E-ISSN 2228-8082

Volume 75, Number 3, March 2023

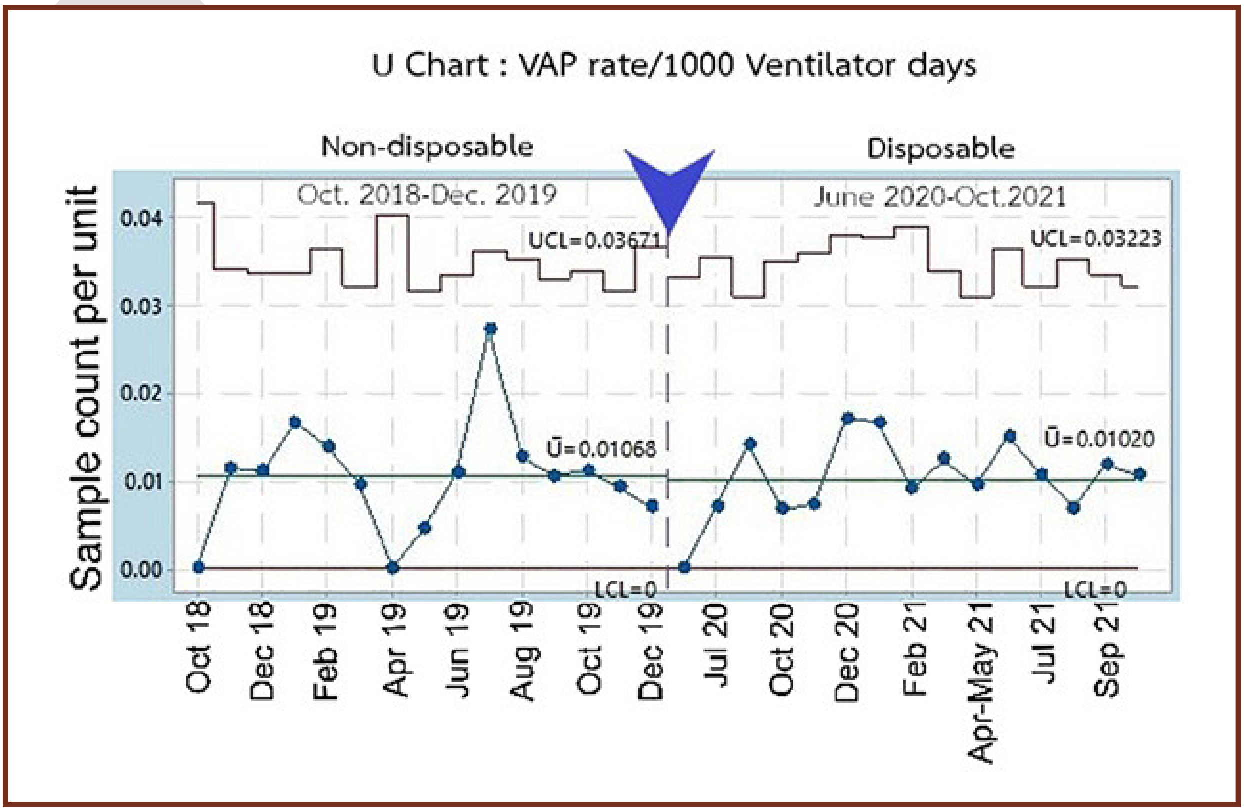
SMJ

Siriraj Medical Journal

The world-leading biomedical science of Thailand

MONTHLY

ORIGINAL ARTICLE



By Jarunee Songmuang, et al.

Indexed by Scopus®

PRINCE MAHIDOL AWARD CONFERENCE 2023

SETTING A NEW HEALTH AGENDA At the Nexus of Climate Change, Environment, and Biodiversity

Thai Association for Gastrointestinal

AIOSG Thailand

COLLEGE INTERNATIONAL DE CHIRURGIENS THAILAND SECTION 1954

MESDA Mekong Endo-Surgery Development Association

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

E-mail: sijournal92@gmail.com

Siriraj Medical Journal SMJ

Volume 75, Number 3, March 2023

ORIGINAL ARTICLE

SIRIRAJ MEDICAL JOURNAL

www.smj.si.mahidol.ac.th

SIRIRAJ MEDICAL JOURNAL is published bimonthly, 6 issues a year (Jan-Feb, Mar-Apr, May-Jun, Jul-Aug, Sep-Oct and Nov-Dec) and distributed by the end of the last month of that issue.
SIRIRAJ MEDICAL JOURNAL is listed as a journal following the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (URM) by the International Committee of Medical Journal Editors (ICMJE) since 9 July 2010 [http://www.icmje.org/journals.html].

An Assessment of the Validity and Reliability of the Social-Media Addiction Screening Scale (S-MASS)

Sirinda Chanpen, M.Sc.*, Chanvit Pornnoppadol, M.D.*, Asara Vasupanrajit, M.Sc., Quankamon Dejatiwongse Na Ayudhya, M.D., Ed.M.***

**Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand, **Department of Psychiatry, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand.*

ABSTRACT

Objective: The excessive use of social media can lead to addiction among many vulnerable individuals. Hence, the utilization of a valid and reliable screening test to assess social media addiction is warranted.

Materials and Methods: The Social-Media Addiction Screening Scale (S-MASS) is a newly developed, self-report screening scale containing 16 items that assess the three main components of behavioral addiction: giving priority, impaired control, and negative consequences. The S-MASS reliability was measured using Cronbach's alpha. An exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) were employed to assess the S-MASS factorial validity. A latent profile analysis (LPA) was also carried out to identify the classes of problematic social media users.

Results: In all, 5,068 participants aged ≥ 13 years were recruited from five high schools and an online survey. Cronbach's alpha for the S-MASS was 0.90 (95% CI: 0.89–0.90), indicating excellent test reliability. The EFA and CFA revealed a good factorial validity for the S-MASS. Based on the LPA, the participants were classed as "low-risk" (n = 1,227; 24.2%), "moderate-risk" (n = 2,757; 54.4%), and "high-risk" (n = 1,084; 21.4%) problematic social media users. The key differences between these classes were gender, age, necessity to use social media for work, self-perception of addiction, and time spent on social media.

Conclusion: The S-MASS is a valid and reliable screening scale for social media addiction. The criterion validity of the S-MASS should be evaluated once formal diagnostic criteria for social media addiction become available.

Keywords: Assessment; social media; addiction; screening; test (Siriraj Med J 2023; 75: 167-180)

INTRODUCTION

In recent years, there has been a dramatic increase in the use of social media, defined as forms of electronic communication through which users create online communities to share various types of contents. Using social media has become an essential part of the daily routines of many people. On average, people spend 2 hours 24 minutes daily on social networks, excluding other internet usage.[1] People are spending increasing amount of time gaining the benefits from the platforms while being exposed to the risks they bring.

Growing evidence suggests problematic social media use or social media addiction (SMA) can cause negative effects to vulnerable individuals similarly to those found in other behavioral addictions.[2-4] SMA is associated with several mental health conditions, including attention deficit hyperactivity disorder (ADHD), obsessive compulsive disorder (OCD), depression, and anxiety.[5-7] Hence, early SMA detection and intervention in people with SMA-especially those with comorbid psychiatric disorders-would yield improved patient outcomes.

To date, social media addiction has not yet been

Corresponding author: Chanvit Pornnoppadol
E-mail: chanvit.por@mahidol.edu
Received 22 September 2022 Revised 5 December 2022 Accepted 13 December 2022
ORCID ID:http://orcid.org/0000-0002-5181-9679
https://doi.org/10.33192/smj.v75i3.261044

established as an official clinical diagnosis in DSM guidelines or ICD-11. Neither its clinical definition has arrived at a consensus.[2] The term is also used widely in non-clinical contexts.[8] However, its operational definition in clinical study is often derived from Griffith's six core components of behavioral addiction: salience, mood modification, tolerance, withdrawal symptoms, conflict, and relapse.[2,5]

ICD-11 also describes three key features of disorders due to addictive behaviors as 1) *impaired control*, a persistent pattern of repetitive behavior in which the individual exhibits impaired control over the behavior; 2) *increasing priority*, given to the behavior to the extent that it takes precedence over other life interests and daily activities; and 3) *negative consequences*, continuation, or escalation of the behavior despite negative consequences.[9] This research operationalized social media addiction using ICD-11 essential features of disorder due to addictive behaviors for the benefit of clinical relevancy.

The integration of factors is the likely explanation of social media addiction: dispositional, sociocultural and reinforcing behavior factors.[2] Neurological and personality factors are example of dispositional explanation. From neurological point of view, both chemical and behavioral addiction is explained through brain's reward systems.[2,10] While personality factors often refer to the big-five personalities and their correlations with social media addiction.[2,11] Sociocultural perspectives posit that certain family dynamics such as parental pressure influences SMA.[2,12] However, more research is yet to be done. Lastly, SMA is also explained through some learning theories such as operant conditioning, classical conditioning, and social learning.[2,12,13] For instance, positive outcomes from using social media, namely, entertainment and attention, are positive reinforcements that influence the same behavior (using social media) to be more likely to repeat.

Many previously validated SMA measurement tools (except for the Bergen Social Media Addiction Scale) were developed using small, homogeneous, and narrow age-range samples, which mostly comprised adolescents or young adults.[5,14-24,25-28,29-33] Moreover, approximately half of the tools were specifically designed for Facebook addiction, rather than SMA generally, and many lack a comprehensive factor-structure assessment (Table 1). In addition, only a few used standard statistical analyses to identify appropriate cut-off scores to differentiate problematic from normal social media use. Therefore, there is the need for a measurement tool that is applicable to social media use in general, which has been validated with a larger sample size and more-standardized analytical methods.

The aim of this study was to develop a new screening scale for the assessment of SMA-namely, the Social-Media Addiction Screening Scale (S-MASS). The authors based the development of the S-MASS on the operational definition of behavioral addiction by ICD-11, for the purpose of clinical diagnosis relevancy. The authors set out to comprehensively explore the reliability and validity of the S-MASS, based on a large-scale and heterogeneous sample with a wide age range, and to find an appropriate cut-off point to identify high-risk problematic social media users. Once the psychometric properties of this new screening instrument are established, S-MASS will be another useful tool for epidemiological and clinical studies of SMA.

MATERIALS AND METHODS

Participants

Participants were included if they were at least 13 years old, which is the minimum age required to register for most social networking sites. Participants needed to have used social media for at least 3 months preceding the study. Participants were randomly recruited from 2 sources: 1) five high schools in Bangkok; and 2) an online survey posted on the Facebook fan page of the Division of Child and Adolescent Psychiatry, Department of Psychiatry, Faculty of Medicine Siriraj Hospital, Mahidol University. Over a six-month study period, 5,437 participants were recruited. 369 participants were excluded due to incomplete S-MASS data.

Measures and Procedure

The S-MASS is a newly developed, 16-item, self-report questionnaire to assess the severity of SMA. The initial item pool was generated by the principal investigator. The final items were selected and reduced by the consensus of all investigators. The development of the S-MASS was theoretically based on three key features of ICD-11 behavioral addictions. Items in each domain were derived from the 9 criteria for Internet Gaming Disorder (IGD) that are outlined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5)[34], including preoccupation, withdrawal, tolerance, unsuccessful attempts to control, loss of interests, continued excessive use, deception, escape, and jeopardized function. The jeopardized function was extended beyond previously validated tools by also asking about disturbed functions in areas of life other than relationships. An item having to do with deceitful behavior was omitted from the final pool of items due to it having the least sensitivity among all criteria.[35]

A content validation process was then performed. All 16 S-MASS items were examined for their relevance,

TABLE 1. Comparison among measurements developed to assess social media addiction.

Year	Name of measurement	Country	No. of items	No. of participants	Participants	Age, years mean ± SD (range)	Cronbach's alpha coefficient
2010	Addictive Tendencies Scale (ATS)[14]	Australia	3	201	College students	19.1 ± 1.9 (17–24)	0.76
2011	Facebook Intrusion Questionnaire (FIQ)[15]	Australia	8	342	Undergraduate students	19.8 ± 1.8 (18–25)	0.85
2012	Bergen Facebook Addiction Scale (BFAS)[16]	Norway	6 18 (original)	423	College students	22.0 ± 4.0 (N/A)	0.83
2012	Social Networking Website Addiction Scale (SNWAS)[17]	USA	5	194	College students	–	0.86
2013	Facebook Dependence Questionnaire (FDQ)[18]	Peru	8	418	College students	20.1 ± 2.5 (N/A)	0.67
2013	Facebook Addiction Scale (FAS)[19]	Turkey	8	447	College students	21.6 ± 1.9 (18–30)	0.86
2013	Addictive Tendencies Toward Social Networking Sites[20]	China	20	316	Adults	26.6 ± 4.4 (18–40)	0.92
2015	Arabic Social Media Addiction Scale (SMAS)[21]	Kuwait	14	1,327	Undergraduate students	21.9 ± N/A (18–31)	0.61–0.75
2015	Bergen Facebook Addiction Scale – Thai Version (Thai-BFAS)[22]	Thailand	6	874	High school students (10^{th}–12^{th} grade)	16.7 ± 1.0 (N/A)	0.91
2016	Facebook Addiction Test (F-AT)[23]	Germany and Austria	7 (short) 20 (long)	1,019	Online survey	27.5 ± 9.1 (N/A)	0.92 (long)
2016	Social Media Disorder (SMD) Scale[24]	Netherlands 27 (long)	9 (short)	2,198	Online survey	14.05 ± 2.1 (10–17) 14.36 ± 2.2 (10–17)	0.76 (short) 0.90–0.92 (long)
2016	Bergen Social Media Addiction Scale (BSMAS)[5]	Norway Hungary	6 6	23,533 5,961	Online survey High school students (9^{th}–10^{th} grade)	35.8 ± 13.3 (16–88) 16.62 ± 0.96 (15–22)	0.88 0.85

TABLE 1. Comparison among measurements developed to assess social media addiction. (Continued)

Year	Name of measurement	Country	No. of items	No. of participants	Participants	Age, years mean ± SD (range)	Cronbach's alpha coefficient
2017	Bergen Social Media Addiction Scale – Italian version[25]	Italy	6	734	High school and college students	21.6 ± 3.9 (16–40)	0.88
2017	Bergen Social Media Addiction Scale – Persian version[26]	Iran	6	2,676	High school students	15.5 ± 1.2 (14–19)	0.86
2018	Chinese Social Media Addiction Scale (Liu & Ma, 2018)[27]	China	58 28	619	College students	20.4 ± 1.5 (18–25)	0.94
2018	Turkish Adaptation of the Social Media Disorder Scale in Adolescents[28]	Turkey	9	553	Adolescents	N/A (14–18)	0.83–0.86
2018	Cross-Sectional and Longitudinal Evaluation of the Social Network Use Disorder and Internet Gaming Disorder Criteria[29]	German	9	Study 1: 192 Study 2: 2,316	Phone interview and online survey Adults	Study 1 Female: 22 (21–24) Study 1 Male: 23 (21–27) Study 2 Female: 32 (25–27) Study 2 Male: 37 (27–53)	0.690–0.774
2019	Psychometric Testing of Three Chinese Online-Related Addictive Behavior Instruments among Hong Kong University Students[30]	China	6 (BSMAS)	307	University students	21.64 ± 8.11 (17–30)	0.819
2019	Cross-cultural validation of the Social Media Disorder scale[31]	China	9	903	College students	20.56 ± 2.75 (N/A)	0.753
2019	Spanish version of the Facebook Intrusion Questionnaire (FIQ-S)[32]	Spain	8	567	Adults	29.09 ± 12.03 (18–67)	0.9
2020	Social Networking Addiction Scale (SNAS)[33]	India	24	Study 1: 525 Study 3: 334	N/A	20.33 ± 1.70 (17-25)	N/A

Abbreviations: N/A, not available; SD, standard deviation clarity, and simplicity by a total of eight experts consisting of two general psychiatrists, three child and adolescent psychiatrists, two clinical psychologists, and an expert in social media communication. Each expert rated each item using one-to-four Likert scale on three measures: relevance, clarity, and simplicity. Rating score of three or four by an expert was counted as one score, resulting in eight score points at most and zero point at least for each item. The score was converted into the item content validity index or I-CVI for each item (Max = 1.00, Min = 0.00). The content validity index for the whole scale was an average of all I-CVIs (Max = 1.00, Min = 0.00).

Focus group interviews were conducted with a group of ten high school students ages from fifteen to eighteen years old to ensure clarity of understanding by those completing the scale. Each student was given a printed copy of the scale. They were provided five minutes to go through the whole paper. Then, the investigators asked how students interpreted each item, one by one. The investigators took notes and used comments to adjust language to ensure each question item convey the intended meanings.

All items are scored on a four-point Likert scale ranging from 0 (“definitely not true”) to 3 (“definitely true”). The total S-MASS score, obtained by summing the participants’ responses, ranges from 0-48. The higher the S-MASS score is, the greater the likelihood that a participant is addicted to social media.

Statistical analysis

Data were analyzed using PASW Statistics for Windows (version 18; SPSS Inc., Chicago, Ill., USA). Cronbach’s alpha was used to measure the internal consistency reliability. A Cronbach’s alpha coefficient between 0.7 and 0.9 indicated acceptable reliability.[36] Content validity was measured using a content validity index (CVI) generated by an expert committee.

The sample of 5,068 subjects was randomly divided into two using the SPSS to produce two groups: one for an exploratory factor analysis (EFA; n = 2,534), and the other for a confirmatory factor analysis (CFA; n = 2,534). The EFA was performed using SPSS.

To test the three-factor model of the S-MASS, the CFA was performed with LISREL for Windows (version 9.10; Scientific Software International Inc., Skokie, Ill., USA). The model included three factors reflecting giving priority, impaired control, and negative consequences. The model’s goodness-of-fit was assessed with the following indices: chi-square test of model fit, standardized root means square residual (SRMR), root mean square error of approximation (RMSEA), and comparative fit index (CFI). To indicate a good fit of the model, the CFI values had to be > 0.95, while the values for RMSEA and SRMR were required to be < .06 and < .08, respectively.[37]

Participants were classified into groups based on their S-MASS scores. The classification model was derived from latent profile analysis (LPA) or a Gaussian finite mixture model by expectation-maximization (EM) algorithm. *The R package mclust (R Foundation, Vienna, Austria) was also used in the classification process*. Bayesian Information Criterion (BIC), Bootstrap Likelihood Ratio Test (BLRT), Log- likelihood, and Integrated Complete-data Likelihood (ICL) were conducted as means to accurately reveal number of latent classes. A better model fit is reflected in lower BIC and AIC values, and higher log likelihood values. The BLRT was conducted to compare models’ fitness. Models with significant changes in −2 log-likelihood implied greater fitness to the data.

A Kruskal–Wallis test, followed by a Dunn–Bonferroni post-hoc pairwise comparison, were used to compare the S-MASS scores among classes. Pearson’s chi-squared test was used to compare the categorical variables of the classes. Spearman’s rank correlation coefficient was also employed to measure the strengths of association between the average daily time spent on social media and the S-MASS scores. All tests of significance were two tailed, and a *p*-value < 0.05 was considered statistically significant.

Ethics

The study was approved by the Institutional Review Board of the Faculty of Medicine Siriraj Hospital (COA no. Si 701/2013). The Declaration of Helsinki was obliged by investigators throughout the study. Informed consents were obtained from all subjects prior to participating in the study. Parental consents were obtained in case of minorities under 18 years old.

RESULTS

Demographic statistics

Table 2 details the sociodemographic characteristics of the participants, the majority of whom were adolescents and young adults. Participants from each age group show different social media use patterns such as estimated hours of use per day ($p<0.05$). Younger participants had higher S-MASS scores than older participants ($p < 0.001$). Two-thirds of participants (66.8%) reported that it was necessary for them to use social media for their work or study. Interestingly, approximately 30% of participants stated that they perceived themselves to be addicted to social media. The five most popular social media platforms used were Facebook (87.9%),

TABLE 2. Sociodemographic characteristics of participants and Social Media Addiction Screening Scale (S-MASS) score.

Characteristics	n (%) mean ± SD	S-MASS score	*P*-value
Source of recruitment			
Total	5068 (100.0)	22.58 ± 9.67	
High schools	3672 (72.5)	23.17 ± 9.54	< 0.001
Online surveys	1396 (27.5)	21.05 ± 9.86	
Gender			
Male	2058 (41.0)	21.52 ± 9.30	< 0.001
Female	2963 (59.0)	23.28 ± 9.84	
Age [mean ± SD (range) = 19.93 ± 9.96 (13–75) yrs]			
13–17 yrs	3406 (69.8)	23.21 ± 9.50[a]	< 0.001
18–25 yrs	525 (10.8)	22.94 ± 9.60[a]	
26–45 yrs	760 (15.6)	20.78 ± 9.85[b]	
46 yrs and above	186 (3.8)	16.26 ± 8.39[c]	
Necessity to use social media for work			
Necessary	3029 (66.8)	22.95 ± 9.73	< 0.001
Not necessary	1503 (33.2)	21.62 ± 9.59	
Self-perception of addiction			
Not addicted	1092 (22.8)	14.65 ± 7.49[a]	< 0.001
Probably addicted	2245 (46.8)	21.90 ± 7.56[b]	
Addicted	1462 (30.5)	29.70 ± 8.68[c]	
Amount of time spent on social media (not including gaming)			
Weekdays [mean ± SD (range) = 3.23 ± 2.63 (0.08–18) hrs/day]			
Light use (< 2 hrs/day)	1148 (26.7)	17.87 ± 8.65[a]	< 0.001
Moderate use (2–4 hrs/day)	2264 (52.6)	22.92 ± 8.79[b]	
Heavy use (> 4 hrs/day)	892 (20.7)	27.80 ± 9.68[c]	
Weekends [mean ± SD (range) = 4.91 ± 4.12 (0.08–24) hrs/day]			
Light use (< 3 hrs/day)	1367 (31.7)	17.63 ± 8.66[a]	< 0.001
Moderate use (3–5 hrs/day)	1692 (39.2)	22.68 ± 8.38[b]	
Heavy use (> 5 hrs/day)	1255 (29.1)	27.65 ± 9.30[c]	
Average daily time spent [mean ± SD (range) = 3.68 ± 2.77 (0.08–19.71) hrs/day]			
Light use (< 3 hrs/day)	2126 (50.6)	19.38 ± 8.66[a]	< 0.001
Moderate use (3–5 hrs/day)	1203 (28.6)	24.06 ± 8.68[b]	
Heavy use (> 5 hrs/day)	876 (20.8)	28.31 ± 9.50[c]	

Note: Different superscript letters (a, b, c) in the same column reflect a significant (*p*-value < 0.05) difference between the means, while the same letter in one column reflects a non-significant difference between the means.

