	0.726**	0.770**	-0.225*	0.149	0.247*
PS	1.000	0.188	-0.104	0.147	0.087	
PV	1.000	-0.276**	-0.055	0.287**		
EE	1.000	0.597**	0.077			
DP	1.000	0.038				
PA	1.000					

Abbreviations: PS, passion; PV, perseverance; EE, emotional exhaustion; DP, depersonalization; PA, personal accomplishment
*p<0.05, **p<0.01

TABLE 5. Factors Associated with Burnout Syndrome using logistic regression analysis.

Factors	Crude odd ratio (95% CI)	P-value	Adjusted odd ratio (95% CI)	*p*-value
Female	0.966(0.446, 2.089)	0.929	0.875(0.376, 2.034)	0.756
GPAX < 3.00	1.626(0.693, 3.815)	0.264	1.435(0.565, 3.645)	0.448
Intention to pursue medicine	0.400(0.117, 1.368)	0.144	0.370(0.098, 1.404)	0.144
High-demand ward	2.881(1.234, 6.727)	0.014*	3.189(1.288, 7.897)	0.012*
Low grit	2.288(0.888, 5.897)	0.087	7.147(1.311, 38.964)	0.023*

*$p<0.05$

level, along with the pressures of ward duties, exams, and expectations from other healthcare providers and supervisors.[1] However, personal accomplishments remained high, with 96.2% reporting high levels of achievement. This highlights the complex nature of burnout, where students may feel both exhausted and detached from their work, while still experiencing a sense of fulfillment from their accomplishments.[21,22]

To place our results in a more recent context, a systematic review by Di Vincenzo et al. (2024) reported that burnout prevalence among medical students can vary widely, ranging from 5.6% to as high as 88%.[4] While our observed prevalence of 52.8% among clinical medical students falls within this broad spectrum, it still highlights that burnout remains a substantial concern. Additionally, potential methodological issues in our study—such as the single-institution context, the relatively small sample size, and possible selection bias—may have influenced our results. Future research should consider larger, multi-institutional samples to enhance the generalizability and accuracy of prevalence estimates.

One of the key findings of this study is the significant relationship between grit and burnout. Grit, especially perseverance, was negatively correlated with emotional exhaustion, indicating that students with higher levels of perseverance are better able to cope with the stresses of medical training. This is consistent with prior studies demonstrating that grit plays a protective role in high-pressure environments by fostering resilience.[10,23] Furthermore, the positive correlation between grit and personal accomplishment suggests that students with higher levels of grit not only experience less emotional exhaustion, but also feel a stronger sense of achievement in their work.[9,11] This underlines the potential role of grit in mitigating some of the negative impacts of burnout.

These findings are in line with more recent studies involving medical students. For example, Lee et al. (2023) demonstrated an inverse association between grit and burnout, indicating that students with higher grit levels tend to report lower burnout symptoms.[18] Similarly, Jumat et al. (2020) found that students with higher grit had a lower likelihood of experiencing burnout, with an odds ratio of 0.84 (95% CI: 0.74–0.96), and an area under the ROC curve of 0.76 (95% CI: 0.62–0.89) when using grit as a single predictor of burnout.[17] Although our study's context differs (e.g., geographic location, clinical environment), the converging evidence suggests a consistent protective effect of grit against burnout in medical trainees.

Furthermore, multivariate logistic regression analysis indicated that low levels of grit were strongly associated with burnout. This suggests that grit can serve as a protective factor, reducing the likelihood of burnout in medical students. This finding aligns with recent studies in medical education that highlight the importance of noncognitive traits, such as grit, in predicting well-being and success in stressful settings.[15,24] Therefore, interventions that cultivate grit in medical students could be valuable in preventing burnout.

In addition, this study also found that students assigned to high-demand ward responsibilities were significantly more likely to experience burnout compared to those with low-demand ward duties. This finding is consistent with previous studies in Thailand, where clinical rotations on high-demand wards are associated with elevated stress levels.[5,6] These results emphasize the

need for targeted interventions for students in these intensive training environments. Fostering resilience and perseverance through grit-enhancing programs may help students manage the pressures of ward duties and reduce the risk of burnout.[12,25]

Interestingly, unlike some previous reports,[1,2,8,26] we did not find significant associations between burnout and variables such as gender or year of study. While some studies suggest that these demographic factors influence burnout risk[1-4,8], our findings differ. This discrepancy may reflect variations in sample size, institutional culture, or educational policies that influence how stressors are distributed or perceived. Beyond demographic and institutional differences, broader educational contexts, such as the frequency and nature of examinations, the availability of mental health support services, the structure of advisory and mentorship systems, and the provision of scholarships or financial assistance, may all contribute to how clinical stressors are experienced.[5,14] These factors could shape students' resilience and coping mechanisms, thereby influencing the prevalence and severity of burnout. Future research should explore why these differences occur—whether they are due to local conditions, unique support structures, or other moderating factors.

While this study provides valuable insights, it has some limitations. Being cross-sectional, it does not allow for the determination of causal relationships. Future longitudinal studies are needed to assess how grit influences the development of burnout over time. Additionally, the small sample size limits the generalizability and statistical power of our findings. Expanding the sample and including multiple institutions could provide a more comprehensive understanding of the relationships between grit, burnout, and contextual factors among clinical medical students.

This study utilized the validated Thai version of the Maslach Burnout Inventory-General Survey (MBI-GS), which is based on the 1989 edition of the MBI Manual. While this version is widely accepted and has been used extensively in Thai research, we acknowledge that the MBI Manual has undergone updates, with the latest being the 4th Edition (1996–2018).[27] These updates may include refinements in scoring or interpretation that were not incorporated into the Thai version. Future studies may consider adapting or validating the most recent version of the MBI for the Thai context to align with the latest international standards.

CONCLUSION

Higher grit, particularly perseverance, is associated with lower emotional exhaustion and greater personal accomplishment, suggesting that grit serves as a protective factor against burnout among clinical medical students. Despite this potential buffering effect, burnout is still prevalent in this population, especially among students with high-demand ward responsibilities and those exhibiting low grit levels. These findings underscore the importance of interventions aimed at enhancing grit as a strategy to mitigate burnout and improve the well-being of medical trainees.

Data Availability Statement

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

ACKNOWLEDGEMENTS

We express our gratitude to Arunrat Tangmunkongvorakul for permitting us to use the 8-item Short Grit Scale (Thai version). Furthermore, this study could not have been completed without the valuable cooperation of the participants and the director of Chonburi Hospital.

DECLARATION

Grants and Funding Information

No funding was provided by any institutions or individuals for this study.

Conflict of Interest

The authors declare no conflict of interests.

Author Contributions

Conceptualization and methodology, N.T., S.T., and P.W. ; Investigation, N.T., and S.T.; Formal analysis, N.T. and S.T. ; Visualization and writing – original draft, N.T., and S.T. ; Writing – review and editing, P.W.; Supervision, P.W. All authors have read and agreed to the final version of the manuscript.

Use of Artificial Intelligence

Not applicable

REFERENCES

1. Thamissarakul S, Hongkan W, Wannapaschaiyong P. Factors Associated with Burnout Syndrome among Clinical Medical Students at Chonburi Hospital, Thailand. J Health Sci Med Res. 2024;42(3):e20231020.
2. Auchayasawat S, Priamnong SP, Praken T, Ritdej N, Yingsanongchat P, Usaneengamcharoen P. Prevalence and associated factors of burnout syndrome among clinical medical students, Ubon Ratchathani University. J Psychiatr Assoc Thai. 2022;67(1): 101-12.