Line (79.2%), YouTube (77.0%), Google+ (38.7%), and Instagram (37.6%).

Correlation with time spent on social media

A moderately positive correlation was observed between average daily time spent on social media and S-MASS scores (*rs*[38] = 0.412; $p < 0.001$). The heavy-use group (using social media for > 4 hours/day on weekdays, or for > 5 hours/day on weekends) had higher S-MASS scores than the light- and moderate-use groups ($p < 0.001$; Table 2).

Content validity

The calculated content validity index or CVI for the S-MASS relative to the relevance, clarity, and simplicity of all items were 0.992, 0.938, and 0.977, respectively. For items with item content validity index (I-CVI) less than 1.00, experts gave specific comments by identifying words which might be unclear to interpretation and suggested alternative words. Some experts also rewrote or rearranged questions as examples. Others pointed out double-barreled questions.

Reliability analysis

The S-MASS had excellent internal consistency, with a Cronbach's alpha coefficient of 0.90 (95% CI: 0.89–0.90). In an item analysis, all 16 S-MASS items had a corrected item-total correlation above 0.3, which confirmed that each of the 16 items was correlated with the overall scale. If deleted, no item had a Cronbach's alpha greater than 0.90; this suggested that no item disproportionately affected the overall reliability.

EFA

Prior to performing the EFA, the suitability of the data for the factor analysis was evaluated. Five factorability assessment criteria were applied. First, the correlation matrix revealed the presence of several (82.03%) coefficients of 0.30 or above, suggesting appropriate factorability. Second, the Kaiser–Meyer–Olkin measure of sampling adequacy was 0.943, which exceeded the recommended value of 0.6.[39] Third, Bartlett's test of sphericity achieved statistical significance (χ^2[40] = 29,879.585; $p < 0.001$), supporting the factorability of the correlation matrix. Fourth, the diagonals of the anti-image correlation matrix were over 0.5, suggesting the inclusion of each item in the factor analysis. Finally, each item shares common variances with other items demonstrated by the communalities above 0.3 (Table 3). Given all of the above criteria were satisfied, all items of the S-MASS were included in the EFA.

A principal components analysis (PCA) revealed three components that the authors named "giving priority", "negative consequences", and "impaired control"; they had eigenvalues exceeding 1 and explained 39.53%, 7.38%, and 6.41% of the variance, respectively (Table 3). The promax rotation method with Kaiser normalization was chosen because the factor correlation matrix for all three factors exceeded 0.32, which indicated that the factors in the analysis were correlated. All 16 items had factor loadings greater than 0.4. In addition, the Cronbach's alphas for giving priority, negative consequences, and impaired control were 0.789, 0.770, and 0.803, respectively, indicating internal consistency of these three components (Table 3).

CFA

The three-factor model with the sixteen components as indicator variables was tested with a CFA. The analysis provided an acceptable fit to the data (χ^2 = 120.77, df = 62, $p < 0.001$; CFI = 0.99; RMSEA = 0.01; SRMR = 0.01). Factor loadings ranged from 0.32 to 0.74 (Table 4 and Fig 1). Item 9 and item 13 of the S-MASS had factor loadings less than 0.4 (0.37 and 0.32, respectively).

LPA

The LPA was performed on the sixteen items of the S-MASS, and according to the Bayesian Information Criterion (BIC) and integrated complete-data likelihood (ICL) criterion, the three-class solution was selected as the best-fitting model. The features of the three classes are presented in Fig 2. The three classes were named "low-risk" (S-MASS scores 0–15), "moderate-risk" (S-MASS scores 16–30), and "high-risk" (S-MASS scores 31–48), representing 24.2% (n = 1,227), 54.4% (n = 2,757), and 21.4% (n = 1,084) of social media users, respectively.

In the "high-risk" class, item 11— "I ignore or fail when people tell me to cut down my social media use"— showed elevated levels compared to the other items, as shown in Fig 2. Interestingly, 67.6% of the participants in the high-risk class believed that they were addicted to social media, whereas only 26.2% and 6.9% in the moderate-risk and low-risk classes, respectively, believed so (Table 5).

DISCUSSION

The purpose of the present study was to measure the psychometric properties of the S-MASS, a newly developed, 16-item, self-report questionnaire developed to screen for SMA. Results showed that the S-MASS has excellent internal consistency (Cronbach's alpha coefficient, α = 0.90). This indicated that all 16 items

TABLE 3. Exploratory factor analysis (EFA) of Social Media Addiction Screening Scale (S-MASS).

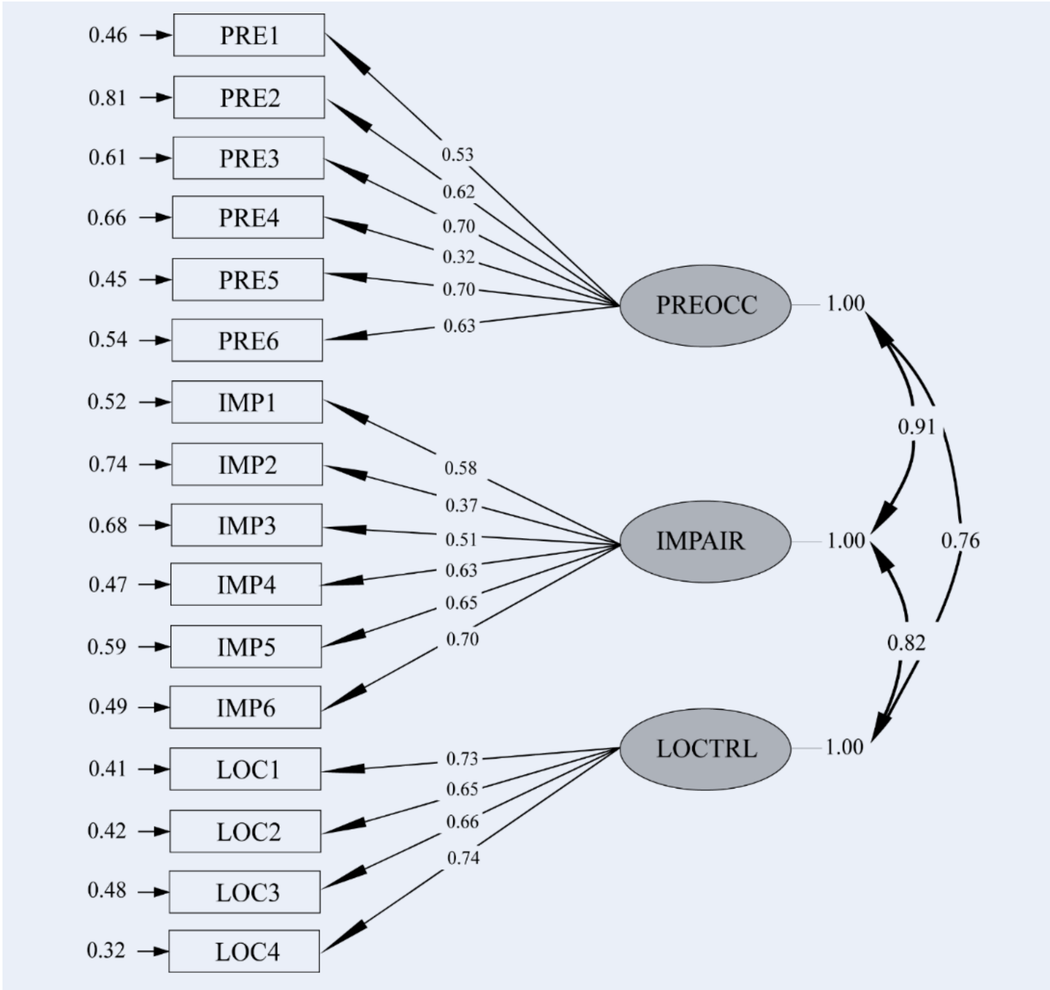
Questions	Loadings Factor 1: Giving priority	Factor 2: Negative consequences	Factor 3: Impaired control	Communality
Since I started using social media…				
(15) My friends regularly see me online.	0.730			0.572
(14) I feel that social media is a part of my life that I can't lose.	0.686			0.564
(1) I use social media whenever I have a chance.	0.677			0.506
(7) I keep checking all the time to see if anyone has “liked” or commented on the pictures/ statuses I have posted.	0.613			0.435
(2) I use social media as soon as I wake up in the morning.	0.603			0.421
(13) I use social media to ease my stress.	0.596			0.393
(11) I ignore or fail when people tell me to cut down my social media use.		0.763		0.587
(16) People around me say I am addicted to social media.		0.705		0.580
(10) I use social media during circumstances when I should not use it (e.g., while in the classroom, doing daily activities, working, meeting with friends or colleagues, walking on the sidewalk, driving, etc.).		0.688		0.490
(12) I get agitated or irritable when I can’t use social media.		0.670		0.505
(8) I talk to people on social media more often than in real life.		0.645		0.420
(9) My social media use negatively impacts my life in some ways.		0.626		0.563
(5) I often spend more time using social media than I originally intended to.			0.799	0.647
(3) I spend all of my free time using social media.			0.791	0.653
(4) I have lost interest in other activities.			0.762	0.602
(6) I spend more time using social media now than I used to.			0.732	0.593
Eigen value	6.324	1.181	1.026	
Percentage of variance explained (total = 53.322)	39.528	7.384	6.410	
Cronbach’s alpha	0.789	0.770	0.803	

Extraction method: principal component analysis
Rotation method: Promax with Kaiser normalization
Factor loadings < 0.4 were suppressed

TABLE 4. Factor loadings, R2, and factor loading coefficient of Social Media Addiction Screening Scale (S-MASS).

Components	Factor loadings				R^2	Factor loading coefficient
	b	B	SE (b)	t		
Giving priority						
(1) I use social media whenever I have a chance. (PRE1)	0.53	0.61	0.02	31.44*	0.37	0.17
(2) I use social media as soon as I wake up in the morning. (PRE2)	0.62	0.57	0.02	28.93*	0.32	0.08
(7) I keep checking all the time to see if anyone has "liked" or commented on the pictures/statuses I have posted. (PRE3)	0.70	0.67	0.02	34.56*	0.45	0.20
(13) I use social media to ease my stress. (PRE4)	0.32	0.36	0.02	17.13*	0.13	0.03
(14) I feel that social media is a part of my life that I can't lose. (PRE5)	0.70	0.72	0.02	37.78*	0.52	0.25
(15) My friends regularly see me online. (PRE6)	0.63	0.65	0.02	34.51*	0.42	0.15
Negative consequences						
(8) I talk to people on social media more often than in real life. (IMP1)	0.58	0.63	0.02	31.62*	0.40	0.22
(9) My social media use negatively impacts my life in some ways. (IMP2)	0.37	0.40	0.02	17.94*	0.15	0.04
(10) I use social media during circumstances when I should not use it (e.g., while in the classroom, doing daily activities, working, meeting with friends or colleagues, walking on the sidewalk, driving, etc.). (IMP3)	0.51	0.53	0.02	26.00*	0.28	0.08
(11) I ignore or fail when people tell me to cut down my social media use. (IMP4)	0.63	0.68	0.02	33.31*	0.46	0.17
(12) I get agitated or irritable when I can't use social media. (IMP5)	0.65	0.65	0.02	33.17*	0.42	0.14
(16) People around me say I am addicted to social media. (IMP6)	0.70	0.71	0.02	37.86*	0.50	0.17
Impaired control						
(3) I spend all of my free time using social media. (LOC1)	0.73	0.75	0.02	40.51*	0.57	0.31
(4) I have lost interest in other activities. (LOC2)	0.65	0.71	0.02	37.58*	0.50	0.25
(5) I often spend more time using social media than I originally intended to. (LOC3)	0.66	0.69	0.02	37.23*	0.48	0.18
(6) I spend more time using social media now than I used to. (LOC4)	0.74	0.79	0.02	40.82*	0.63	0.38

*p-value < 0.01



Chi-Square = 120.77, df = 62, P-value = 0.0001, RMSEA = 0.019

Fig 1. Factor loadings of Social-Media Addiction Screening Scale (S-MASS)

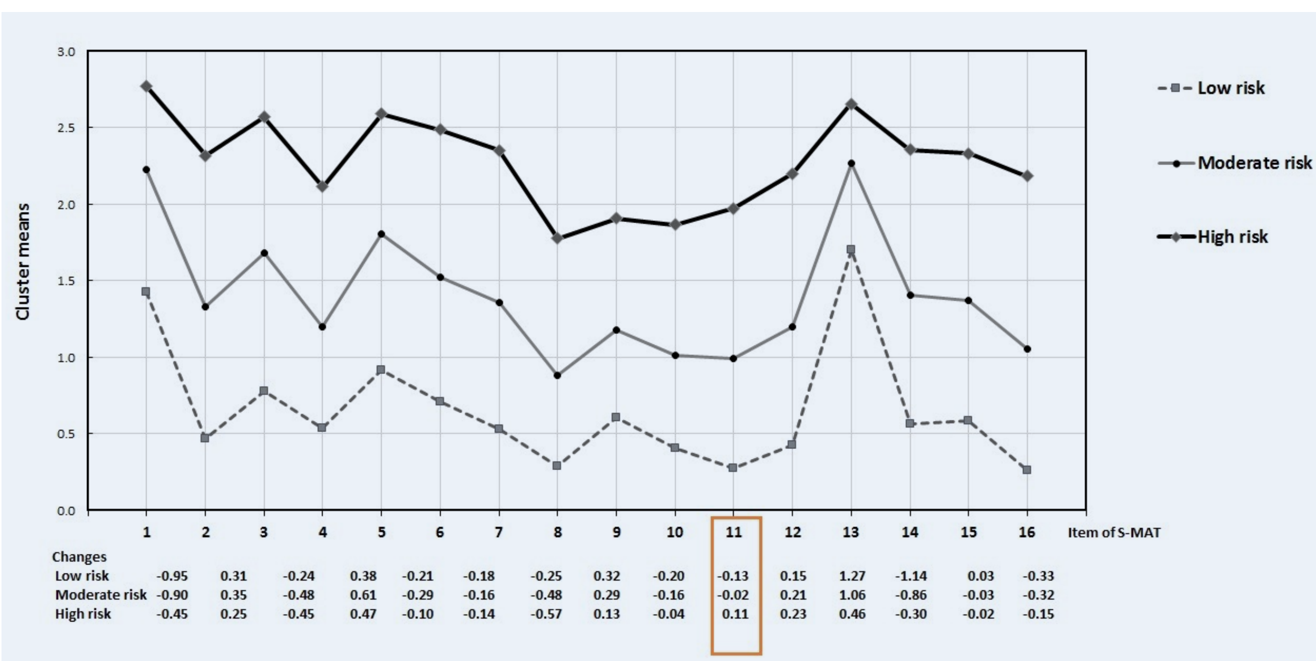


Fig 2. The three classes of social-media users obtained from the latent profile analysis

TABLE 5. Comparison of the three latent classes.

Characteristics	Total (n = 5,068)	Low-risk (n = 1,227)	Moderate-risk (n = 2,757)	High-risk (n = 1,084)	Overall test $\chi^2(df)$	*P*-value
Gender (Female); n (%)	2,963 (59.0)	687 (56.4)[a]	1557 (57.0)[a]	719 (67.2)[b]	37.77 (2)	< 0.001
Age (years) Median 16.00, min 13.00, max 75.00, SE 0.14: Mean (SD)	19.93 (9.96)	22.34 (11.91)[a]	19.49 (9.63)[b]	18.25 (7.55)[c]	34.30 (2)*	< 0.001
Necessity to use social media for work (Necessary); n (%)	3,029 (66.8)	715 (63.8)[a]	1,621 (66.3)[a]	693 (71.7)[b]	15.51 (2)	< 0.001
Self-perception of addiction n (%)						< 0.001
Probably addicted	2,245 (46.8)	452 (39.2)[a]	1,493 (57.0)[b]	300 (29.2)[c]	475.48 (2)	< 0.001
Addicted	1,462 (30.5)	80 (6.9)[a]	687 (26.2)[b]	695 (67.6)[c]	1,042.90 (2)	< 0.001
Amount of time spent on social media (not including gaming)						< 0.001
Weekdays (hrs/day) Median 2.50, min 0.08, max 18.00, SE 0.04: Mean (SD)	3.23 (2.63)	2.24 (1.97)[a]	3.16 (2.42)[b]	4.57 (3.20)[c]	526.81 (2)*	< 0.001
Weekends (hrs/day) Median 4.00, min 0.08, max 24.00, SE 0.06: Mean (SD)	4.91 (4.12)	3.05 (2.79)[a]	4.78 (3.60)[b]	7.41 (5.27)[c]	672.17 (2)*	< 0.001
Average daily time spent (hrs/day) Median 2.86, min 0.08, max 19.71, SE 0.04: Mean (SD)	3.68 (2.77)	2.46 (1.97)[a]	3.61 (2.51)[b]	5.28 (3.37)[c]	614.49 (2)*	< 0.001
Total Social Media Addiction Screening Scale (S-MASS) Score Median 22.00, min 0, max 48, SE 0.14: Mean (SD)	22.58 (9.67)	10.48 (3.72)[a]	22.51 (4.10)[b]	36.47 (4.52)[c]	4,133.80 (2)*	< 0.001

Note: Different superscript letters (a, b, c) in the same row reflect a significant (p-value < 0.05) difference between the means, while same superscript letters in one row reflect a non-significant difference between the means, according to the Pearson chi-square or (*) Kruskal–Wallis test, followed by a Tukey post-hoc pairwise comparison.

of the S-MASS measured the same problem, namely, social media addiction (SMA). The EFA using a PCA with promax rotation demonstrated that the S-MASS has good factorial validity as all items had loading factors above 0.4. The EFA also revealed that 3 factors were foundational to the S-MASS and covered the essential characteristics of behavioral addiction (giving priority, impaired control, and negative consequences). The CFA also confirmed the three-factor model of the S-MASS.

According to the LPA, three classes of social media users were identified, based on their risk of addiction: high risk, moderate risk, and low risk. Members of the high-risk class were likely to (i) be female, (ii) be younger, (iii) have necessary work-related use, (iv) perceive themselves as being addicted to social media, and (v) spend more than 5 hours daily on social media (Table 5). These risk factors will help clinicians accurately identify social media users who might be at risk for SMA. Interestingly, item 11 of the S-MASS ("I ignore or fail when people tell me to cut down my social media use") showed an elevated level in the high-risk class relative to the other items. This item may be helpful in distinguishing the high-risk class from the low-risk and medium-risk classes. Furthermore, 21.4% (n = 1,084) of participants belonged to the high-risk class, based on the LPA; this is consistent with the prevalence of social-networking-site addiction (29.5%) reported by Tang and Koh.[41] On the other hand, the high-risk class proportion found in our study is somewhat lower than the previously reported prevalence of Facebook addiction in Thailand (41.8%)[42] and much higher than the rate for the at-risk group of SMA in Hungary (4.5%).[43] The disparity in the prevalence of SMA among countries is probably due to differences in the measurement tools and the sample populations used by the various studies. Nevertheless, cultural influence might also contribute to the disparity found between countries.[44]

The study also discovered that the S-MASS score is moderately positively correlated with average daily time spent on social media (*rs*[38] = 0.412; $p < 0.001$). In other words, the greater the amount of time spent on social media, the greater the risk of becoming addicted to social media. This finding may imply that the S-MASS can determine the severity of addiction.

Strengths of this study include its relatively large sample size, the heterogeneity of the participants, and the comprehensive assessment of S-MASS reliability and validity. Cut-off scores were also identified for three-level risk classification. This fulfills the gap in previous studies on assessment tools. More importantly, SMA manifests in a spectrum, not in binary categories. Having cut-points is beneficial for clinical practice. The diversity of sources from which the participants were recruited, especially the online survey, allowed us to recruit and enroll participants with relatively diverse sociodemographic backgrounds and a wider age range (13–75 years) than other studies (although the majority of participants were aged between 13 and 25 years; Table 1). The authors set out to develop the S-MASS for use in screening for SMA in general, not Facebook addiction only. S-MASS can, therefore, be applied to subgroups of social media users who interact with social networking sites other than Facebook.

Limitations

The present study has some limitations. First, results were based on a convenient sample, limiting the extent to which findings can be generalized to a broader population. Second, the S-MASS is a self-report questionnaire which is subject to several biases (such as social desirability and short-term recall). Third, the test-retest reliability was not evaluated to determine stability. Forth, the same cut-off might not be a one-size-fit-all since each age group shows different patterns of social media usage. Further studies to identify age-specific cut-offs are warranted. Finally, although a few forms of validity were tested in this study, other important types of validity should also be examined (e.g., concurrent, predictive, convergent, and discriminant validities).

Future studies

The S-MASS should be further validated—most notably, its criterion validity. This validation process can be undertaken after a formal diagnosis of SMA becomes available. In its present form, the S-MASS is best suited for use in epidemiologic studies. However, to test whether the S-MASS is sensitive to change after interventions is an interesting and worthwhile research pursuit. Once the S-MASS is proven to be adequately sensitive to change, it can also be used in clinical or interventional studies.

CONCLUSION

The Social-Media Addiction Screening Scale (S-MASS) is a psychometrically reliable and valid screening test for SMA. Two cut-offs are identified for risk classification. Further studies assessing the concurrent, predictive, convergent, and discriminant validity of the S-MASS in a more heterogeneous population are warranted. In addition, the criterion validity of the S-MASS to determine its sensitivity, specificity, and the appropriateness of the current recommended cut-off scores should be evaluated once formal diagnostic criteria for SMA become available.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Ms.Naratip Sanguanpanich, Ms.Thapanee Sangsawang, and Dr. Sasima Tongsai for their assistance with the statistical analyses.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

1. Georgiev D. How Much Time Do People Spend on Social Media? 2022 [Available from: https://review42.com/how-much-time-do-people-spend-on-social-media/.
2. Andreassen CS. Online social network site addiction: A comprehensive review. Current Addiction Reports. 2015;2(2): 175-84.
3. Griffiths MD, Kuss DJ, Demetrovics Z. Social networking addiction: An overview of preliminary findings. In: Kenneth PR, Laura CF, editors. Behavioral Addictions: criteria, evidence, and treatment: Academic Press; 2014. p. 119-41.
4. Hou Y, Xiong D, Jiang T, Song L, Wang Q. Social media addiction: Its impact, mediation, and intervention. Cyberpsychology: Journal of Psychosocial Research on Cyberspace. 2019;13(1).
5. Andreassen CS, Billieux J, Griffiths MD, Kuss DJ, Demetrovics Z, Mazzoni E, et al. The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: A large-scale cross-sectional study. Psychol Addict Behav. 2016;30(2):252-62.
6. Robinson A, Bonnette A, Howard K, Ceballos N, Dailey S, Lu Y, et al. Social comparisons, social media addiction, and social interaction: An examination of specific social media behaviors related to major depressive disorder in a millennial population. Journal of Applied Biobehavioral Research. 2019;24(1):e12158.
7. Shensa A, Sidani JE, Dew MA, Escobar-Viera CG, Primack BA. Social media use and depression and anxiety symptoms: A cluster analysis. Am J Health Behav. 2018;42(2):116-28.
8. Sun Y, Zhang Y. A review of theories and models applied in studies of social media addiction and implications for future research. Addict Behav. 2021;114:106699.
9. World Health Organization. ICD-11 for mortality and morbidity statistics 2018 [Available from: https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1448597234.
10. Karim R, Chaudhri P. Behavioral Addictions: An Overview. J Psychoactive Drugs. 2012;44(1):5-17.
11. Kircaburun K, Griffiths MD. Instagram addiction and the Big Five of personality: The mediating role of self-liking. J Behav Addict. 2018;7(1):158-70.
12. Smith MA. Social learning and addiction. Behav Brain Res. 2021;398:112954.
13. Wang X. Mobile SNS Addiction as A Learned Behavior: A Perspective from Learning Theory. Media Psychology. 2020;23(4):461-92.
14. Wilson K, Fornasier S, White KM. Psychological predictors of young adults' use of social networking sites. Cyberpsychol Behav Soc Netw. 2010;13(2):173-7.
15. Elphinston RA, Noller P. Time to face it! Facebook intrusion and the implications for romantic jealousy and relationship satisfaction. Cyberpsychol Behav Soc Netw. 2011;14(11):631-5.
16. Andreassen C, Torsheim T, Brunborg G, Pallesen S. Utvikling av en Facebook-avhengighetsskala. Psychol Rep. 2012;110(2): 501-17.
17. Turel O, Serenko A. The benefits and dangers of enjoyment with social networking websites. European Journal of Information Systems. 2012;21(5):512-28.
18. Wolniczak I, Caceres-DelAguila JA, Palma-Ardiles G, Arroyo KJ, Solís-Visscher R, Paredes-Yauri S, et al. Association between Facebook dependence and poor sleep quality: a study in a sample of undergraduate students in Peru. PLoS One. 2013;8(3):e59087.
19. Koc M, Gulyagci S. Facebook addiction among Turkish college students: the role of psychological health, demographic, and usage characteristics. Cyberpsychol Behav Soc Netw. 2013;16(4): 279-84.
20. Wu AM, Cheung VI, Ku L, Hung EP. Psychological risk factors of addiction to social networking sites among Chinese smartphone users. J Behav Addict. 2013;2(3):160-6.
21. Al-Menayes J. Psychometric properties and validation of the Arabic social media addiction scale. J Addict. 2015;2015:291743.
22. Phanasathit M, Manwong M, Hanprathet N, Khumsri J, Yingyeun R. Validation of the Thai version of Bergen Facebook addiction scale (Thai-BFAS). J Med Assoc Thai. 2015;98:108-17.
23. Dantlgraber M, Wetzel E, Schützenberger P, Stieger S, Reips U-D. Simple construct evaluation with latent class analysis: An investigation of Facebook addiction and the development of a short form of the Facebook Addiction Test (F-AT). Behav Res Methods. 2016;48(3):869-79.
24. van den Eijnden RJ, Lemmens JS, Valkenburg PM. The Social Media Disorder Scale. Computers in Human Behavior. 2016; 61:478-87.
25. Monacis L, De Palo V, Griffiths MD, Sinatra M. Social networking addiction, attachment style, and validation of the Italian version of the Bergen Social Media Addiction Scale. J Behav Addict. 2017;6(2):178-86.
26. Chung-Ying Lin AB, Per Nilsen, Mark D Griffiths, Amir H Pakpour. Psychometric validation of the Persian Bergen Social Media Addiction Scale using classic test theory and Rasch models. J Behav Addict. 2017;6(4):620-9.
27. Liu C, Ma J. Development and validation of the Chinese social media addiction scale. Personality and Individual Differences. 2018;134:55-9.
28. Savci M, Ercengiz M, Aysan F. Turkish adaptation of the Social Media Disorder Scale in adolescents. Noro Psikiyatr Ars. 2018;55(3):248-55.
29. Bouna-Pyrrou P, Aufleger B, Braun S, Gattnar M, Kallmayer S, Wagner H, et al. Cross-sectional and longitudinal evaluation of the social network use disorder and internet gaming disorder criteria. Front Psychiatry. 2018;9:692.
30. Yam C-W, Pakpour AH, Griffiths MD, Yau W-Y, Lo C-LM, Ng JM, et al. Psychometric testing of three Chinese online-related addictive behavior instruments among Hong Kong university students. Psychiatr Q. 2019;90(1):117-28.
31. Fung SF. Cross-cultural validation of the Social Media Disorder scale. Psychol Res Behav Manag. 2019;12:683-90.
32. Bendayan R, Blanca Mena MJ. Spanish version of the Facebook intrusion questionnaire (FIQ-S). Psicothema. 2019;31(2):204-9.
33. Shahnawaz MG, Rehman U. Social Networking Addiction Scale. Cogent Psychology. 2020;7(1):1832032.
34. American Psychiatric Association. Diagnostic and statistical

manual of mental disorders (DSM-5®): American Psychiatric Association; 2013.
35. Ko CH, Yen JY, Chen SH, Wang PW, Chen CS, Yen CF. Evaluation of the diagnostic criteria of Internet gaming disorder in the DSM-5 among young adults in Taiwan. J Psychiatr Res. 2014;53:103-10.
36. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53-5.
37. Brown TA. Confirmatory factor analysis for applied research: Guilford publications; 2015.
38. Griffiths M. A 'components' model of addiction within a biopsychosocial framework. Journal of Substance Use. 2005; 10(4):191-7.
39. Kaiser HF, Rice J. Little Jiffy, Mark IV. Educational and psychological measurement. Educational and Psychological Measurement 1974;34(1):111-7.
40. Young K. Internet addiction: diagnosis and treatment considerations. Journal of Contemporary Psychotherapy. 2009;39(4):241-6.
41. Tang CS-K, Koh YYW. Online social networking addiction among college students in Singapore: Comorbidity with behavioral addiction and affective disorder. Asian J Psychiatr. 2017;25: 175-8.
42. Khumsri J, Yingyeun R, Mereerat M, Hanprathet N, Phanasathit M. Prevalence of Facebook Addiction and Related Factors Among Thai High School Students. J Med Assoc Thai. 2015;98: S51-60.
43. Bányai F, Zsila Á, Király O, Maraz A, Elekes Z, Griffiths MD, et al. Problematic Social Media Use: Results from a Large-Scale Nationally Representative Adolescent Sample. PloS One. 2017;12(1):e0169839.
44. Cheng C, Lau Y-C, Chan L, Luk JW. Prevalence of social media addiction across 32 nations: Meta-analysis with subgroup analysis of classification schemes and cultural values. Addict Behav. 2021;117:106845.

Changes in the VARK Learning Style from the First to the Second Preclinical Year of Medical Students: A Follow-up Cross-sectional Questionnaire Study in a Thai Medical School

Punyapat Maprapho, M.D.*, Chantacha Sitticharoon, M.D., Ph.D.**, Pailin Maikaew, B.Sc.**, Vasu Lertsiripatarajit, M.D.**, Issarawan Keadkraichaiwat, M.Sc.**, Nipith Charoenngam, M.D.***

*Siriraj Health Science Education Excellence Center, Department of Education, **Department of Physiology, ***Department of Internal Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

ABSTRACT

Objective: This study aimed to determine percentage of students who changed (change group) and did not change (no-change group) in the visual (V)–aural (A)–reading/writing (R)–kinesthetic (K) learning style from the first (Preclinic1) to the second (Preclinic2) preclinical year; and compare academic performance (GPA, percentile of GPA, and achievement of study targets) and stress levels between these groups.

Materials and Methods: The VARK and research questionnaires were sent to students of the 2019 class at the end of Preclinic1 and again at Preclinic2. GPA and percentile of GPA were obtained from academic records while achievement of study targets and stress levels were from the research questionnaire.

Results: Most students were multimodal learners (65.03% in Preclinic1 and 69.51% in Preclinic2). From Preclinic1 to Preclinic2, 69.3% of students changed and 30.7% of students did not change their learning preferences. In Preclinic1 and Preclinic2, GPA and percentile of GPA were higher in the no-change compared with the change group ($p<0.01$ all). GPA in the change group was lower in Preclinic2 than that in Preclinic1 ($p<0.001$). Achievement of study targets and stress levels of the change group in Preclinic2 were lower than theirs in Preclinic1 and those of the no-change group in Preclinic2 ($p<0.05$ all). The students who changed their learning preferences might need to adapt to their new learning preferences probably leading to lower academic capability but less stress.

Conclusion: Students changed their learning preferences around 70% from Preclinic1 to Preclinic2. The change group exhibited lower academic capability but less stress.

Keywords: VARK; preclinic; learning preference; academic performance (Siriraj Med J 2023; 75: 181-190)

INTRODUCTION

There are a vast number of theories describing the learning styles according to various psychological constructs.[1] The VARK learning style, which categorizes students into four learning modules, comprises visual (V), aural (A), reading/writing (R), and kinesthetic (K).[2] This model proposes that students use different sensory modalities for processing knowledge and information.[1] The model is popularly used among educators to figure out what learning modalities their learners predominantly prefer.[3] Flemming described these four modalities of students' learning preferences in 1992.[4]

"V" students prefer to learn with graphics and symbols; they like using figures, pictures, and symbolic tools,

Corresponding author: Chantacha Sitticharoon
E-mail: chantacha.sit@mahidol.ac.th
Received 3 December 2022 Revised 21 December 2022 Accepted 27 December 2022
ORCID ID:http://orcid.org/0000-0001-8790-1300
https://doi.org/10.33192/smj.v75i3.260867

while movements or actions could disturb their focus.[5] "A" students prefer "heard" information and best learn through lectures and discussions with others,[4] although the noise could easily trouble them.[6] "R" students prefer printed words and text to obtain information. They learn effortlessly via text, books, and notes.[7] Unlike others, "K" students use all perceptual modalities in the process of learning.[4] They understand things seamlessly by doing, practicing, and having real hands-on experience. They prefer learning through physical skills and movement.[8] They are good at coordination and have a good sense of rhythm.[8] Nevertheless, one does not have to fall in only one type of the VARK learning style, as for many students, their learning preferences are multimodal.[4,9]

Keefe, a former president of the Learning Environments Consortium International, mentioned that it is important for the teachers to not only teach content but also understand the learning strategies of students and facilitate them to know their learning preferences.[10] Moreover, he stated that students' learning environment is one of the influencing factors that causes different learning preferences and responses of each individual.[10] Furthermore, different backgrounds, demographic data, gender, academic levels, cultures, and individual creativities affect the different distribution of students' VARK preferences.[7,11-13] Besides, VARK preferences have been shown to be associated with academic success.[14] "V" modality was positively correlated while "A" modality was negatively correlated with assessment performance in the first year nursing students.[14] However, another study reported that there was no significant impact of each VARK learning module on examination scores.[9] Accordingly, the association between each VARK module and academic performance is still open to question.

During preclinical years, students have to study loads of learning content and experience various learning environments,[12] which might cause adaptation of students' learning styles. This study aimed to determine 1) distribution of the VARK learning style in the first (Preclinic1) and the second (Preclinic2) preclinical years; 2) the mean VARK score of total, male, and female students in Preclinic1 and Preclinic2; 3) percentage of students who changed (change group) and did not change (no-change group) their VARK learning preferences from Preclinic1 to Preclinic2; 4) comparisons of academic performance (GPA, percentile of GPA, and achievement of study targets) and stress levels between the change and no-change groups; 5) correlations between each VARK score and other factors; and 6) factors that contributed to GPA of students by multivariate regression analysis.

MATERIALS AND METHODS

This study was approved by the Siriraj Institutional Review Board (COA no. Si 022/2015). This is a follow-up cross-sectional questionnaire study. Prior to the study, informed consent forms and questionnaires were sent out to all students of the 2019 class at the end of Preclinic1 (academic year 2014) and again at Preclinic2 (academic year 2015). The inclusion criteria were medical students of the 2019 class who voluntarily returned both the VARK and research questionnaires.

Type of curriculum and course setup

The Doctor of Medicine program at the Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand, is a six-year curriculum in which medical students were enrolled in the medical school after high school graduation.

For the program 2014, the first year, also known as a premedical year, consisted of basic sciences and general education subjects. The second year was Preclinic1 in which gross anatomy 1, gross anatomy 2, neuroanatomy, embryology, histology, physiology, biochemistry, and minor subjects were taught. The third year was Preclinic2 in which subjects including clinical pathology, pharmacology, immunology, pathology, microbiology, parasitology, and minor subjects were taught. The fourth to sixth years were clinical years.

The teaching methods used in the preclinical years included lectures, group discussions, and practical sessions, together with clinical case studies associated with delivered content.

Questionnaire

The English language VARK questionnaire version 7.8 used in this study was from a previously published paper[4] with no translation, consisting of 16 four-choices questions which allows participants to select more than one choices if a single answer does not match their perception. Students were classified according to their VARK modules into 4 main VARK categories and 15 subcategories including unimodal (V, A, R, or K); bimodal (VA, VR, VK, AR, AR, AK, or RK); trimodal (VAR, VAK, VRK, or ARK); and quadrimodal (VARK).[15] The change group represented students who had different VARK learning subcategories between Preclinic1 and Preclinic2, while the no-change group was defined as students who had the same VARK learning category between Preclinic1 and Preclinic2.

The research questionnaire, written in Thai, was a self-report form consisting of questions regarding gender, time spent on the recorded-e-lecture study, time spent on study materials rather than the recorded-e-lecture study, percentage of achievement of study targets, and stress levels of students. Recorded-e-lecture study is one of the methods of students' lesson review using the recorded video from a regular class that is provided on the intranet after the class. Gender of participants was obtained because there were gender differences in VARK learning preferences, academic achievement, and study habits.[13,16] Stress levels were determined using a Likert scale (1=extremely low, 2=low, 3=medium, 4=high, and 5=extremely high). The legibility and clarity of the questionnaire were initially reviewed by medical students. Then the questionnaire was submitted to the committee of experts for revision and validation on readability, clarity, rational analysis, and comprehensiveness. The internal consistency of data collection, calculated from Cronbach's alpha, was 0.893.

Academic achievement

Academic achievement represented as GPA, percentile of GPA, and scores of all subjects was obtained officially from the undergraduate education department.

Subgroup analysis

Students who had the same learning preferences between Preclinic1 and Preclinic2 were allocated into the "no-change" group, while students who had different learning preferences between Preclinic1 and Preclinic2 were allocated into the "change" group.

Statistical analysis

Data were analyzed using the Statistical Package for Social Science version 18. Descriptive statistics were used to analyze the percentage of the VARK learning style. For comparisons between two independent groups, the independent sample T-test was used. The paired sample T-test was performed to analyze the comparisons between two related data. As there was no previous study that compared academic performance and stress levels between the change and no-change groups, we could not calculate sample size from comparisons between two groups. However, we calculated sample size from correlations between two-factor analysis with the equation $N= ((Z\alpha/2+Z\beta)^2)/C+3$ by setting $\alpha=0.05$, $Z\alpha=1.96$, $\beta=0.2$, $Z\beta=0.84$, type I error=0.05, type II error=0.2, H0: rho=0, H1: rho=0.25, and $C=0.5 \times \ln [(1+r)/(1-r)]$ leading to N=124. So, at least 124 questionnaire respondents are enough for the analysis. To determine correlations between two factors, represented as correlation coefficient (R value), the Pearson product-moment correlation coefficient was used for continuous variables, and Spearman's rank-order correlation coefficient was used for rank or non-normal distributed variables. Non-normal distributed data or rank data were analyzed using the nonparametric tests. Multiple regression analysis was used to determine factors that significantly contributed to students' GPA in each academic year. A p-value less than 0.05 is considered statistical significance.

RESULTS

The VARK and research questionnaires were returned 87.20% (286/328) and 92.99% (305/328) at the end of Preclinic1 and Preclinic2, respectively. The age range of Preclinic1 and Preclinic2 students was 19–21 years.

Distribution of the VARK learning style

Distribution of the VARK learning style in male and female students in Preclinic1 and Preclinic2 is presented in Fig 1. The unimodal learning preferences comprises V, A, R, and K, while the multimodal learning styles consist of bimodal, trimodal, and quadrimodal learning preferences. Most students were multimodal learners (65.03%, 67.66%, and 61.34% of total, male, and female students, respectively, in Preclinic1; and 69.51%, 66.10%, and 74.22% of total, male, and female students, respectively, in Preclinic2), while the rest of them were unimodal learners (34.97%, 32.34%, and 38.66% of total, male, and female students, respectively, in Preclinic1; and 30.49%, 33.90%, and 25.78% of total, male, and female students, respectively, in Preclinic2).

The mean VARK score

The mean of each VARK score and the standard error of the mean (S.E.M.) of total, male, and female students in Preclinic1 and Preclinic2 are shown in Table 1. In Preclinic1, K was the highest mean score in total, male, and female students while in Preclinic2, V was the highest mean score in total and male students and A was the highest mean score in female students.

Changes in the VARK learning style from Preclinic1 to Preclinic2

When these students progressed from Preclinic1 to Preclinic2, one-third of them (n=85, 30.7%) did not change their learning styles, while two-thirds of them (n=192, 69.3%) did. The same trend was observed with regard to gender; 49 males (30.2%) and 36 females (31.3%) did not change their learning styles while 113 males (69.8%) and 79 females (68.7%) did.

Changes in the VARK learning style of students

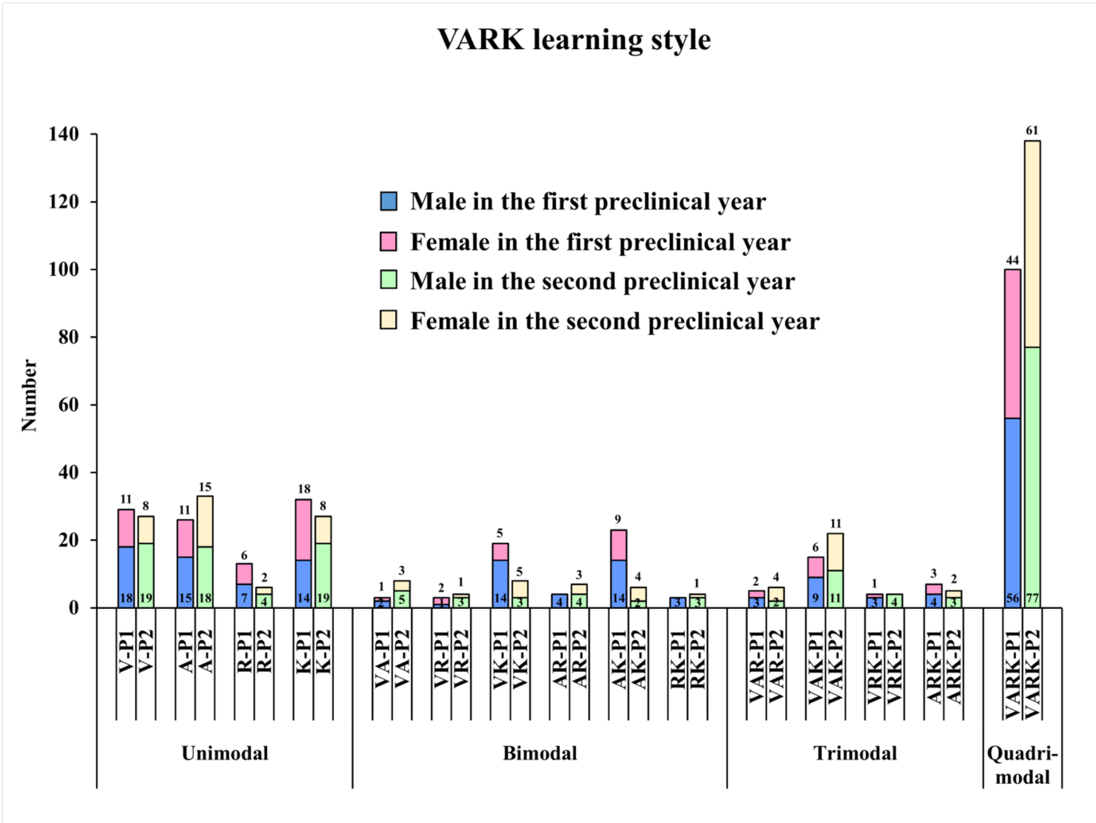


Fig 1. Distribution of the VARK learning style in male and female students in the first (P1) and the second (P2) preclinical years. The unimodal learning preferences comprises visual (V), aural (A), reading/writing (R), and kinesthetic (K) learning styles, while the multimodal learning styles consist of bimodal (VA, VR, VK, AR, AK, or RK), trimodal (VAR, VAK, VRK, or ARK), and quadrimodal (VARK) learning preferences.

TABLE 1. The mean of each VARK score of total, male, and female students in the first and the second preclinical years.