3. Frajerman A, Morvan Y, Kreb M-O, Gorwood P, Chaumette B. Burnout in medical students before residency: A systemic review and meta-analysis. Eur Psychiatry. 2019;55:36-42.
4. Di Vincenzo M, Arsenio E, Della Rocca B, Rosa A, Tretola L, Torica R, et al. Is There a Burnout Epidemic among Medical Students? Results from a Systematic Review. Medicina (Kaunas). 2024;60(4):575.
5. Wannapaschaiyong P, Kallawicha K. Stress and stress coping strategies among foreign bachelor's medical interns in Bangkok Metropolitan Administration (BMA) hospitals. J Health Res. 2023;37:306-16.
6. Apiwatanasiri C, Somaketrarin K, Suraprayoon K, Leumprasert K, Wankaew N, Homchampa P, et.al. Stress and coping in medical students at clinical level, KhonKaen University. Srinagarind Med J. 2007;22(4):416-24.
7. Rayakeaw R, Satayasai W. 4-6th year medical students' stress: Case study at medical centre, Buddhachinaraj Hospital, Phitsanulok. Thammasat Medical Journal. 2013;13(1):17-23.
8. Ishak W, Nikravesh R, Lederer S, Perry R, Ogunyemi D, Bernstein C. Burnout in medical students: a systematic review. Clin Teach. 2013;10(4):242-5.
9. Duckworth AL, Peterson C, Matthews MD, Kelly DR. Grit: perseverance and passion for long-term goals. J Pers Soc Psychol. 2007;92(6):1087-101.
10. Eskreis – Winkler L, Shulman EP, Beal SA, Duckworth AL. The grit effect: predicting retention in the military, the workplace, school, and marriage. Front Psychol.2014;5:36.
11. Kim J, Jang H. Psychological strengths predict college students' academic grit to reduce burnout. JSS. 2020;31(4):27-44.
12. Halliday L, Walker A, Vig S, Hines J, JJPmj B. Grit and Burnout in UK doctors: a cross-sectional study across specialties and stages of training. Postgrad Med J. 2017;93(1101):389-94.
13. Bailey J, Thomas C, McDowall A. Grit and burnout in UK emergency medicine trainees. BMJ Leader. 2018;2:A38.
14. Hamdan M, Haddad BI, Alshrouf MA, Al-Ani A, Alisi MS, Hammed Y, et al. Burnout, grit, and resilience among Jordanian orthopaedic surgeons: a cross-sectional study. BMC Med Educ. 2023;23:593.
15. Salles A, Cohen GL, Mueller CM. The relationship between grit and resident well-being. Am J Surg. 2014;207:251-4.
16. Lee DH, Reasoner K, Lee D, Davidson C, Pennings JS, Blazar PE, et al. Is Grit Associated with Burnout and Well-being in Orthopaedic Resident and Faculty Physicians? A Multi-institution Longitudinal Study Across Training Levels. Clin Orthop Relat Res. 2021;479:2576-86.
17. Jumat MR, Chow PKH, Allen JC, Lai SH, Hwang N-C, Iqbal J, et al. Grit protects medical students from burnout: a longitudinal study. BMC Med Educ. 2020;20:266.
18. Lee D, Reasoner K, Davidson C, Pennings JS, Lee DH. The Relationships Between Grit, Burnout, and Demographic Characteristics in Medical Students. Psychol Rep. 2023;126(5): 2511-29.
19. Tangmunkongvorakul A, Kelly M, Thongpibul K, Musumari PM, Srithanaviboonchai K, Banwell C. The association between connectedness and grit among Thai in-school adolescents in urban Chiang Mai, Thailand. Front Psychol. 2022;13:809508.
20. Summawart S. Burnout among the staff nurses in Ramathibodi Hospital [Master's Thesis]: Mahidol University; 1989.
21. Liselotte D, Tait S. A narrative review on burnout experienced by medical students and residents. Med Educ. 2016;50(1):132-49.
22. Auchayasawat S, Parichart SP, Praken T, Ritdejch N, Yingsanongchat P, Usaneengamcharoen P. Prevalence and Associated Factors of Burnout Syndrome among Clinical Medical Students, Ubon Ratchathani Univdersity. J Psychiatr Assoc Thailand. 2022;67(1): 101-12.
23. Bakker AB, Demerouti E. The Job Demands-Resources model: state of the art. Journal of Managerial Psychology. 2007;22(3): 309-28.
24. Harpaz G, Vaizman T, Yaffe Y. University students' academic grit and academic achievements predicted by subjective well-being, coping resources, and self-cultivation characteristics. Higher Educ Q. 2024;78(1):192-211.
25. Donald H, Lee MD, Kaitlyn Keasoner MD, Diane Lee BS, Claudia Davidson MPH, Jacquelyn S, et al. Is Grit Associated with Burnout and well-being in Orthopaedic Resident and Faculty Physicians? A Multi-institution longitudinal study Across Training Levels. Clin Orthop Relat Res. 2021;479: 2576-86.
26. Soonthornkes N, Sakornwattananon O, Jirapitakkul J, Nuallaong W. Burnout and associated factors among Thai anesthesiology residents. Siriraj Med J. 2023;75(10):736-43.
27. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory™ Manual. 4th ed. Mind Garden, Inc; 2018.