VARK Score	V Mean±S.E.M.	A Mean±S.E.M.	R Mean±S.E.M.	K Mean±S.E.M.
1st Preclinical year				
Total	5.09±0.15	4.98±0.15	3.91±0.14	5.53±0.14
Male	5.31±0.22	5.13±0.21	4.05±0.20	5.61±0.19
Female	4.79±0.20	4.83±0.21	3.68±0.21	5.53±0.21
2nd Preclinical year				
Total	6.37±0.19	6.21±0.18	4.79±0.16	6.32±0.17
Male	6.67±0.26	5.95±0.24	4.93±0.22	6.43±0.23
Female	5.98±0.27	6.59±0.26	4.63±0.24	6.24±0.27

Highlighted areas represent the highest mean score in each category.

from Preclinic1 to Preclinic2 are summarized and shown in details in Table 2. The number of students who did not change their learning preferences from Preclinic1 to Preclinic2 was 6 for V, 10 for A, 8 for K, 1 for VA, 1 for VR, 2 for VK, 2 for AK, 1 for VAR, 3 for VAK, and 51 for VARK (Table 2).

Comparisons of academic performance and stress levels between the no-change and change groups

Comparisons of academic factors and stress levels between the no-change and change groups are shown in Fig 2. In Preclinic1 and Preclinic2, GPA (Fig 2A) and percentile of GPA (Fig 2B) in the no-change group were

TABLE 2. Details of changes in the VARK learning style from the first to the second preclinical year.

1st Preclinical year			2nd Preclinical year: Unimodal V	Unimodal A	Unimodal R	Unimodal K	Bimodal VA	Bimodal VR	Bimodal VK	Bimodal AR	Bimodal AK	Bimodal RK	Trimodal VAR	Trimodal VAK	Trimodal VRK	Trimodal ARK	Quadrimodal VARK	Total
Unimodal	V	N	6	3									1	1	1		16	**28**
		%	21.4	10.7									3.6	3.6	3.6		57.1	**100**
	A	N		10	1							1		2			10	**24**
		%		41.7	4.2							4.2		8.3			41.7	**100**
	R	N				2	1			1		1	1				6	**12**
		%				16.7	8.3			8.3		8.3	8.3				50.0	**100**
	K	N	3	2		8			2		1		3	2			9	**30**
		%	10.0	6.7		26.7			6.7		3.3		10.0	6.7			30.0	**100**
Bimodal	VA	N					1							1			1	**3**
		%					33.3							33.3			33.3	**100**
	VR	N						1									2	**3**
		%						33.3									66.7	**100**
	VK	N	2			4			2		1			2	1	1	6	**19**
		%	10.5			21.1			10.5		5.3			10.5	5.3	5.3	31.6	**100**
	AR	N			1											1	2	**4**
		%			25.0											25.0	50.0	**100**
	AK	N	2	3		2	1			1	2			2			10	**23**
		%	8.7	13.0		8.7	4.3			4.3	8.7			8.7			43.5	**100**
	RK	N				1		1							1			**3**
		%				33.3		33.3							33.3			**100**
Trimodal	VAR	N	1							2			1				1	**5**
		%	20.0							40.0			20.0				20.0	**100**
	VAK	N	2	1		2	1							3			5	**14**
		%	14.3	7.1		14.3	7.1							21.4			35.7	**100**
	VRK	N			1												3	**4**
		%			25												75	**100**
	ARK	N	1	2	1	1								1			1	**7**
		%	14.3	28.6	14.3	14.3								14.3			14.3	**100**
Quadrimodal	VARK	N	8	8	2	4	4		2	3	2	2		8	1	3	51	**98**
		%	8.2	8.2	2.0	4.1	4.1		2.0	3.1	2.0	2.0		8.2	1.0	3.1	52.0	**100**
Total																		**277**

Abbreviations: V=visual, A=aural, R=reading/writing, and K=kinesthetic learning styles.
Highlighted areas represent the number and percentage of students who had the same VARK learning styles between the first and the second preclinical years

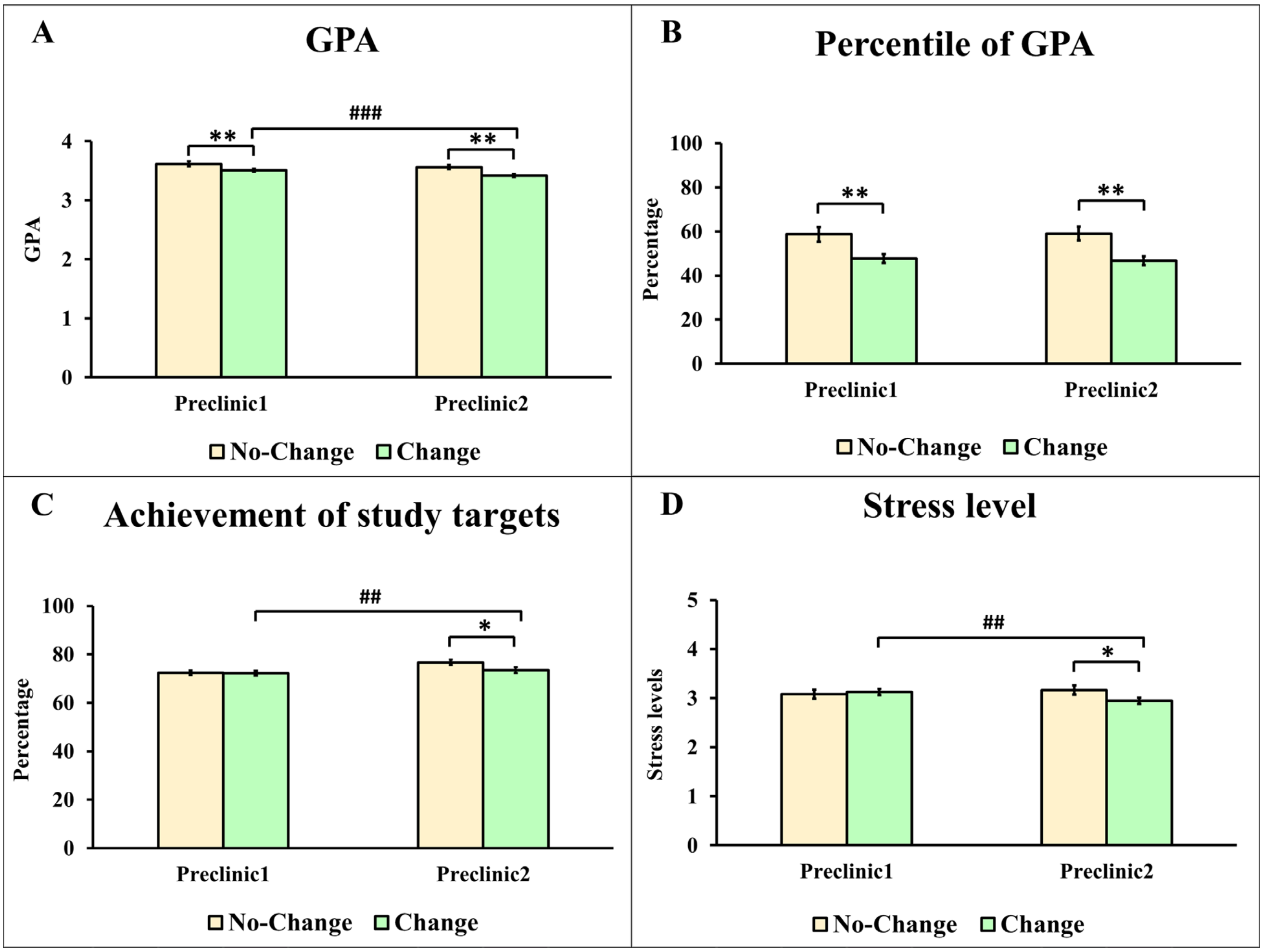


Fig 2. Comparisons between students who had the same (no-change) and different (change) learning preferences between the first and the second preclinical years with regard to GPA (A), percentile of GPA (B), percentage of achievement of study targets (C), and stress levels (D). Data are shown as mean (S.E.M.), *p<0.05 **p<0.01, ***p<0.001 compared between the no-change and change groups; ##p<0.01, ###p<0.001 compared between the first and the second preclinical years.

significantly higher than those in the change group (p<0.01 all). Interestingly, the change group had significantly lower GPA in Preclinic2 than Preclinic1 (Fig 2A; p<0.001).

In Preclinic1, achievement of study targets (Fig 2C) and stress levels (Fig 2D) in the no-change group were comparable to those in the change group, but in Preclinic2 these factors were significantly higher in the no-change group compared with the change group (p<0.05 all). Furthermore, the change group had significantly lower achievement of study targets (Fig 2C) and stress levels (Fig 2D) in Preclinic2 than Preclinic1 (p<0.01 all).

Correlations between each VARK score and other factors

Correlations between each VARK score and other factors are shown in Table 3. In Preclinic1, the V score had positive correlations with GPA (R=0.134), percentile of GPA (R=0.119), and scores of gross anatomy 1 (R=0.135), gross anatomy 2 (R=0.187), histology (R=0.143), embryology (R=0.127), neuroanatomy (R=0.146), and biochemistry (R=0.121, p<0.05 all); and showed a positive correlation trend with the physiology score (R=0.115, p=0.052; Table 3). The A score was positively correlated with time spent on the recorded-e-lecture study (hours/week; R=0.126, p<0.05; Table 3). The R score had a positive correlation with time spent on the non-recorded-e-lecture study (hours/week; R=0.122, p<0.05; Table 3). The K score was positively correlated with the embryology score (R=0.135, p<0.05; Table 3). In Preclinic2, the R score was positively correlated with time spent on the non-recorded-e-lecture study (hours/week; R=0.160), while the K score was negatively correlated with time spent on the recorded-e-lecture study (hours/week; R=–0.113, p<0.05 all; Table 3).

Multivariate regression analysis of GPA of students in Preclinic1

Multivariate regression analysis of GPA of students in

TABLE 3. The mean of each VARK score of total, male, and female students in the first and the second preclinical years.

VARK score	V score		A score		R score		K score	
Factors	R	P	R	P	R	P	R	P
1st Preclinical Year								
GPA	0.134	0.025*	-0.024	0.687	0.089	0.139	0.072	0.232
Percentile GPA	0.119	0.047*	-0.026	0.661	0.085	0.158	0.018	0.764
Gross Anatomy 1 score	0.135	0.022*	-0.043	0.473	0.047	0.428	0.086	0.145
Gross Anatomy 2 score	0.187	0.001**	-0.084	0.158	0.069	0.242	0.075	0.207
Histology score	0.143	0.016*	-0.038	0.523	0.107	0.071	0.113	0.056
Embryology score	0.127	0.032*	-0.052	0.378	0.083	0.161	0.135	0.022*
Neuroanatomy score	0.146	0.014*	-0.047	0.424	0.047	0.426	0.056	0.349
Biochemistry score	0.121	0.042*	-0.040	0.501	0.066	0.263	0.090	0.131
Physiology score	0.115	0.052	-0.070	0.236	0.089	0.135	0.070	0.239
Time spent on non-recorded-e-lecture study (hours/week)	0.056	0.343	0.056	0.343	0.122	0.038*	0.109	0.064
Time spent on recorded-e-lecture study (hours/week)	-0.057	0.334	0.126	0.032*	-0.021	0.718	-0.084	0.152
2nd Preclinical Year								
GPA	0.054	0.346	-0.060	0.294	-0.004	0.939	-0.052	0.362
Percentile GPA	0.028	0.625	-0.036	0.535	0.008	0.890	-0.062	0.281
Time spent on non-recorded-e-lecture study (hours/week)	0.092	0.107	-0.035	0.541	0.160	0.005**	0.010	0.866
Time spent on recorded-e-lecture study (hours/week)	-0.029	0.609	0.058	0.313	-0.012	0.839	-0.113	0.048*

*p<0.05, **p<0.01; R=correlation coefficient

Preclinic1 is shown in Table 4. The factor that significantly contributed to students' GPA in Preclinic1 was the V score (R=0.134, p<0.05; Table 4).

DISCUSSION

This study determined the VARK learning style of students in two consecutive years and the changes in their learning preferences from Preclinic1 to Preclinic2 associated with their academic performance and stress levels, to reveal factors associated with the change. To the best of our knowledge, this is the first follow-up study that compared factors between students who changed and did not change their learning preferences in two preclinical years.

With regard to the distribution of the VARK learning style, most students were multimodal learners (60-75% in Preclinic1 and Preclinic2), while the rest of them were unimodal learners. Our results were consistent with previous studies showing that the majority of students' learning preferences were multimodal.[9,17-22]

When observing only the unimodal preferences, the dominant learning module was K in Preclinic1 but was A in Preclinic2. In Preclinic1, students needed to study content together with practical sessions including preclinical knowledge, cadaver dissection, and laboratory studies. These kinds of studies might be suitable for the K module.

By contrast, in Preclinic2, there were much more subjects, types of content, and clinical case scenarios with many more hours of lectures compared with the previous

TABLE 4. Multivariate regression analysis of GPA of students in the first preclinical year.

Factor	R	R^2	P value		Coefficient	Standard Error	T value	P value
GPA	0.134	0.018	0.025*	(Constant)	3.433	0.052	65.528	<0.001***
				Visual score	0.021	0.009	2.250	0.025*

*p<0.05, ***p<0.001

year. These study modes might be more effective with the A learning preference as A students like to learn via lectures and discussions.[4] These findings support that students might change their learning preferences to fit the learning environments they experience. However, quadrimodal preference is the majority of students' learning preferences in both years, indicating that to handle vast medical content in preclinical years, the combination of multiple modes of learning might be needed.

There were tremendous changes in the VARK learning preferences (around 70%) from Preclinic1 to Preclinic2, which probably might be because the learning they used in Preclinic1 did not fit with the new learning environments. However, around 30% of these students did not change their learning styles. The most preferred learning styles of the students who had the same learning preferences from Preclinic1 to Preclinic2 was quadrimodal (51/98, 52%). These results emphasize what we mentioned earlier that multiple skills are needed to study in preclinical years in accordance with previous studies.[14,23,24] Furthermore, A learning preference was the second highest retained learning styles, which might be because the A module matches the learning environments of Preclinic2 as previously mentioned.

When comparing the no-change group with the change group, the no-change group had higher GPA and percentile of GPA than the change group in both preclinical years. Furthermore, GPA of students in the no-change group was comparable between Preclinic1 and Preclinic2, while that of students in the change group was significantly lower in Preclinic2 compared with Preclinic1. In addition, percentage of achievement of study targets was comparable between Preclinic1 and Preclinic2 in the no-change group, but this factor was lower in Preclinic2 compared with Preclinic1 in the change group. These results suggest that students who did not change their learning preferences had better academic outcomes and academic performance on reaching their study goals than students who changed their learning preferences. A better academic outcome in the no-change group might be explained by the supporting evidence showing that the achievement of study targets was positively associated with academic performance.[16,25,26]

We hypothesized that students who did not change their learning styles had been practicing their skills to the level of high proficiency letting them be skillful in what they used to learn. Therefore, when they used the same skill day by day, they could effectively learn with less effort. By contrast, students who changed their learning preferences could not learn effectively because the previous learning styles did not suit new learning environments or did not help them reach academic performance at their level of expectation or satisfaction. Nevertheless, these students had to practice their new learning styles causing less competency or efficiency leading to less achievement of their study targets and academic outcomes. Therefore, students who found their compatible learning preferences and kept practicing the skills to the master level could maintain their academic performance.

Interestingly, stress levels of the no-change group were not different between Preclinic1 and Preclinic2; however, those of the change group were lower in Preclinic2 compared with Preclinic1. These results suggest that the students in the change groups felt less stress when they changed to the new learning styles. Even though these changes did not improve their academic performance, these students might be more comfortable using new learning preferences leading to reduced stress.

For the mean of each VARK score, the K score was highest in Preclinic1 followed by V, A, and R, respectively. While in Preclinic2, the V score was highest followed by K, A, and R, respectively. Interestingly, the mean R score was lowest in both preclinical years. Our results in Preclinic1 were in accordance with a previous study on the first-year preclinical students from Nepal and India showing that the K score was highest and the R score was lowest.[22,27] The possible explanation of why the R score was lowest might be because reading and writing skills are not natural for humans.[28] Reading and writing, unlike other skills, are developed because human brain is not naturally wired to read and write.[28]

The V score was positively correlated with academic outcomes including GPA, percentile of GPA, and scores of gross anatomy 1, gross anatomy 2, neuroanatomy, embryology, histology, and biochemistry in Preclinic1. As the V score is related to figures, pictures, and what students can see,[5] it is not surprising that the V score was associated with anatomy-related subjects similarly to a previous study.[14] In regression analysis, the V score was the only score that was positively associated with GPA. This might be because it was positively correlated with the scores of many subjects in Preclinic1; thus, students who had high V score might get a better GPA in the academic year.

The A score was positively correlated with time spent on the recorded-e-lecture study in Preclinic1. Recorded-e-lecture was provided to students after class, and they could access this material format any time. Students who were A learners might get along well with a lesson review with repeated lectures letting them spend more time on the record-e-lecture study.

The R score was positively correlated with time spent on the non-recorded-e-lecture study in both preclinical years. The non-recorded-e-lecture study refers to a study using other materials, including text, books, notes, and interactive materials, rather than recorded-e-lecture. Therefore, these similar results found in both preclinical years reflect that students with a higher R score might spend a higher amount of time on other materials rather than recorded-e-lecture.

The K score was negatively correlated with time spent on the recorded-e-lecture study in Preclinic2. As the K learning preference is related to practicing and real hands-on experiences,[4] students with a high K score might be less satisfied with a lesson review by recorded-e-lecture.

CONCLUSION

Most students were multimodal learners in Preclinic1 and Preclinic2. The highest mean VARK score was K in Preclinic1 and V in Preclinic2. Around 70% of students changed their VARK learning styles from Preclinic1 to Preclinic2. Students who changed their learning preferences had less stress but lower academic performance than students who did not. Mean V score was positively but weakly associated with academic scores and could slightly contribute to students' GPA in Preclinic1. "A" score had a weakly positive correlation with time spent on the recorded e-lecture study in Preclinic1 while "R" score had weakly positive correlations with time spent on the non-recorded e-lecture study reflecting that the designated VARK preferences corresponded with their preferable learning materials of choices.

Limitations

Stress levels of students were self-reported and not measured with standardized questionnaires; therefore, this information might not be validated. Furthermore, data on achievement of study targets, time spent on the non-recorded-e-lecture study, and time spent on the recorded-e-lecture study were also self-reported, which could probably demonstrate students' estimation, not the exact numbers. The results from this study could not determine a cause-and-effect relationship. Accordingly, the explanation of our observation was made from gathering knowledge/principles/theoretical concepts/results from previous research.

ACKNOWLEDGEMENTS

This study was supported by the Faculty of Medicine Siriraj Hospital Medical Education Research Fund. CS is supported by the "Chalermphrakiat" Grant, Faculty of Medicine Siriraj Hospital, Mahidol University.

Conflicts of interest: The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

1. Biggs J. From Theory to Practice: A Cognitive Systems Approach. Higher Education Research & Development. 1993;12(1):73-85.
2. Prithishkumar IJ, Michael SA. Understanding your student: using the VARK model. J Postgrad Med. 2014;60(2):183-6.
3. Dryden G, Vos J. The learning revolution : to change the way the world learns. Torrance, Calif.: The Learning Web; 1999.
4. Fleming ND, Mills C. Not Another Inventory, Rather a Catalyst for Reflection. 1992;11(1):137-55.
5. Othman N, Amiruddin M. Different Perspectives of Learning Styles from VARK Model. Procedia - Social and Behavioral Sciences. 2010;7:652-60.
6. Drago WA, Wagner R. Vark preferred learning styles and online education. Management Research News. 2004;27:1-13.
7. Murphy RJ, Gray SA, Straja SR, Bogert MC. Student learning preferences and teaching implications. J Dent Educ. 2004;68(8):859-66.
8. Armstrong A. Instructional Design in the Real World: A View from the Trenches: Information Science Pub.; 2004.
9. Farkas GJ, Mazurek E, Marone JR. Learning style versus time spent studying and career choice: Which is associated with success in a combined undergraduate anatomy and physiology course? Anat Sci Educ. 2016;9(2):121-31.

10. Keefe JW. Learning style: theory and practice. Reston, Va.: National Association of Secondary School Principals; 1987.
11. Alkhasawneh IM, Mrayyan MT, Docherty C, Alashram S, Yousef HY. Problem-based learning (PBL): assessing students' learning preferences using VARK. Nurse Educ Today. 2008;28(5):572-9.
12. Baykan Z, Nacar M. Learning styles of first-year medical students attending Erciyes University in Kayseri, Turkey. Adv Physiol Educ. 2007;31(2):158-60.
13. Sarabi-Asiabar A, Jafari M, Sadeghifar J, Tofighi S, Zaboli R, Peyman H, et al. The relationship between learning style preferences and gender, educational major and status in first year medical students: a survey study from iran. Iran Red Crescent Med J. 2015;17(1):e18250.
14. Good JP, Ramos D, D'Amore DC. Learning style preferences and academic success of preclinical allied health students. J Allied Health. 2013;42(4):81-90.
15. Leite WL, Svinicki M, Shi Y. Attempted Validation of the Scores of the VARK: Learning Styles Inventory With Multitrait–Multimethod Confirmatory Factor Analysis Models. Educational and Psychological Measurement. 2009;70(2):323-39.
16. Lertwilaiwittaya P, Sitticharoon C, Maikaew P, Keadkraichaiwat I. Factors influencing the National License Examination step 1 score in preclinical medical students. Adv Physiol Educ. 2019; 43(3):306-16.
17. Aldosari MA, Aljabaa AH, Al-Sehaibany FS, Albarakati S. Learning style preferences of dental students at a single institution in Riyadh, Saudi Arabia, evaluated using the VARK questionnaire. Adv Med Educ Pract. 2018;9:179-86.
18. Panambur S, Nambiar V, Heming T. Learning style preferences of preclinical medical students in oman. Oman Med J. 2014;29(6): 461-3.
19. Samarakoon L, Fernando T, Rodrigo C. Learning styles and approaches to learning among medical undergraduates and postgraduates. BMC Med Educ. 2013;13:42.
20. Almigbal TH. Relationship between the learning style preferences of medical students and academic achievement. Saudi Med J. 2015;36(3):349-55.
21. Narayanan S, Ananthy V. The influence of learning style in understanding analogies and 2D animations in embryology course. Anat Cell Biol. 2018;51(4):260-5.
22. Khanal L, Giri J, Shah S, Koirala S, Rimal J. Influence of learning-style preferences in academic performance in the subject of human anatomy: an institution-based study among preclinical medical students. Adv Med Educ Pract. 2019;10:343-55.
23. Breckler J, Joun D, Ngo H. Learning styles of physiology students interested in the health professions. Adv Physiol Educ. 2009;33(1):30-6.
24. Lujan HL, DiCarlo SE. First-year medical students prefer multiple learning styles. Adv Physiol Educ. 2006;30(1):13-6.
25. Sitticharoon C, Srisuma S, Kanavitoon S, Summachiwakij S. Exploratory study of factors related to educational scores of first preclinical year medical students. Adv Physiol Educ. 2014;38(1): 25-33.
26. Raksadawan Y, Sitticharoon C, Charoenngam N, Maikaew P, Keadkraichaiwat I. Factors influencing academic achievement in preclinical students. International Educational Research. 2020; 3(3):8.
27. Kharb P, Samanta PP, Jindal M, Singh V. The learning styles and the preferred teaching-learning strategies of first year medical students. J Clin Diagn Res. 2013;7(6):1089-92.
28. Liberman IY, Shankweiler D, Liberman AM. The alphabetic principle and learning to read. Phonology and reading disability: Solving the reading puzzle. International Academy for Research in Learning Disabilities monograph series, No. 6. Ann Arbor, MI, US: The University of Michigan Press; 1989.p.1-33.

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

Jarunee Songmuang, M.N.S.*, Pachcharat Akatat, R.N.*, Sasithorn Wattanagulanurak, M.N.S.*, Nuttapol Rittayamai, M.D.**, Siriporn Sripalakit, M.A. (Env.) ***

** Nursing Department, ** Division of Respiratory Disease and Tuberculosis, Department of Medicine, *** Centre for Nosocomial Infection Control, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Objective: This research aimed to compare the incidence of ventilator-associated pneumonia (VAP) and ventilator circuit costs among patients using disposable ventilator circuits and patients using non-disposable ventilator circuits.

Materials and Methods: Observational research was performed consisting of the following: A retrospective chart review of a group of patients who used non-disposable circuits (n=193) and a prospective cohort study of a group of patients using disposable circuits (n=166). The sample was purposively selected based on the following inclusion criteria: patients aged 18 years old and over who were admitted to the Respiratory Care Unit, Siriraj Hospital and ventilated >48 h.

Results: VAP incidence in the group non-disposable circuits was 10.41/1,000 ventilator days *(n=27, 13.8%)* and 10.82 /1,000 ventilator days *(n=24, 14.4%)* in the group disposable circuits *(p=0.871)*. According to the data analysis using the U-control chart, no statistically significant differences were found. The unit cost of the non-disposable circuit was lower than that of the disposable circuit (THB 295.94), while the work unit personnel satisfaction toward working with disposable circuits was at a good level *(Mean=3.83)* and medium level *(Mean=3.12)* in non-disposable circuits *(p=0.002)*.

Conclusion: The type of ventilator circuits had no effects on VAP rate. The unit cost of non-disposable circuits was lower than that of disposable circuits, while the work unit personnel had a higher satisfaction working with disposable circuits than non-disposable circuits.

Keywords: Ventilator-associated pneumonia; health care costs; non-disposable ventilator circuits; disposable ventilator circuits (Siriraj Med J 2023; 75: 191-199)

INTRODUCTION

Ventilator-associated pneumonia (VAP) is pneumonia occurring at least 48 h after intubation and ventilation or within 48 h after extubation. The diagnostic criteria and clinical symptoms are used for its diagnosis.[1,2] VAP is well known as a major hospital infection correlated with the time on ventilators, length of stay, mortality rate, and higher costs. One US study reported that treatment expenses were USD 25,000–28,000 higher per patient with VAP[5], while the mortality rate in connection with VAP was 5%–65%, depending on the patients' condition.[3,4] VAP incidence in intensive care units (ICUs), where most patients use ventilators, has been reported to range from 8%–28%[5,6], or 13–51/1,000 ventilator days.[7] The

Corresponding author: Jarunee Songmuang
E-mail: dah_ns@hotmail.com
Received 9 November 2022 Revised 16 December 2022 Accepted 27 December 2022
ORCID ID:http://orcid.org/0000-0001-9354-7694
https://doi.org/10.33192/smj.v75i3.260866

Respiratory Care Unit (RCU), Department of Medicine, Siriraj Hospital, provides care for patients with respiratory failure, unstable blood circulation, respiratory infections, air-borne infections, difficulty weaning from ventilators and severe complications from diagnostic procedures and treatment of respiratory illnesses. The VAP incidence rate was 0–16.9/1,000 ventilator days in 2018 and 0–27.2/1,000 ventilator days in 2019.[8] In addition to affecting patients' health and costs, the problem increases hospital and staff workloads, with more complex care and treatment needed for the prevention of infections, which then translates into increased hospital expenses and costs.

The etiology of VAP is attributed to obstructions to natural bacteria and mucus disposal mechanisms during intubation, while the formation of bacterial microfilms around the endotracheal tube (ETT) causes choking from secretions that accumulate above the balloon. In addition, the positive pressure from ventilators enables germs to easily enter the lower respiratory tract. This is facilitated by differences in immunity and VAP risk factors in each patient.[9,10] Many factors contribute to VAP, including patient factors and factors related to the examination and treatment procedures, such as the patient's age, disease severity, immunodeficiency, the use of proton-pump inhibitors, the use of narcotics or neuromuscular blocking agents, previous antibiotic exposure, emergency intubation, re-intubation, bronchoscopy, and contaminated medical equipment, such as nebulizers and ventilator circuits.[9,11,12] In the RCU, used ventilator circuits are sterilized before use by the next patient. However, if the cleaning and disinfection processes are performed inadequately, infections can break out among patients. According to a study conducted by Li et al. compared the rates of bacterial contamination in non-disposable and disposable ventilator circuits, and found the bacterial infection rates in the group using non-disposable ventilator circuits were higher than in the group using disposable ventilator circuits, with statistical significance (94.8% and 81.9% $p < 0.01$).[13] In Thailand, Srisan et al. conducted a study to compare the VAP incidence rates among pediatric patients on non-disposable and disposable ventilator circuits, and found the VAP incidence rates were not statistically significantly different (20.53 and 30.77/1,000 ventilator days, $p = 0.24$). Regarding the treatment agency expenses, however, the group using disposable ventilator circuits had higher costs.[14] Apart from the major health issues caused by VAP incidence, hospital costs and expenses, workloads, and infection risks to the staff involved in preparing ventilator circuits are also important. Although trends in the findings from previous studies have shown that the use of disposable or non-disposable ventilator circuits have no effect on VAP incidence, differences in the patients' condition, environment, care standards, and ventilator circuit cleaning guidelines may cause the results to be different. Consequently, the present research was conducted to compare the effects from using non-disposable ventilator circuits to the effects from using disposable ventilator circuits in the RCU in terms of VAP incidence, agency costs, and staff satisfaction with the aim to ensure that the existing resources can be used cost-effectively with the highest benefit. Furthermore, this study was conducted with the aim of creating evidence-based practices that can be used by other institutions.

MATERIALS AND METHODS

This observational study was conducted at Siriraj Hospital in Bangkok, Thailand. The Institutional Review Board of the Faculty of Medicine Siriraj Hospital, approved the study protocol (COA no. Si 136/2020). The research involved a study of the variances of VAP incidence. The population studied was the number of days on ventilators and the incidence of VAP. VAP was diagnozed by using the criteria of the Centers for Disease Control and Prevention.[1] VAP means pneumonia in patients who had been on ventilatures for more than two days. Diagnostic crteria and clinical symptoms consisted of chest x-ray images with new or additional infiltrations and at least two of the following three factors: (1) fever > 38°C or temperature below 36°C; (2) leukopenia (≤ 4,000 WBC/mm^3) or leukocytosis (≥ 12,000 WBC/mm^3); and (3) sputum containing pus or changing characteristics, higher volume or difficulty breathing. The patients who participated in the study had to meet the following inclusion criteria: (1) patients were intubated or on tracheostomy; (2) patients were on ventilators for at least 48 h; (3) patients were aged 18 years old and over; and (4) patients consented to participate in the research project. The exclusion criteria consisted of one of the following criteria: (1) patients were pregnant; and (2) patients were on ventilators for less than 48 h. The sample size calculation criteria of the American Society of Testing and Materials were used. The criteria specified a minimum sample size of more than or equal to 1/U bar and the appropriate sample size was 4/U bar (U bar or average VAP incidence/1,000 ventilator days). In this study, the sample size was calculated by using the VAP incidence rate in the period January 2018 to December 2018 of 10.8/1,000 ventilator days. Therefore, the appropriate sample size per dataset was 1/0.0108–4/0.0108 or 93–370 ventilator days. For the U-control chart to be able to compare and display VAP incidence variances with quality, at least 15 datasets were needed per group. In this study, the sample was divided into the following two groups: (1) A group using non-disposable ventilator circuits (Hamilton). The data of this group were collected retrospectively in the period October 2018 to December 2019 for a total of 15 datasets. In this group, 235 patients used ventilators and 192 patients met the inclusion criteria; (2) A group using disposable ventilator circuits (Fisher & Paykel, model RT 380). Due to the COVID-19 situation, the number of ventilator days was less than 93 days in certain months, causing the data from two months to be merged into one data set. In this group, the data were collected for 17 months between June 2020 and October 2021 (15 datasets). This group had a total of 201 patients who were on ventilators with 166 patients who met the inclusion criteria, 148 patients who were on ventilators and used non-disposable ventilator circuits before changing to use disposable ventilator circuits after joining the research project and 18 patients who began using ventilators in the RCU by using disposable ventilator circuits from the start.

Data collection

In the data collection, the patients' medical records were used to collect the data retrospectively in the group using non-disposable ventilators, while letters were prepared for the patients or their representatives to sign and grant consent to participate in the study before collecting data from the group using disposable ventilators. The researcher then began using the case record form to collect data on the patients' demographics, co-morbidities, APACHE II scores, ventilator indications, VAP risk factors, number of ventilator days, length of hospitalization, ICU mortality, 30-day mortality, survival past one month after transfer from the ICU, hospital charges, and use of antibiotics. Data on the ventilator circuit costs were collected by calculating the equipment prices, depreciation, expendable equipment costs, gassing costs, infected waste disposal costs, and staff wages throughout the process to summarize the average cost per ventilator circuit. Data on staff satisfaction were collected by using the satisfaction assessment form to compare their satisfaction between working with disposable ventilators and non-disposable ventilators. Data were collected for comparison at the same time point at the end of the research project.

Statistics

Fisher's exact test was applied for the data analysis, with the data analyzed using the Minitab (U-control chart) program to assess the variances in VAP incidence. The SPSS program was used for the descriptive and comparison statistics, with $p < 0.05$ indicating statistical significance. The data analysis was as follows: (1) sample characteristics were reported by the mean scores, percentages, and standard deviation; differences were compared using t-test statistics in the case of quantitative data, while chi-square statistics were used for the qualitative data; (2) clinical outcomes were reported as median scores and differences were compared using the Mann–Whitney U-test. Data on the length of hospitalization, length of stay in the ICU, number of ventilator days, expenses from the hospitalization and antibiotics, clinical outcomes in terms of the number of patients with VAP and the treatment outcomes before transfer from the ICU were reported using percentages, while the differences were compared using chi-square test statistics, Kaplan-Meier survival rate analysis by using the log-rank test, analysis of factors with effects on the hazard ratio according to the Cox Regression Model and the VAP incidence fluctuation assessment control graph; (3) staff satisfaction was reported by percentages, mean scores, and standard deviation, while differences were compared using paired sample test statistics and Fisher's exact test.

RESULTS

In this study, 358 participants were enrolled divided into two groups: the non-disposable ventilator circuits group, comprising 192 patients who used non-disposable ventilator circuits (53.6%) for a combined 2,527 ventilator days, and the disposable ventilator circuits group, comprising 166 patients who used disposable ventilator circuits (46.4%) for a combined 2,354 ventilator days. The demographic data of both groups were similar. Most of the study cohort were males, aged over 60 years old and had co-morbidities. The illness severity was assessed based on APACHE II scores. The patients in the disposable ventilator circuits group had lower scores than the patients in the non-disposable ventilator circuits group *($p = 0.007$)*. As a co-morbidity, chronic lung disease was higher in the group using disposable circuits than the group using non-disposable circuits *($p = 0.02$)*. The other co-morbidities showed no statistically significant difference. The main cause of respiratory failure was pneumonia. The risk of VAP from receiving proton-pump inhibitors was the only risk factor in the group using disposable ventilator circuits that was lower than in the group using non-disposable ventilator circuits *($p = 0.001$)*. Other aspects, such as antibiotics, sedatives and re-intubation, were not statistically significantly different (Table 1).

As for clinical outcomes, VAP incidence was reported at 51 episodes divided into 27 episodes in the group

TABLE 1. Patient characteristics.

Patient data *(n = 358)*	Non-disposable *(n = 192)*	Disposable *(n = 166)*	*P*
Male, *n (%)*	120 (62.5)	100 (60.2)	.661
Age *(years) (mean ± SD)*	65.01 ± 17.16	65.83 ± 16.03	.642
BMI *(mean ± SD)*	21.95 ± 6.09	22.97 ± 6.70	.130
APACHE II Score *(mean ± SD)*	24.13 ± 7.16	22.01 ± 7.50	.007*
Co-morbidities, *n (%)*	178 (92.7)	157 (94.6)	.472
Hypertension	94 (49.0)	92 (55.4)	
Chronic Lung Disease	67 (34.9)	78 (47.0)	
Cardiovascular Disease	52 (27.1)	46 (27.7)	
Chronic Renal Failure	46 (24.0)	33 (19.9)	
Diabetes	44 (22.9)	52 (31.3)	
Cancer	44 (22.9)	32 (19.3)	
Stroke	27 (14.1)	20 (12.0)	
Connective Tissue	21 (10.9)	12 (7.2)	
Other	27 (14.1)	23 (13.8)	
Cause of Intubation, *n (%)*			
Pneumonia	81 (42.2)	82 (49.4)	.172
Obstructive Airway Disease	23 (12.0)	13 (7.8)	.193
ARDS	20 (10.4)	13 (7.8)	.399
Heart Failure	10 (5.2)	9 (5.4)	.928
Alteration of Consciousness	8 (4.2)	11 (6.6)	.301
Cardiac Arrest	8 (4.2)	5 (3.0)	.560
Aspiration	7 (3.6)	8 (4.8)	.581
Septicemia	7 (3.6)	4 (2.4)	.499
Risk of VAP Incidence[9,11,12], *n (%)*			
Proton-pump Inhibitors	190 (99.0)	152 (91.6)	.001*
Previous Antibiotic Exposure	183 (95.3)	157 (94.6)	.751
Sedative	148 (77.1)	126 (75.9)	.793
Re-intubation	37 (19.3)	24 (14.5)	.227

*Statistical significance
Abbreviations: BMI = Body mass index, APACHE II = Acute physiology and chronic health evaluation II, ARDS = acute respiratory distress syndrome

using non-disposable ventilator circuits (13.8%) with a VAP incidence rate of 10.68/1,000 ventilator days versus 24 episodes in the group using disposable ventilator circuits (14.4%) with a VAP incidence rate of 10.19/1,000 ventilator days, which showed no statistically significant difference *(p=0.87)*. A survival rate analysis by Kaplan-Meier using the log-rank test found the time to a VAP event of non-disposable ventilator circuits to be 52 days *(95%CI: 35.91-68.09)*, while the time for patients with disposable ventilator circuits was 85 days *(95%CI: 57.37-112.63)*, which had no statistical significance *(p=0.256)*. An analysis of factors with effects on the Hazard ratio according to the Cox Regression Model found disposable ventilator circuits to reduce VAP incidence by 0.72 times *(HR 0.72, 95%CI: 0.407-1.274)*, which similarly had no statistical significance *(p=0.259)*. Other outcome aspects, such as the length of hospitalization, length of stay in the ICU, and number of ventilator days, were also not statistically significantly different. In addition, the outcomes before transfer from the ICU to other wards, discharge, deaths, referrals to other hospitals, and survival at one month after leaving the ICU were also not statistically significantly different (Table 2). According to the culturing results, the main causes of VAP were Carbapenem-resistant *Acinetobacter baumannii* (47%), Carbapenem-susceptible *Pseudomonas aeruginosa* (13.7%), Carbapenem-susceptible *Acinetobacter baumannii* (9.8%) and Gram-negative bacteria (9.8%).

When the U-control chart was used to assess variances in VAP incidence, the data for both the upper control limit (UCL) and the lower control limit (LCL) had normal distributions within a common cause variable, meaning the infections were considered to have been continual without special causes. Therefore, the VAP incidence in patients using non-disposable and disposable ventilator circuits were not statistically significantly different (Fig 1).

In terms of the average cost of the ventilator circuits

TABLE 2. Ventilator-related and Clinical outcomes.

Outcome	Non-disposable *(n = 192)*	Disposable *(n = 166)*	*P*
VAP case, *n (%)*	27 (13.8)	24 (14.4)	.871
Length of Hospitalization *(days)**	32 (16.5–49)	37 (15–59)	.240
Length of Stay in ICU *(days)**	13 (7–23)	14 (8–31)	.121
Ventilator Days *(days)**	8 (4–13)	9 (5–18)	.208
Discharge Status, *n (%)*			
Transfer	111 (57.8)	82 (49.4)	.330
Death	56 (29.2)	53 (31.9)	.330
Discharge	18 (9.4)	24 (14.5)	.330
Transferred Hospital	7 (3.6)	7 (4.2)	.330
Alive at 1 month after leaving the ICU, *n (%)*	104 (55.6)	98 (59.4)	.474

* Median (interquartile range)
Abbreviation: VAP = Ventilator-associated pneumonia

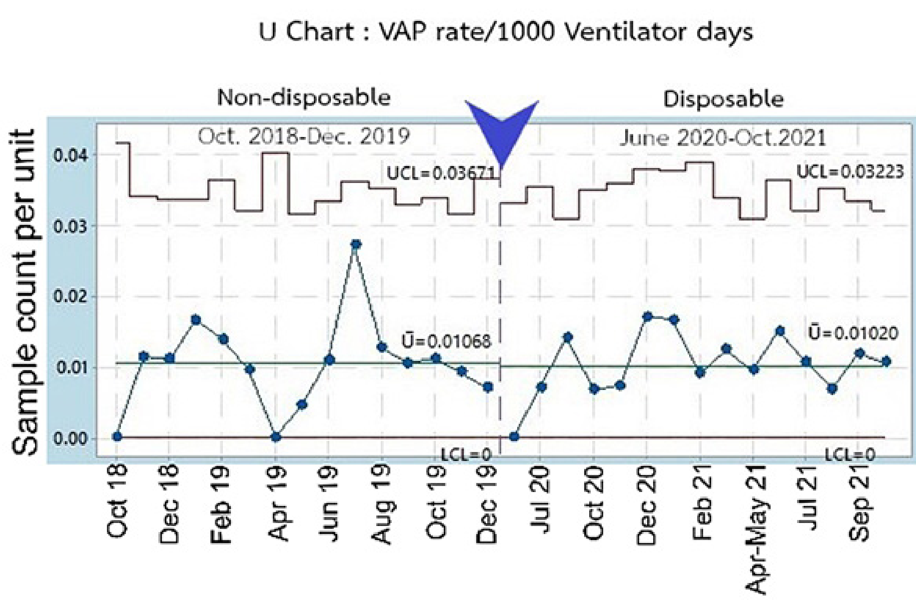


Fig 1. U-control chart of VAP incidence in patients using non-disposable ventilator circuits and patients using disposable ventilator circuits. The arrow indicates the change in patient group.

in one round of change for one month, after calculations based on the activity and time, the non-disposable ventilator circuits were found to have an average cost of THB 1,863.56, while the disposable ventilator circuits were found to have an average cost of THB 2,159.5, meaning the average cost of the non-disposable ventilator circuits was on average THB 295.94 lower (Table 3). In terms of the overall expenses from the hospitalization and antibiotics use in patients who had VAP and patients who did not have VAP, the patients who had VAP had higher overall hospitalization and antibiotic expenses than the patients who did not have VAP *($p < 0.00$)*.

Next, we assessed the satisfaction among the 42 staff members in the hospital work unit toward the ventilator circuits. The work unit personnel consisted of 34 nurses (81%) and 8 nurse assistants (19%) with a mean age *(mean ±SD)* of 37 ± 10 years old, and work experience of <5 years (5 personnel or 11.9%), 5–10 years (14 personnel or 33.3%), or >10 years (23 personnel or 54.8%). In terms of their satisfaction toward ventilator circuits, they expressed a medium level of satisfaction toward non-disposable ventilator circuits in every question item, but a good level of satisfaction toward disposable ventilator circuits in most questions, with only the question on the amount of fluids remaining in the ventilator circuits showing a medium level of satisfaction (Table 4). However, when the work unit personnel were asked for their opinions regarding "if the work unit were to implement a policy to change disposable ventilator circuits", most of the personnel approved (33 personnel or 78.6%) while a minority disapproved (9 personnel or 21.4%), but without a statistically significant difference *($p = 0.168$)*. However, when the opinions of the nurses were separated from the nurse assistants, it was found that 25 nurses (73.5%) approved and 9 nurses disapproved (26.5%) with this change scenario, while every nurse assistant approved (8 personnel or 100%).

TABLE 3. Ventilator circuits related cost per set.

Cost/set (THB)	Non-disposable	Disposable
Cost per set	694.44	2,124
Labor cost	250.69	35.5
Maintenance cost	918.43	0
Total Cost per set	**1,863.56**	**2,159.5**

calculated from the costs in 1 month

TABLE 4. Satisfaction of the personnel with non-disposable and disposable ventilator circuits.

	Level of satisfaction (Mean)		
Question topics	**Non-disposable**	**Disposable**	***P-value***
Convenience in assembling	Medium (2.88)	Good (4.41)	<.000*
Time spent assembling	Medium (2.88)	Good (4.07)	<.000*
Preparatory testing before use	Medium (3.46)	Good (3.98)	.001*
Flexibility during nursing care	Medium (3.50)	Good (3.90)	.061
Circuit weight, tension	Medium (2.98)	Good (3.81)	.005*
Condensate fluids in the ventilator circuits	Medium (3.43)	Medium (3.14)	.279
Removal of the ventilator circuits after use	Medium (2.86)	Good (4.33)	<.000*
Total satisfaction	Medium (3.12)	Good (3.83)	.002*

*Statistically significant

DISCUSSION

The findings indicated that use of non-disposable versus disposable ventilator circuits had no effect on VAP incidence. This was consistent with a study conducted by Srisan among pediatric children.[14] Furthermore, another study found the bacterial contamination of used ventilator circuits to not be correlated with VAP incidence.[13,15] The findings support the evidence-based practice whereby ventilator circuits should not be frequently replaced without appropriate indicators.[16]

The VAP incidence rates in the group using non-disposable ventilator circuits and the group using disposable ventilator circuits were 10.68 and 10.19 per 1,000 ventilator days, respectively. This finding corresponded with the mean VAP incidence rate of the RCU in 2018, which was 10 per 1,000 ventilator days.[8] The findings from this study are also consistent with the findings of most previous studies, which found hospitalization expenses and antibiotic expenses in patients who had VAP to be higher than in patients without VAP.[17,18] Therefore, finding preventive measures is important. VAP prevention guidelines consist of the following three main guidelines: (1) guidelines for preventing and controlling hospital infections, such as providing instructions for staff on infection prevention, proper hand washing, and on the appropriate number of personnel; (2) guidelines for reducing bacterial growth, such as washing hands before contact with patients and keeping the environment clean, proper antibiotic use, the provision of proton-pump inhibitors based on necessity, oral cavity cleaning, In addition to reducing complications from mechanical ventilation, weaning from ventilators as soon as possible also reduces expenses in other areas[19]; and (3) guidelines for reducing aspiration, such as by positioning patients' heads to be at a 30–45 degree angle, controlling the intra-cuff pressure of ETT at 25–30 cmH_2O, suctioning saliva before suctioning in ETT, assessing the amount of gastric content, preventing condensate fluids from entering patients, and not changing the ventilator circuits more than every seven days.[9,12] Many institutions have developed guidelines from evidence-based practices and found the aforementioned practices to be able to reduce the incidence of VAP. The researcher's own agency implemented "guidelines for preventing ventilator-associated pneumonia in adults" from Siriraj Hospital or the WHAP-C Bundle in every patient who was on a ventilator. The guidelines covered the topics: (1) weaning patients off ventilators; (2) hygienic hand washing; (3) aspiration precautions; (4) contamination prevention; and (5) chest physiotherapy. Furthermore, discipline was constantly monitored in terms of compliance with the guidelines. Thus, both the study groups in the present research received care under the same standards. However, because VAP occurs due to multiple causes, including external and personal factors, including accuracy and consistency of compliance with preventive guidelines, VAP remains a major problem requiring further study and correction.[20-22]

According to the present study, the use of disposable ventilator circuits may be unable to reduce VAP incidence and the costs were higher than for non-disposable ventilator circuits. Therefore, in changing from using non-disposable ventilator circuits to using disposable ventilator circuits, treatment agencies must consider cost-efficiency in every area. In the present study, the job descriptions of the assistant nurses included preparing ventilators. From monitoring the work unit personnel, it was found that the ventilator preparation process took 3–4 h. In addition to the risk of personnel infection from contaminated ventilator circuits, during this time, the work unit lost one nurse assistant who provided care for patients. This was evident in the findings reflecting the opinions of the nurse assistants, all of whom approved changing to disposable ventilator circuits. Furthermore, most nursing personnel and nurse assistants were more satisfied with working with disposable ventilator circuits than non-disposable ventilator circuits, with statistical significance, except on the question of flexibility during nursing care and the issue of condensate fluids remaining in the circuits, which were not statistically significantly different. This can be explained by understanding that even though the disposable dual heated-wire circuits had properties that reduced the fluids remaining in the circuits, poor temperature control in patient wards and the nurses connecting unheated flexible lines from the Y-piece of the ventilators before connecting to the ventilator tubes for convenience when working with patients who had a tracheostomy or were in a prone position caused the temperature control efficiency to be poor, leading to more condensate fluids remaining in the circuits. Moreover, the type of circuit used by our patients had no water trap, causing difficulties in pouring water from the ventilator circuit and inflexibility during nursing care, which represented disadvantages of the circuits and change recommendations from the main personnel. However, disposable ventilator circuits are widely used in developed countries and private hospitals due to their convenience, reduced likelihood of infection among personnel, and higher cost tolerance, such as for higher personnel wages, when compared with developing countries or state-owned hospitals.[14] However, recently, Siriraj Hospital started to use disposable ventilator circuits with COVID-19 patients to prevent personnel infections. Many different types of disposable ventilator circuits are available with differences in terms of the price and their properties. The circuits used in the present study were disposable dual heated-wire circuits with the ability to control temperature, provide humidity during inhalation and exhalation, prevent secretion obstructions and condensate fluids collecting in the circuits. However, according to the present study, the average cost of the disposable ventilator circuits used in the study was higher than for the non-disposable ventilator circuits. Changing to other less expensive ventilator circuits with similar properties would reduce the costs of disposable ventilator circuits, which would benefit reducing the work unit's costs without having a negative effect on patients while preventing infections among personnel during ventilator circuit preparation. The RCU of Siriraj Hospital recently changed to the use of disposable dual heated-wire ventilator circuits under the decision of the work unit's leader, who recognizes the benefits, including the cost-efficiency and prevention of infection among personnel, e.g., during the COVID-19 pandemic. However, the work unit is currently procuring disposable ventilator circuits of other types at lower prices with similar efficiency as replacements to achieve higher cost-efficiency in the future.

CONCLUSION

Using non-disposable versus disposable ventilator circuits had no effects on VAP incidence. Although non-disposable ventilator circuits had lower costs than disposable ventilator circuits, most our work unit personnel were more satisfied with disposable circuits than non-disposable circuits.

ACKNOWLEDGMENTS

The research group would like to thank the following people who were involved in helping make this research successful: Miss Sureerat Tamrongsombataku, Value Driven Care Unit (VDCU); Miss Somkid Kamsrimuang, Information Department; Miss Chulaporn Pooniam, Clinical Epidemiology Unit; Miss Neungruthai Torsuwan, Finance Department; and ENDOMED (1999) Co., Ltd. for supplying 150 sets of disposed ventilator circuits. This study was supported by the Siriraj Research Development Fund (Managed by Routine to Research: R2R).

REFERENCES

1. The Centers for Disease Control and Prevention, National Healthcare Safety Network (NHSN). Pneumonia (Ventilator-associated [VAP] and non-ventilatorassociated Pneumonia [PNEU]) Event 2022 19 february 2022.
2. Chastre J, Luyt C-E. 34 - Ventilator-Associated Pneumonia. Philadelphia: W.B. Saunders; 2016. Available from: https://www.sciencedirect.com/science/article/pii/B9781455733835000348.
3. Ladbrook E, Khaw D, Bouchoucha S, Hutchinson A. A systematic scoping review of the cost-impact of ventilator-associated pneumonia (VAP) intervention bundles in intensive care. Am J Infect Control. 2021;49(7):928-36.
4. Klompas M, Branson R, Eichenwald EC, Greene LR, Howell MD, Lee G, et al. Strategies to Prevent Ventilator-Associated Pneumonia in Acute Care Hospitals: 2014 Update. Infect Control Hosp Epidemiol. 2014;35(8):915-36.
5. American Thoracic Society, Infectious Diseases Society of America. Guidelines for the Management of Adults with Hospital-acquired, Ventilator-associated, and Healthcare-associated Pneumonia. Am J Respir Crit Care Med. 2005;171(4): 388-416.
6. Altinsoy S, Catalca S, Sayin MM, Tutuncu EE. The risk factors of Ventilator Associated Pneumonia and relationship with type of tracheostomy. Trends in Anaesthesia and Critical Care. 2020;35:38-43.
7. Patil HV, Patil VC. Incidence, bacteriology, and clinical outcome of ventilator-associated pneumonia at tertiary care hospital. J Nat Sci Biol Med. 2017;8(1):46-55.
8. Centre for Nosocomial Infection Control of Siriraj Hospital. Annual disease surveillance report 2018-2021. 2022.
9. Waters B, Muscedere J. A 2015 Update on Ventilator-Associated Pneumonia: New Insights on Its Prevention, Diagnosis, and Treatment. Curr Infect Dis Rep. 2015; 17(8):496.
10. Goel V, Gupta S, Goel T. Ventilator-associated pneumonia: a review of the clinically relevant challenges in diagnosis and prevention. British Journal of Medical Practitioners. 2016; 9(2):18.
11. Keyt H, Faverio P, Restrepo MI. Prevention of ventilator-associated pneumonia in the intensive care unit: a review of the clinically relevant recent advancements. Indian J Med Res. 2014; 139(6):814-21
12. Michael Klompas. Risk factors and prevention of hospital-acquired and ventilator-associated pneumonia in adults. UpToDate. 2022.
13. Li Y-C, Lin H-L, Liao F-C, Wang S-S, Chang H-C, Hsu H-F, et al. Potential risk for bacterial contamination in conventional reused ventilator systems and disposable closed ventilator-suction systems. PLoS One. 2018;13(3):e0194246.
14. Srisan P, Meechaijaroenying K. Comparison of ventilator-associated pneumonia in children using disposable and nondisposable ventilator circuits. Pediatric Respirology and Critical Care Medicine. 2017; 1(4):77-80.
15. Pen DL, Yan GF, Chen WM, Liu J, Ying JY, Lu GP, et al. The role of bacterial colonization of ventilator circuit in development of ventilator-associated pneumonia: a prospective observational cohort study. Clin Microbiol Infect. 2021;27(3):467.e1-e7.
16. Koenig SM, Truwit JD. Ventilator-associated pneumonia: diagnosis, treatment, and prevention. Clin Microbiol Rev. 2006; 19(4):637-57
17. Rodrigues J, Sousa P. Economic and Clinical Impact of Ventilator-Associated Pneumonia in Intensive Care Units of a University Hospital Center. Cham: Springer International Publishing; 2019.
18. Kollef MH, Hamilton CW, Ernst FR. Economic Impact of Ventilator-Associated Pneumonia in a Large Matched Cohort.

Infect Control Hosp Epidemiol. 2012; 33(3):250-6.
19. Innok S, Dokphueng W, Udol K, Slisatkorn W, Sawasdiwipachai P. Clinical Outcomes and Cost of Ventilator Weaning and Endotracheal Extubation Guided by An Established Ventilator Weaning Protocol in Patients Undergoing Elective Cardiac Surgery. Siriraj Med J. 2021;73(12):815-22.
20. Kheiawsawang M, Sopaporml O, Siripitayakunkit A, Sutherasan Y. The effects of nursing practice guidelines to prevent ventilator associated pneumonia (VAP) in Medical Intensive Care Unit, Ramathibodi Hospital. Mahidol R2R e-Journal. 2020; 7(1).
21. Siriraj_KM. Siriraj Concurrent Trigger Tool by Modified Early Warning Signs (SiCTT by MEWS) 2021 26 Feb 2022. Available from:https://www2.si.mahidol.ac.th/km/cop/sicop/sictt/16798/.
22. Thamcharoentrakul B, Saensom D. Effects of Using Nursing Practice Guideline for Ventilator Associated Pneumonia Prevention: Medical Department. The Journal of Baromarajonani College of Nusing, Nakhonratchasima. 2019; 25(1).

Factors Associated with Cervical Cancer Screening Overuse and Underuse, and Attitude towards Human Papillomavirus Self-sampling among Hospital Staffs

Phaitoon Laowjan*, Kla Maichonklang*, Pronnapa Permpool*, Pattarawalai Talungchit, M.D., Ph.D., Nida Jareemit, M.D.****

**Department of Preventive and Social Medicine, **Department of Obstetrics and Gynecology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*

ABSTRACT

Objective: To investigate the rates of and factors associated with cervical cancer screening overuse and underuse, and the attitude towards human papillomavirus (HPV) self-sampling among hospital staffs.

Materials and Methods: This cross-sectional study included female hospital staff undergoing an annual health check-up. A questionnaire was used to collect sociodemographic characteristics, indication for cervical cancer screening, reasons for screening decision, and attitude toward HPV self-sampling. Screening overuse was defined as having cervical cancer screening without indication while screening underuse was having indication for screening but not receiving it.

Results: Among the 600 included participants, 220 and 380 women decided to screen and not to screen for cervical cancer, respectively. The prevalence of screening overuse and underuse was 11.8% and 33.0%, respectively. Multivariate analysis revealed age was associated with screening underuse (aOR: 1.06, 95% CI: 1.04-1.09; $p<0.001$), whereas married status was associated with screening overuse (aOR: 3.73, 95% CI: 2.05-6.79; $p<0.001$). The common reasons for screening were "add-on to annual health check-up" (93.2%), "fear of cancer" (84.1%), and "family/organizational support" (54.5%). The common reasons for not screening were "feeling healthy" (73.7%), "fear of pain" (58.7%), and "embarrassment" (57.1%). Most women (65.1%) expressed interest in screening via HPV self-sampling.

Conclusion: Overuse and underuse of cervical cancer screening were common. An accurate information regarding screening indication should be provided and indication restriction should be implemented to a health system to avoid screening overuse. In addition, an encouragement should be enhanced to the target population to attend the screening program.

Keywords: Attitude; human papillomavirus cervical self-sampling; cervical cancer screening overuse and underuse; female staff members (Siriraj Med J 2023; 75: 200-207)

INTRODUCTION

Cervical cancer is the fourth most common cancer and the leading cause of cancer-related death among women worldwide.[1] In Thailand, the age-standardized incidence (16.4%) and mortality (7.4%) rates are greater than the global rates. The global age-standardized incidence

Corresponding author: Nida Jareemit
E-mail: nida.jareemit@gmail.com
Received 12 November 2022 Revised 1 January 2023 Accepted 4 January 2023
ORCID ID:http://orcid.org/0000-0003-1913-6613
https://doi.org/10.33192/smj.v75i3.260868

and mortality rates were reported to be 13.3% and 7.3%, respectively.[2] Cervical cancer screening is one of the most effective types of cancer screening as evidenced by a substantial decrease in the incidence and mortality rates of cervical cancer during the past few decades.[3] Although increased, cervical cancer screening coverage in Thailand increased from 46.3% in 2007 to only 59.7% in 2009.[4] Low screening coverage is commonly found in less-educated, poor, young, unmarried, and non-Buddhist women.[5]

Women that work in a hospital might be expected to have good knowledge about cervical cancer and good cervical cancer screening practices. These women are viewed as a reliable source of high-quality healthcare information, and as a role model that has the ability to inspire and encourage their friends and families to adopt and demonstrate good healthcare practices. This hypothesis, if it were found to be true, would results in increased cervical cancer screening coverage in general population. However and surprisingly, prior studies reported that only 20.7% of nurses have ever undergone cervical cancer screening.[6] In Siriraj Hospital, a university-based national tertiary referral center, cytology-based cervical cancer screening is provided free of charge; however, the proportion of female hospital staff members that agree to undergo screening remains low. Our 5-year data showed that only 2.7-6.5% of all female hospital staff attended the center's cervical cancer screening program. On the contrary, some women who repeat screening every year instead of every 2 years according to cervical cancer screening recommendation were observed.[7] Yet, the data regarding screening overuse has never been collected and published. To better understand the cervical cancer screening attitudes and habits of female staff at our center, the aim of this study was to investigate the rates of, and factors associated with cervical cancer screening overuse and underuse, and the attitude towards human papillomavirus (HPV) cervical self-sampling among female hospital staff members at a single-center, university hospital.

MATERIALS AND METHODS

Study design and study participants

This cross-sectional, questionnaire-based study was conducted at a single center, tertiary university hospital. Female hospital staff members who attended their annual health check-up during January 2020 to December 2020 study period were invited to participate. Those unwilling to join the study were excluded. The protocol for this study was approved by local Institutional Review Board (COA no. Si 616/2019), and written informed consent was obtained from all study volunteers. The study protocol complied in all ways with the principles set forth in the 1964 Declaration of Helsinki and all its subsequent amendments.

Study questionnaire

The study questionnaire was created by the authors, after which it was reviewed and approved by local experts in gynecologic oncology. The questionnaire comprised the 4 following parts: 1) sociodemographic characteristics, including age, marital status, education level, occupation, income, HPV vaccination status, and family history of cancer; 2) indication for cervical cancer screening, including age at first sexual intercourse and history of cervical cancer screening; 3) risk factors for or symptoms of cervical cancer, including first sexual intercourse at age ≤18 years, having given birth to three or more children, having multiple sexual partners, using oral combined contraceptive pills for longer than 5 years, cigarette smoke exposure, immunosuppressed status, and having abnormal vaginal bleeding or discharge; 4) decision to undergo or not undergo cervical cancer screening, and the reason(s) for their decision; and, 5) attitude towards HPV self-sampling for cervical cancer screening. Following the Royal Thai College of Obstetricians and Gynaecologists (RTCOG) guideline, indication for cervical cancer screening in this study was satisfied if all the following criteria was met: 1) age 25 to <30 years and currently sexually active or age 30 years or older regardless of sexual activity, and 2) not having cervical cancer screening within the past 2 years.[7] Screening underuse was defined as having the aforementioned indication for cervical cancer screening, but the screening had not been performed. Screening overused was defined as not having the indication for cervical cancer screening but the screening was performed.

Study process

After the purpose of the study was fully explained, instructions were given, and written informed consent was obtained to join the study, the study volunteer was given the questionnaire to be completed. Any questions regarding the questionnaire were addressed until the study volunteers were clearly understood. Then, the study participants were asked if they wanted to undergo cervical cancer screening. Those women were further subdivided into the declined cervical cancer screening or decided for cervical cancer screening groups. Among the women scheduled for screening, data specific to who did and did not attend the scheduled screening appointment were collected. Reasons for declining screening and not attending the scheduled screening visit were collected and analyzed. Screening result data were also collected and analyzed.

Statistical analysis

Descriptive statistics were used to summarize study participant data. Chi-square test or Fisher's exact test was performed to compare differences in categorical variables, and those results are reported as frequency and percentage. Normally distributed continuous variables were compared using Student's *t*-test, and the findings are reported as mean plus or minus standard deviation. Non-normally distributed continuous data were compared using Mann-Whitney U test, and those results are given as median and interquartile range. Univariate and multivariate binary regression analysis was used to identify factors independently associated with the underuse or overuse of cervical cancer screening services. A p-value <0.05 was considered statistically significant for all tests. Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.

RESULTS

A total of 600 female hospital staff members were enrolled in this study. There were 220 women who decided to attend cervical cancer screening, and 380 women who decided not to do so (Fig 1). Of the 220 women who were scheduled for cervical cancer screening, 125 (56.8%) women showed up for screening at their appointed screening date and time. Taking the indication for cervical cancer screening into account, screening underuse was 33% while screening overuse was 11.8%. Among women having screening overuse, 21 (29.6%) women were aged less than the screening recommendation, 54 (76%) women had negative screening within 2 years, and 30 (42.3%) women had both.

Baseline sociodemographic and clinical characteristics of all study participants and compared between those who decided to and not to screen for cervical cancer are shown in Table 1. The comparative analysis between those two groups revealed a significant difference relative to age, marital status, occupation, and risk factors for cervical cancer.

The reasons given for screening or not screening are given in Fig 2. Among those who decided to screen, the top three reasons were "add-on of cervical screening to annual health check-up program" (91.9-93.2%), "fear of cancer" (79.1-86.3%) and "family/organizational support"(54.1-58.9%) in all women, women without indication, and women without risk. For women who decided not to screen, the top three reasons for not doing so were "screening is not necessary because I am healthy" (73.7-74.4%), "fear of pain" (52.7-61.5%), and "embarrassment" (52.1-68.6%) in all women, women without indication, and women without risk

Table 2 and Table 3 shows the factors significantly and independently associated with screening overuse and screening underuse, respectively. Univariate analysis showed

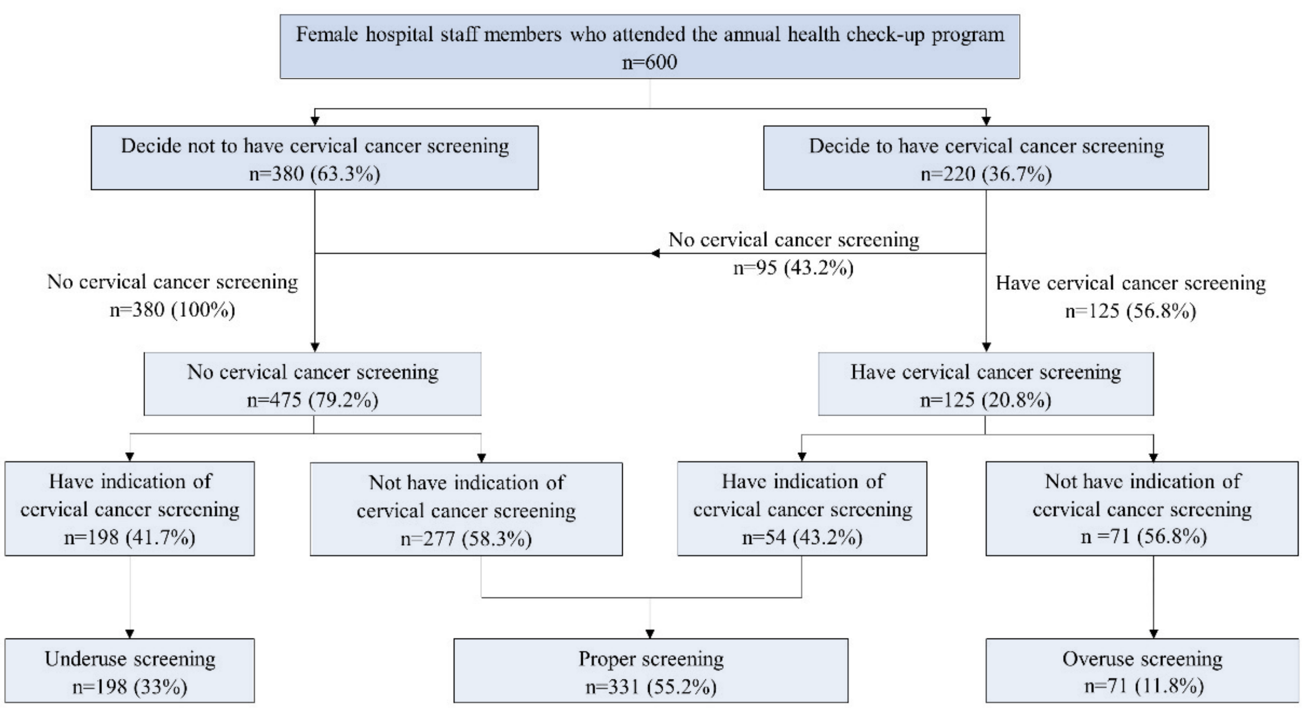


Fig 1. The flow of female staff members through the study.

TABLE 1. Baseline sociodemographic and clinical characteristics of all study participants, and compared between those who decided to and not to screen for cervical cancer.

Characteristics	All (N=600)	Decided not to screen (n=380)	Decided to screen (n=220)	*P*-value
Age (years)	34.5±9.3	33±9.2	37.2±9.0	**<0.001**
Marital status				**<0.001**
Single	393 (65.5%)	295 (77.6%)	98 (44.5%)	
Married	181 (30.2%)	69 (18.2%)	112 (51.0%)	
Widowed/divorced/separated	26 (4.3%)	16 (4.2%)	10 (4.5%)	
Level of education				0.569
Secondary school	20 (3.3%)	11 (2.9%)	9 (4.1%)	
High school	96 (16.0%)	58 (15.3%)	38 (17.3%)	
Bachelor's degree or higher	484 (80.7%)	311 (81.8%)	173 (78.6%)	
Occupation				**<0.001**
Doctor, dentist, nurse, or health scientist	393 (65.5%)	269 (70.8%)	124 (56.4%)	
Back office or others	207 (34.5%)	111 (29.2%)	96 (43.6%)	
Income (US dollar)				0.482
≤ 600	111 (18.5%)	68 (17.9%)	43 (19.5%)	
> 600-900	300 (50.0%)	198 (52.1%)	102 (46.4%)	
> 900	189 (31.5%)	114 (30.0%)	75 (34.1%)	
Had HPV vaccination	62 (10.3%)	40 (10.5%)	22 (10.0%)	0.840
Family history of cancer	163 (27.2%)	106 (27.9%)	57 (25.9%)	0.600
Had risk for cervical cancer	280 (46.7%)	146 (38.4%)	134 (60.9%)	**<0.001**
Median number of cervical cancer risks	0 (0, 1)	0 (0, 1)	1 (0, 2)	**<0.001**
Had indication for cervical cancer screening	252 (42.0%)	156 (41.1%)	96 (43.6%)	0.537

Data are presented as mean plus/minus standard deviation, number and percentage, or median (P25, P75)
A *p*-value<0.05 indicates statistical significance
Abbreviation: HPV, human papillomavirus

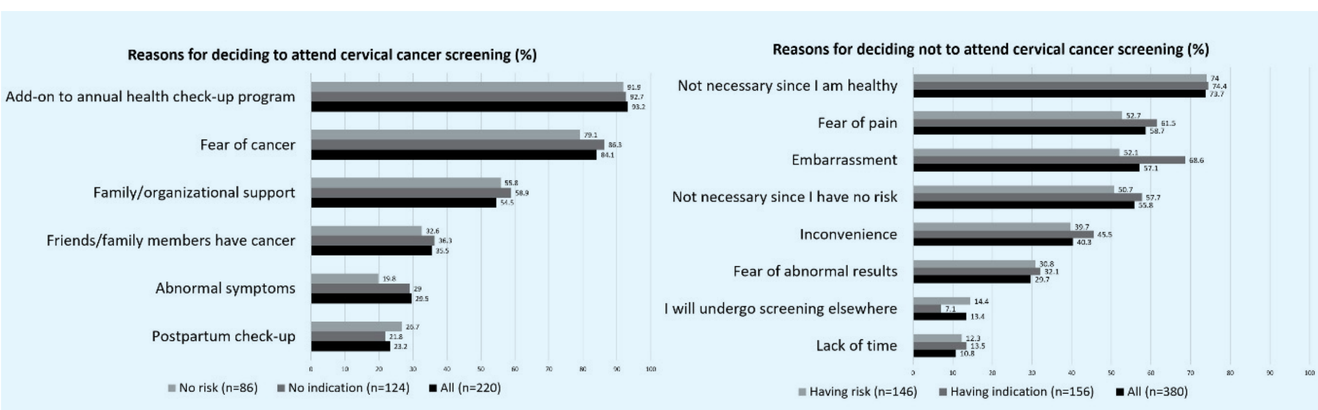


Fig 2. Reasons given for attending or not attending cervical cancer screening.

TABLE 2. Univariate and multivariate analysis to identify factors independently associated with overuse of cervical cancer screening.

Factors	N (%)	OR (95%CI)	*P*	Adjusted OR (95%CI)	*P*
Age		1.03 (1.01-1.06)	**0.011**		
Marital status					
Single	29 (40.8)	1		1	
Married	39 (54.9)	3.45 (2.05-5.79)	***<0.001***	3.73 (2.05-6.79)	***<0.001***
Widowed/divorced/separated	3 (4.2)	1.64 (0.46-5.78)	**0.444**	1.77 (0.47-6.69)	0.402
Level of education					
Secondary school	2 (2.8)	1			
High school	8 (11.3)	0.82 (0.16-4.18)	0.809		
Bachelor's degree or higher	61 (85.9)	1.30 (0.29-5.73)	0.731		
Occupation					
Doctor, dentist, nurse, and scientist	37 (52.1)	1			
Back office or others	34 (47.9)	1.89 (1.15-3.12)	***0.012***		
Income (US dollar)					
≤ 600	12 (16.9)	1			
> 600-900	40 (56.3)	1.27 (0.64-2.52)	0.538		
> 900	19 (26.8)	0.92 (0.43-1.98)	0.835		
Had HPV vaccination	10 (14.1)	1.50 (0.73-3.11)	0.272		
Family history of cancer	15 (21.1)	0.69 (0.38-1.26)	0.225		
Number of risk factors for cervical cancer		1.15 (0.90-1.47)	0.256		

TABLE 3. Univariate and multivariate analysis to identify factors independently associated with underuse of cervical cancer screening.

Factors	N (%)	OR (95%CI)	*P*	Adjusted OR (95%CI)	*P*
Age		1.06 (1.04-1.08)	***<0.001***	1.06 (1.04-1.09)	***<0.001***
Marital status					
Single	118 (59.6)	1			
Married	64 (32.3)	1.28 (0.88-1.85)	0.202		
Widowed/divorced/separated	16 (8.1)	3.73 (1.64-8.46)	***0.002***		
Level of education					
Secondary school	8 (4)	1			
High school	30 (15.2)	0.68 (0.25-1.84)	0.450		
Bachelor's degree or higher	160 (80.8)	0.74 (0.30-1.85)	0.520		
Occupation					
Doctor, dentist, nurse, and scientist	132 (66.7)	1			
Back office or others	66 (33.3)	0.93 (0.65-1.33)	0.673		
Income (US dollar)					
≤ 600	35 (17.7)	1			
> 600-900	78 (39.4)	0.76 (0.47-1.23)	0.266		
> 900	85 (42.9)	1.78 (1.09-2.90)	***0.022***		
Had HPV vaccination	13 (6.6)	0.51 (0.27-0.96)	0.036		
Family history of cancer	59 (29.8)	1.22 (0.83-1.77)	0.310		
Number of risk factors for cervical cancer		0.95 (0.79-1.14)	0.580		

age, married status, and back-office staff or others to be significantly associated with screening overuse. That same analysis demonstrated age, widowed/divorced/separated status, and income > 900 US dollar to be significantly associated with screening underuse. Multivariate analysis revealed age (aOR: 1.06, 95%CI: 1.04-1.09; $p<0.001$) to be independent predictors of screening underuse, whereas married status was found to associate with screening overuse (aOR: 3.73, 95%CI: 2.05-6.79; $p<0.001$).

Regarding the attitude of the hospital staff towards HPV self-sampling for cervical cancer screening, 391 of 600 women (65.1%) reported being interested in this screening method. The percentages of women interested in HPV self-sampling were similar between women who decided to screen (64.5%) and women who decided not to screen (65.5%).

Of 125 women undergoing cervical cancer screening, 2 (1.6%) women had abnormal cervical cytology. Yet, no cervical intraepithelial neoplasia or more severe lesions presented in this cohort. Among those who did not show up for their scheduled screening appointment, the most common reasons given were inconvenience (57.9%), having menstrual period (15.8%), and plan to seek or already had cervical screening elsewhere (6.3%).

DISCUSSION

Screening overuse

In the present study, 71 (11.8%) women decided to undergo cervical cancer screening despite the fact that they were not clinically indicated to do so. Among this group, 76% had negative screening within 2 years, 29.6% were aged less than the screening recommendation, and 42.3% women had both.

In this cohort, the screening without indication was found the most in women who underwent screening even though they had negative screening within 2 years. Prior studies from the United States also reported rates of screening overuse ranging from 45-65%.[8-10] A large-scale population-based study that was conducted in the US reported that up to two-thirds of women had repeated screening within three years of their index test.[10] The cumulative incidence of repeat cervical cancer screening was 17.7% (95% CI: 17.6-17.7%) at 12 months, 51.1% (95% CI: 51.0-51.2%) at 24 months, and 65.8% (65.7-65.8%) at 36 months. According to US guideline, cytology-based screening should be performed every 3 years. A multicenter European study reported a 0.28% risk of cervical intraepithelial neoplasia (CIN) 3 or cancer after negative cervical cytology.[11] A shorter duration screening interval yielded a negligible reduction in cancer risk, but it substantially increased unnecessary procedures and treatments, which resulted in significantly increased costs.[12,13] The number of colposcopies would be decreased by 50% if cervical cancer screening was performed every 3 years instead of every year.[14] In Thailand, cervical cytology is the most commonly used technique for cervical cancer screening. Due to the higher incidence of cervical cancer in Thailand, a 2 year-screening interval is recommended by Thai national guideline.[7,15] Considering the benefits and risks, annually screening is not recommended.[15]

Approximately one-third of women in this study were aged less than the screening recommendation. Prior studies revealed a low burden of cervical cancer in women aged less than 25 years. More specifically, the rates of cervical cancer-related incidence and mortality in this young age group was reported to be 0.8% and 0.5%, respectively.[16] In some contrast, the highest prevalence of HPV infection was found among younger aged women. Nevertheless, the rates of both persistent infection and cancer progression were lower in the younger age group compared to the older age group.[17-19] For the reasons, the American Cancer Society and the Royal Thai College of Obstetricians and Gynaecologists both recommend that cervical cancer screening should start at 25 years of age.[7,16]

Screening overuse was less common in our study compared to prior studies from the US.[8-10] This may be due to the fact that the present study enrolled hospital staff members who might be more aware of screening indication compared to the general populations enrolled in the US studies. The main reasons for cervical screening without screening indication in our study were add-on of cervical screening to the annual health check-up program, fear of cancer, and family/organizational support. Our study women who lived with their spouses were more likely to participate in cervical cancer screening. In contrast, a US study reported younger age, screening with cytology alone, more medical visits, contraceptive management visits, and gynecology provider specialty to be factors associated with screening overuse.[9] Those findings correspond other studies that found physician-related factors also to be related to screening overuse.[20-22] Since the cervical cancer screening guideline has changed overtime, updated information regarding screening interval and indication should be provided to healthcare providers, and physicians should be aware of the drawbacks of overscreening. Specific indication, as well as the risks and benefits of screening should be discussed with each patient. Interventions, such as reimbursement only in indicated cases, might decrease screening overuse.

Screening underuse

Screening underuse was found in 33% of women in this cohort. Prior studies reported that 20.7-53.0% of nurses had previously undergone cervical cancer screening.[6,23,24] Common reasons given for not undergoing screening included embarrassment and fear of pain. To overcome these concerns, HPV cervical sampling for cervical cancer screening may be a more attractive screening alternative for these women. Initially, HPV self-sampling was developed to increase screening coverage in rural or remote areas. Various methods were used for specimen collection, including brushes, swabs, vaginal patches, and lavage.[25] Regarding test accuracy, a meta-analysis reported the sensitivity and specificity of HPV self-sampling for detecting high-grade cervical lesion or cancer to be slightly lower than the sensitivity and specificity of clinician-collected sampling (ratio = 0.88, 95% CI 0.85-0.91: and 0.96, 95% CI 0.95-0.97, respectively).[26] Previous studies found that HPV self-sampling increased cervical cancer screening coverage to a level comparable to that observed in high-income countries.[25] The reported advantages of HPV self-sampling were privacy, convenience, less embarrassment and anxiety, user-friendly, and less discomfort and pain compared to clinician-collected specimen.[27-29] HPV self-sampling appears to be a potentially viable screening alternative in our setting because 65.1% of our study women expressed positive interest in HPV self-sampling for cervical cancer screening. However, some concerns about HPV self-sampling have been reported. For example, in low- and middle-income countries, the higher price of an HPV self-sampling test would be a financial barrier for many women. More study is needed to illuminate the benefits and drawbacks of HPV self-sampling for cervical cancer screening compared to traditional screening via clinician-collected sampling.

This study found that screening underuse was associated with age which was in accordance with the prior study.[9] To improve screening coverage among those who don't undergo screening often enough, and in those who have never undergone screening, interventions, such as a risk or indication assessment checklist, might be used to ensure that individual women are aware of their own screening indication and/or cancer risks, if any exist.

Strengths and limitations

In addition to providing cervical cancer screening overuse and underuse data, our study also explored the attitude of our study women towards the use of HPV self-sampling cervical cancer screening among our center's female staff. Concerning potential limitations, our data were collected among medical staff, so our data may not necessarily be applicable in non-medical settings. Lastly and given the sensitivity of some of our survey questions, it is possible that some women were reluctant to provide accurate data due to concerns about divulging such personal information.

CONCLUSION

Screening overuse and underuse were both found to be common among female hospital staff. An accurate information regarding screening indication should be provided and indication restriction should be implemented to a health system to avoid screening overuse. For women having screening indication, interventions such as a risk or indication assessment checklist should be implemented to enhance adherence to cervical cancer screening recommendation. HPV self-sampling for cervical cancer screening is a private and attractive screening technique that needs to be approved for use to improve a screening coverage.

Statements and declarations

Funding disclosure

This study was funded by a grant from the Routine to Research Unit of the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (grant no. [IO] R016235028).

Conflict of interest statement

All authors declare no conflicts of interest.

Authors' contributions

Phaitoon Laowjan: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – review and editing, Supervision, Project administration; Kla Maichonklang: Methodology, Investigation, Resources; Pronnapa Permpool: Methodology, Investigation, Resources; Pattarawalai Talungchit: Conceptualization, Methodology, Formal analysis, Writing – review and editing, Supervision; Nida Jareemit: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review and editing

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Assoc. Prof. Winai Ratanasuwan, the Chair of the Department of Preventive and Social Medicine and Mrs. Dujpratana Pisalsarakij for supporting the project, all staff members from the Routine to Research Unit for ever encouraging and motivating guidance, Dr. Orawan Supapueng for assistance with statistical analysis, Kevin P. Jones for language editing and proofreading, and our families for unconditional and loving support.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.
2. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Pineros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021;149(4):778-89.
3. Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F. Worldwide trends in cervical cancer incidence: impact of screening against changes in disease risk factors. Eur J Cancer. 2013;49(15):3262-73.
4. Mukem S, Meng Q, Sriplung H, Tangcharoensathien V. Low Coverage and Disparities of Breast and Cervical Cancer Screening in Thai Women: Analysis of National Representative Household Surveys. Asian Pac J Cancer Prev. 2015;16(18):8541-51.
5. Niu L, Virani S, Bilheem S, Sriplung H. The effect of Pap smear screening on cervical cancer stage among southern Thai women. Sci Rep. 2019;9(1):16921.
6. Ifemelumma CC, Anikwe CC, Okorochukwu BC, Onu FA, Obuna JA, Ejikeme BN, et al. Cervical Cancer Screening: Assessment of Perception and Utilization of Services among Health Workers in Low Resource Setting. Int J Reprod Med. 2019;2019:6505482.
7. Chittithaworn S, Charakorn C, Kongsawatvorakul C. Cervical Cancer Screening Guidelines: An updated review. Thai Journal of Obstetrics and Gynaecology. 2021;29:186-90.
8. Parekh N, Donohue JM, Men A, Corbelli J, Jarlenski M. Cervical Cancer Screening Guideline Adherence Before and After Guideline Changes in Pennsylvania Medicaid. Obstet Gynecol. 2017;129(1):66-75.
9. Almeida CM, Rodriguez MA, Skootsky S, Pregler J, Steers N, Wenger NS. Cervical cancer screening overuse and underuse: patient and physician factors. Am J Manag Care. 2013;19(6):482-9.
10. Wright JD, Chen L, Tergas AI, Melamed A, St Clair CM, Hou JY, et al. Overuse of Cervical Cancer Screening Tests Among Women With Average Risk in the United States From 2013 to 2014. JAMA Netw Open. 2021;4(4):e218373.
11. Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. BMJ. 2008;337:a1754.
12. Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. Br J Cancer. 2003;89(1):88-93.
13. Kulasingam SL, Havrilesky L, Ghebre R, Myers ER. Screening for Cervical Cancer: A Decision Analysis for the US Preventive Services Task Force. U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Rockville (MD); 2011.
14. Stout NK, Goldhaber-Fiebert JD, Ortendahl JD, Goldie SJ. Trade-offs in cervical cancer prevention: balancing benefits and risks. Arch Intern Med. 2008;168(17):1881-9.
15. Practice Bulletin No. 168: Cervical Cancer Screening and Prevention. Obstet Gynecol. 2016;128(4):e111-e30.
16. Fontham ETH, Wolf AMD, Church TR, Etzioni R, Flowers CR, Herzig A, et al. Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society. CA Cancer J Clin. 2020;70(5):321-46.
17. Ronco G, Dillner J, Elfstrom KM, Tunesi S, Snijders PJ, Arbyn M, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. Lancet. 2014;383(9916):524-32.
18. Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. Lancet Oncol. 2010;11(3):249-57.
19. Moscicki AB, Shiboski S, Hills NK, Powell KJ, Jay N, Hanson EN, et al. Regression of low-grade squamous intra-epithelial lesions in young women. Lancet. 2004;364(9446):1678-83.
20. Darwish-Yassine M, Garvin AD, Johnston CM, Zoschnick L, Conners A, Laing S, et al. An assessment of gynecological cytology screening practices among health care providers nationwide. Arch Pathol Lab Med. 2015;139(5):650-5.
21. Gerend MA, Shepherd MA, Kaltz EA, Davis WJ, Shepherd JE. Understanding women's hesitancy to undergo less frequent cervical cancer screening. Prev Med. 2017;95:96-102.
22. Rendle KA, Schiffman M, Cheung LC, Kinney WK, Fetterman B, Poitras NE, et al. Adherence patterns to extended cervical screening intervals in women undergoing human papillomavirus (HPV) and cytology cotesting. Prev Med. 2018;109:44-50.
23. Singh E, Seth S, Rani V, Srivastava DK. Awareness of cervical cancer screening among nursing staff in a tertiary institution of rural India. J Gynecol Oncol. 2012;23(3):141-6.
24. Srisubat A, Limwattanayingyong A, Chaisewikul R, Anannamcharoen S. Factors Affecting the Utilization of Pap Smear Screening Program among Nurses in the Department of Medicine, Siriraj Hospital. Siriraj Med J. 2006;58(8):1002-5.
25. Yeh PT, Kennedy CE, de Vuyst H, Narasimhan M. Self-sampling for human papillomavirus (HPV) testing: a systematic review and meta-analysis. BMJ Glob Health. 2019;4(3):e001351.
26. Tota JE, Bentley J, Blake J, Coutlee F, Duggan MA, Ferenczy A, et al. Introduction of molecular HPV testing as the primary technology in cervical cancer screening: Acting on evidence to change the current paradigm. Prev Med. 2017;98:5-14.
27. Madzima TR, Vahabi M, Lofters A. Emerging role of HPV self-sampling in cervical cancer screening for hard-to-reach women: Focused literature review. Can Fam Physician. 2017; 63(8):597-601.
28. Bais AG, van Kemenade FJ, Berkhof J, Verheijen RH, Snijders PJ, Voorhorst F, et al. Human papillomavirus testing on self-sampled cervicovaginal brushes: an effective alternative to protect nonresponders in cervical screening programs. Int J Cancer. 2007;120(7):1505-10.
29. Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, et al. HPV screening for cervical cancer in rural India. N Engl J Med. 2009;360(14):1385-94.

Prevalence of Metabolic Syndrome and Its Associated Factors among Thai Police Officers – A Population-based Study

Mahesh Gurung, M.D.*, Runyawan Chotenimitkhun, M.D., Kasem Ratanasumawong, M.D.**, Bundit Prommete, M.D.**, Wichai Aekplakorn, M.D.*****

**Cardiology Division, Police General Hospital, Thailand, and Jigme Dorji Wangchuck National Referral Hospital, Thimphu, Bhutan, **Cardiology Division, Police General Hospital, Bangkok, Thailand, ***Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.*

ABSTRACT

Objective: The prevalence of Metabolic Syndrome (MetS) in Thai police officers is unknown. This study aims to accurately determine the prevalence of MetS in this population compared to the Thai general population.

Materials and Methods: We conducted a population-based cross-sectional study of 107,933 Thai police officers during the 2012 annual Police Health Care Center check-up. Metabolic syndrome was defined using the harmonized criteria of six international expert groups.

Results: The prevalence of MetS among Thai police officers was 39.24%, with a higher prevalence among males compared to female police officers (40.9% male and 14.3% female). The prevalence was higher in male police officers (40.65% versus 23.80%, P<0.001), whereas it was lower in female officers (16.30% versus 31.59%, P<0.001) when compared to the general population. High blood pressure was the most common abnormal metabolic component in both genders (male 67.4 % female 33.3%). Whereas observation of low high-density lipoprotein cholesterol (HDL) in males (11.6%) and high fasting plasma glucose (FG) in females (16.3%). The adjusted odds ratio for MetS increased with age, wide pulse pressure, male gender, lower rank, alcohol drinking, and being in a non-metropolitan city.

Conclusion: The prevalence of MetS is higher among Thai police officers compared to the general Thai population. These findings underscore the need for effective preventive measures and continuous monitoring to reduce the risk and burden of cardiovascular diseases.

Keywords: Metabolic syndrome; prevalence; police officers; Thai (Siriraj Med J 2023; 75: 208-217)

INTRODUCTION

Metabolic Syndrome (MetS) is associated with increased risk of atherosclerotic cardiovascular disease (ASCVD) (2-fold) and all-cause mortality (1.5 fold).[1] MetS is emerging as a dominant public health concern due to its relationship to cardiovascular disease and the dramatic increase in its global prevalence. MetS can be considered a pandemic as it variably ranges from 11.9 % to 49 % in Pacific Asia and 24-34 % in the United States and Europe.[2-5] In Thailand, the Thai National Health Examination Survey (TNHES) 2009 revealed that approximately one-quarter of adults of age more than 19 years had MetS. MetS was observed with higher prevalence in urban versus rural males and the reverse in females.[6,7] Cardiovascular risk factors are higher in Thailand, with an estimated 5.1 million individuals with

Corresponding author: Mahesh Gurung
E-mail: dr.mgurung@gmail.com
Received 8 December 2022 Revised 2 January 2023 Accepted 10 January 2023
ORCID ID:http://orcid.org/0000-0003-4279-0859
https://doi.org/10.33192/smj.v75i3.260869

hypertension, 4.4 million with elevated cholesterol, 8.9 million classified as overweight and obese, 2.4 million with diabetes, and 6.2 million current smokers.[8] Law enforcement is a unique occupation due to the demands of irregular work hours, work-life imbalance, higher rates of tobacco and alcohol use, and disturbances in the sleep cycle leading to continual physical and mental stress. This type of lifestyle is often associated with MetS.[9-14]

Prior global epidemiological studies demonstrated an inconsistent prevalence of MetS in police officers compared to other armed forces or the general population.[15-19] A few studies in Thailand have shown an increasing prevalence of overweight or obesity in Thai army personnel.[20,21] However, no reports on the prevalence of MetS and its components in Thai Police officers are available.

This study aims to update the prevalence of MetS and evaluate factors associated with this condition in Thai police officers. The results will hopefully guide future preventive strategies and may, in turn, help prevent future cardiovascular disease in this unique population.

MATERIALS AND METHODS

Study Population

The Police Health Check Up Program 1 was the first large-scale non-communicable disease screening program conducted in Thailand among police officers throughout the country in 2012. There were approximately 213,664 active-duty Thai police officers working in 76 provinces during the year of study. Of this amount, 123,400 officers (60%) had completed the 2012 Police Health Check Up Program conducted by the Police Health Care Center at Police General Hospital, Bangkok.[22] After excluding those subjects with incomplete data, the remaining 107,933 officers comprised the study population.

Methods

Before the initiation of the study, the Ethical Review Committee, Police General Hospital, Bangkok, approved the study design. Data regarding alcohol consumption, cigarette smoking, prior diagnosis of diabetes, hypertension, use of glucose-lowering medications, anti-hypertensive drug use, and lipid-lowering drugs were obtained through a written questionnaire. Each subject had anthropometric measurements such as waist circumference, height, body weight, along with blood pressure (BP), and pulse measurement obtained by a nurse or medical assistant at the Police Health Care center. We measured waist circumference (WC) from the horizontal plane at a level halfway between the iliac crest and the lower costal margin.[23] Brachial artery blood pressure and pulse rate were measured using an automatic machine after participants rested for at least 5 minutes in a sitting position according to standard procedures.[24] A limited physical examination, including assessment of the head, ear, eye, nose, and throat (HEENT), auscultation of the heart, lungs, and palpation of the abdomen, was performed by the supervising physician. Fasting venous blood samples for glucose (FG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-c), and triglyceride (TG) were analyzed using an ADVIA 1800 Clinical Chemistry System. The Friedewald formula was used to calculate Low-density lipoprotein cholesterol (LDL-c).[25]

All laboratory techniques were standardized to Center for Disease Control reference methods.[26]

Definitions

We defined metabolic syndrome according to the definition of the joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention, the National Heart, Lung and Blood Institute, the American Heart Association, the World Heart Federation, the International Atherosclerosis Society, and the International Association for the Study of Obesity.[27] The diagnosis required three or more of the following five components: 1) FG ≥ 100 mg/dl (≥ 5.6 mmol/L) or having diabetes; 2) TG ≥ 150 mg/dl (≥ 1.69 mmol/L) or taking lipid-lowering medication; 3) HDL-c < 40 mg/al (< 1.04 mmol/L) in men or < 50 mg/dl (< 1.29 mmol/L) in women; 4) BP ≥ 130/85 mm Hg or treated hypertension; 5) WC ≥ 90 cm in men and ≥ 80 cm in women (reflecting the Asian population) or a body mass index ≥ 25 kg/m^2.[28,29] Pre-metabolic syndrome (pre-MetS) was defined as any two of the five MetS components. Alcohol drinking was defined as any self-reported regular consumption of any alcoholic beverage. The definition of pulse pressure (PP) was the difference between systolic and diastolic blood pressures during one cardiac cycle. Participants were categorized into groups based on PP as follows: < 35, 35-45; 46-55 and > 55 mmHg.

Statistical analysis

The prevalence and age-adjusted mean prevalence of MetS were calculated based on the direct method using standard techniques and compared to the 2010 general Thai population.[30] Participants were categorized into age - groups based on the above comparison population as follows: 35-39; 40-44; 45-49; 50-54; and 55-59 years of age. Differences between groups were calculated using the t-test for continuous data and the chi-square for categorical data. The chi-square test was used to calculate differences in the distribution of covariates for police officers in Bangkok (metropolitan) and non-Bangkok (non-metropolitan). Multivariate logistic regression was used to examine the association of demographic variables (age, sex, rank, province, pulse pressure, smoking, and alcohol use) and MetS. Stata software version 11 was used for all calculations.

RESULTS

A total of 107,933 police officers (101,846 males and 6,087 females) constituted the entire study population with mean age (SD) of 43.60 (0.08) years (range, 35 – 59 years). The reported gender baseline characteristics of the study population are in Table 1. Highest age adjusted prevalence of MetS by regions in Thailand was observed in the western region (57%) followed by central region (49%) and lowest was observed in Bangkok (30%) (Fig 1). Lower-rank police officers were predominant (79.5 % Vs. 20.5 %) and were posted in non-metropolitan cities. Compared to female police officers, a higher percentage of male police officers were overweight (39% Vs. 21 %), current smokers (46.7 % Vs. 0.3%), and alcohol drinkers (46.7 % Vs. 6.6 %).

The overall age-adjusted prevalence of metabolic syndrome among active-duty police officers aged 35 to 59 years old was 39.4 % (40.9% male and 14.3 % female) (Table 2) which in comparison to the Thai general population (27.8 %)[6], was significantly high (Fig 2). However, on the gender-based comparison, the prevalence of metabolic syndrome was significantly higher in male police officers compared to males in the general population (40.65% Vs. 23.80%, p<0.001) but significantly lower in female officers compared to females in the general population for all age groups (16.30% Vs. 31.59%, p<0.001) (Fig 2).

Prevalence of MetS increased by age in both genders (male, 32.7%, female, 8.9% in the age group of 35-39 years to male 50.9% and female 23.9% in age group of 55-59 years). The prevalence of MetS in police officers increased with age to a greater degree due to its prevalence among male police officers compared to the general population. The highest common abnormal metabolic component was high BP, whereas the lowest common was low HDL for both genders (Table 2). The prevalence of low HDL and high WC was significantly higher among females compared to male police officers (low HDL 15% Vs. 11.6%; high WC 27.7% Vs. 19.3%, p <0.001), whereas high BP, high TG, and high FG was observed in male police officers (high BP 67.4% Vs. 33.3 %; high TG 54.7% Vs. 17%; high FG 36.6% Vs. 16.3%, p< 0.001).

Table 3 shows a multivariable analysis of crude and adjusted factors associated with the prevalence of MetS in men and women. The factors included were age range, pulse pressure, gender, rank, location (metropolitan or non-metropolitan), and alcohol consumption. After controlling for potential confounding factors in the multivariable analysis, increasing age, pulse pressure above 45, male gender, lower rank, non-metropolitan location, and alcohol drinking were positively associated with MetS. Male police officers had approximately triple more risk of MetS compared to female police officers (OR 2.59, 95% CI 2.30 - 2.92, p < 0.001). Police officers at lower-ranking posts had an additional odds of getting MetS compared to the police officers at higher-ranking posts (OR 1.17, 95% CI 1.12 - 1.22, p < 0.001). The MetS risk doubled when the pulse pressure increased to more than 55 (OR 2.24, 95% CI 2.10 - 2.40, p < 0.001). Compared to Metropolitan police officers, non-metropolitan police officers had a higher risk of MetS (OR 1.25, 95% CI 1.17 -1.34, p < 0.001). Alcohol drinking was significantly associated with MetS (OR 1.29, 95% CI 1.25 - 1.34, p < 0.001).

DISCUSSION

This is the first study conducted in Thailand reporting the findings indicating a high prevalence of MetS in active-duty police officers aged between 35 to 59 years at 39.24%. Erstwhile studies conducted on the prevalence of MetS among police officers vary according to the countries and populations under study. The predominance of MetS in this study was approximately similar to Brazilian police officers (38.54%)[31] but higher than in Iowa, US police officers (27.5%), with a higher cut-off point for abdominal obesity.[32] A study in Japan, using a higher FG cut-off point, reported a prevalence of MetS of 25% in the police officers of the age group of 45-59 years which was lower than our study (45% in the same age group).[33] Studies from India reported varying prevalence of MetS among police officers, 16.8% in Kerala[34] and 57.3% in Chennai[16], India. The difference in MetS prevalence between countries may be due to race-specific diagnostic criteria for MetS.

Compared to the prevalence of MetS in the Thai general population, specifically among those in the age range from 35 to 59[6], the overall age-standardized prevalence of Mets was significantly higher in Thai police officers (39.24% Vs. 27.82%, p<0.001). Nevertheless, MetS was substantially higher in male officers (40.65% Vs. 23.80%) but lower in female officers (16.30% Vs. 31.59%), and the overall prevalence of MetS in police officers increased with age more remarkably than in the general population. This finding was consistent with previous studies in India, Indonesia, and China comparing

TABLE 1. Baseline characteristics of Thai police officers aged ≥ 35 years according to gender.

Parameters	Male, n (%)	Female, n (%)	Total, n (%)	P value
Total	101,846 (94.4)	6,087 (5.6)	107,933 (100)	<0.001
Age, years, mean (SD)	46.3 ± 0.02	46.2 ± 0.02	43.6 ± 0.08	<0.001
Age – group, years				
35-39	15216 (14.9)	2004 (32.9)	17220 (15.9)	<0.001
40-44	32316 (31.7)	1671 (27.5)	33987 (31.5)	<0.001
45-49	19767 (19.4)	1315 (21.6)	21082 (19.5)	<0.001
50-54	20889 (20.9)	758 (12.5)	21647 (20.1)	<0.001
55-59	13658 (13.7)	339 (5.6)	13997 (13.0)	<0.001
Body mass index, kg/m^2, mean (SD)	24.9±0.01	23.3 ± 0.05	24.9 ± 0.01	<0.001
Obese (BMI 25 kg/m^2)	46500 (46)	1598 (26)	40658 (45)	<0.001
Waist circumference, cm, mean (SD)	85.80 ± 0.03	77.40 ± 0.12	85.3 ± 0.03	<0.001
Blood pressure, mmHg, mean (SD)				
Systolic blood pressure	133.72 ± 0.05	120.9 ± 0.20	133.0 ± 0.05	<0.001
Diastolic blood pressure	85.2 ± 0.04	76.96 ± 0.14	84.7 ± 0.04	<0.001
Mean arterial pressure	101.4 ± 0.04	91.60 ± 0.15	100.8 ± 0.04	<0.001
Pulse Pressure	48.5 ± 0.04	43.9 ± 0.13	48.3 ± 0.04	<0.001
Total Cholesterol, mg/dL, mean (SD)	227.5 ± 0.15	215.9 ± 0.53	226.9 ± 0.14	<0.001
Triglycerides, mg/dL, mean (SD)	204.7 ± 0.54	108.1 ± 0.86	199.3 ± 0.51	<0.001
HDL cholesterol, mg/dL, mean (SD)	53.3 ± 0.04	65.5 ± 0.20	54.0 ± 0.04	<0.001
LDL cholesterol, mg/dL, mean (SD)	136.7 ± 0.14	128.4 ± 0.46	136.3 ± 0.13	<0.001
Glucose, mg/dL, mean (SD)	102.9 ± 0.11	91.4 ± 0.26	102.3 ± 0.11	<0.001
Smoking	26590 (26.1)	19 (0.3)	26609 (24.7)	<0.001
Alcohol drinking	47567 (46.7)	404 (6.6)	47971 (44.5)	<0.001
Rank				
High (Commissioned)	19,665 (19.3)	2494 (41.0)	22159 (20.5)	<0.001
Low (Non-Commissioned)	82,181 (80.7)	3593 (59.0)	85774 (79.5)	<0.001
Duty location				
Metropolitan city	11606 (11.4)	2987 (49.0)	14593 (13.5)	<0.001
Non-Metropolitan city	90240 (88.6)	3100 (51.0)	93340 (86.5)	<0.001

Abbreviations: BMI= body mass index; HDL= high density lipoprotein; LDL= low density lipoprotein
**Figures in bracket indicate percentage

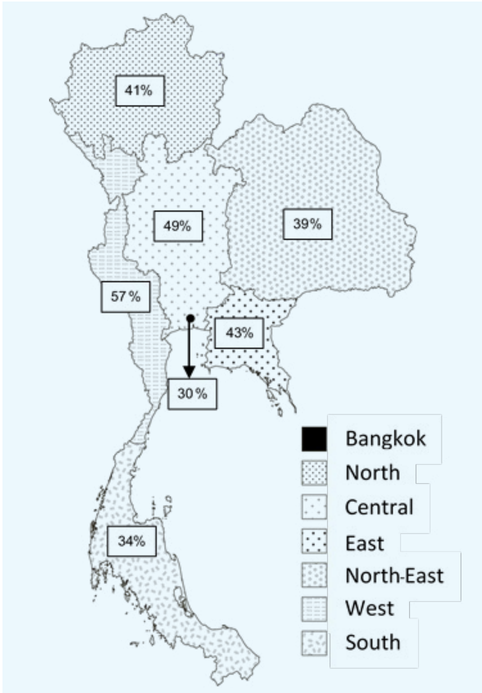


Fig 1. Age-adjusted prevalence of Metabolic Syndrome in percentage (%) by regions in Thailand

TABLE 2. Age-adjusted prevalence of metabolic syndrome and abnormal metabolic components among Thai police officers aged ≥ 35 years.

MetS and its components	Total, n (%)	Age group, n (%)					P for trend
		35-39	40-44	45-49	50-54	55-59	
Men	101846 (100)	15216 (14.9)	32316 (31.7)	19767 (19.4)	20889 (20.5)	13658 (13.4)	-
MetS	41671 (40.9)	4979 (35.7)	11529 (32.7)	8147 (41.2)	10065 (48.2)	6951 (50.9)	<0.001
WC > 90 cm	19663 (19.3)	2357 (19.7)	5499 (21.3)	3754 (23.7)	4657 (29.6)	3396 (32.5)	<0.001
High BP	68593 (67.4)	8996 (59.1)	20194 (62.5)	13186 (66.7)	15583 (74.6)	10634 (77.9)	<0.001
TG>150 mg/dl	55755 (54.7)	7537 (49.5)	16329 (50.5)	11561 (58.5)	12507 (59.9)	7821 (57.3)	<0.001
HDL<40 mg/dl	11796 (11.6)	1664 (10.9)	3579 (11.1)	2217 (11.2)	2559 (12.3)	1777 (13)	<0.001
FG>100 mg/dl	37281 (36.6)	4175 (27.5)	9954 (30.8)	6889 (34.9)	9417 (45.1)	6846 (50.2)	<0.001
Women	6087 (100)	2004 (32.9)	1671 (27.5)	1315 (21.6)	758 (12.5)	339 (05.6)	-
MetS	873 (14.3)	178 (8.9)	217 (13)	232 (17.6)	165 (21.8)	81 (23.9)	<0.001
WC > 80 cm	1684 (27.7)	464 (26.5)	426 (30.1)	371 (33.2)	273 (40.1)	150 (49.3)	<0.001
High BP	2027 (33.3)	494 (24.7)	519 (31.1)	505 (38.4)	356 (47)	153 (45.1)	<0.001
TG>150 mg/dl	1035 (17.0)	275 (13.7)	228 (13.6)	289 (22)	172 (22.7)	71 (20.9)	<0.001
HDL<50mg/dl	911 (15.0)	274 (13.7)	271 (16.2)	214 (16.3)	103 (13.6)	49 (14.5)	0.61
FG>100 mg/dl	995 (16.3)	226 (11.3)	257 (15.4)	240 (18.3)	176 (23.2)	96 (28.3)	<0.001

Abbreviations: MetS= Metabolic syndrome; WC= waist circumference; BP= blood pressure; TG= triglyceride; HDL= high density lipoprotein; FG= fasting glucose
** Figures in bracket indicate percentage

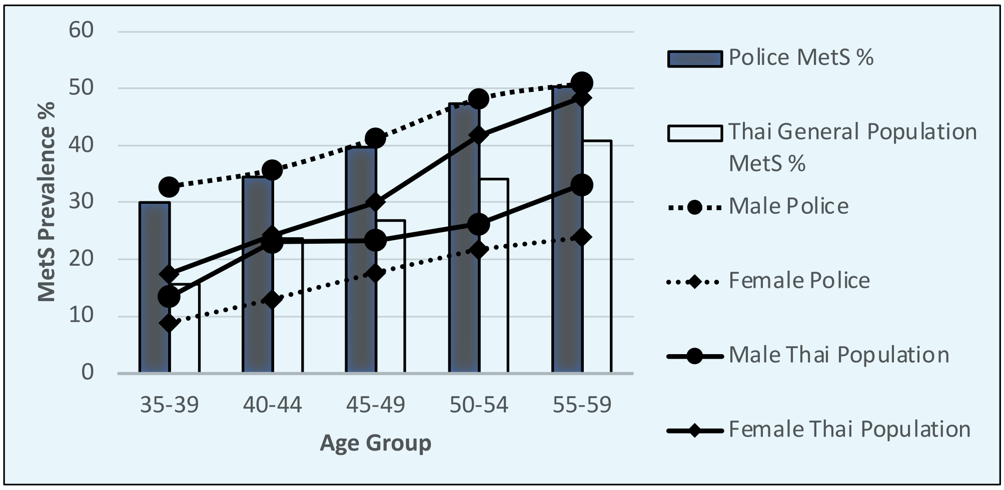


Fig 2. Age-adjusted and gender based prevalence of MetS between Thai police officers and Thai general (NHES 4) population.
Abbreviations: MetS = Metabolic syndrome; NHES= National Health Examination Survey

TABLE 3. The crude and age adjusted odds ratio (OR) with 95% confidence interval (CI) of socio-demographic characteristics and physical activity level on the risk of metabolic syndrome using multivariable logistic regression model.

Variables	Crude OR (95 % CI)	P value	Adjusted OR (95 % CI)	P value
Age (year)				
35-39	1 (-)	-	1 (-)	-
40-44	1.24 (1.19, 1.29)	<0.001	1.18 (1.13, 1.24)	<0.001
45-49	1.54 (1.48, 1.61)	<0.001	1.37 (1.30, 1.44)	<0.001
50-54	2.1 (2.01, 2.19)	<0.001	1.69 (1.60, 1.79)	<0.001
55-59	2.36 (2.25, 2.47)	<0.001	1.84 (1.73, 1.96)	<0.001
Pulse Pressure				
< 35	1 (-)	-	1 (-)	-
35 – 45	1.03 (0.98, 1.08)	0.304	0.99 (0.93, 1.06)	0.764
46 – 55	1.66 (1.59, 1.75)	<0.001	1.53 (1.43, 1.62)	<0.001
> 55	2.79 (2.66, 2.94)	<0.001	2.24 (2.10, 2.40)	<0.001
Gender				
Male	4.14 (3.85, 4.45)	<0.001	2.59 (2.30, 2.92)	<0.001
Female	1 (-)	-	1 (-)	-
Rank				
Low (Non-commissioned)	1.19 (1.15, 1.23)	<0.001	1.17 (1.12, 1.22)	<0.001
High (Commissioned)	1 (-)	-	1 (-)	-
Province				
Metropolitan	1 (-)	-	1 (-)	-
Non-Metropolitan	1.59 (1.53, 1.65)	<0.001	1.25 (1.17,1.34)	<0.001
Smoking				
Yes	1 (-)	0.002	1 (-)	-
No	1.05 (1.02, 1.08)	-	1.13 (1.09, 1.17)	<0.001
Alcohol drinking				
Yes	1.29 (1.25, 1.33)	<0.001	1.29 (1.25, 1.34)	<0.001
No	1 (-)	-	1 (-)	-

MetS prevalence between police officers and the general population reporting a higher prevalence of MetS among police officers.[19,33-36] A case control study conducted in Japan investigated the prevalence of risk factors for ischemic heart disease (IHD) among police officers and examined its association with working conditions and lifestyles. It demonstrated that the frequency of MetS and the risk factors of IHD increased with age and was more rampant in police officers than in office workers. The higher prevalence of MetS in police officers is likely to be attributed to their sporadic working conditions due to shift work, long working time, and unhealthy behaviors including alcohol drinking[11,13,18,33], which in turn predisposes them to develop MetS at an early age. A further study would be required to clarify the mutual effects about the significant relationship between smoking, sleep habits, work stress, and development of MetS.

A lower prevalence of MetS in female officers compared to the general female population, especially office workers, similar to our study, was seen in a study conducted in Germany[37], possibly due to differences in the job characterized by a more sedentary workload and less physical activity. Nonetheless, the exact mechanism for the difference is beyond the scope of this study.

The prevalence was significantly higher among male police officers compared to female police officers by 2.59-fold. This finding was similar to studies conducted in Taiwan and China[19,38] The gender difference in MetS prevalence is observed in many studies attributing to the complex interplay of effects of female and male sex hormones (such as estrogens, androgens, etc), sex- specific gene expression that determines body fat distribution, and attendant cardiometabolic abnormalities regulated by hormones.[39-42] Therefore leading to a higher risk of MetS in males than in premenopausal females, whereas post menopause, the estrogen decline has implications for a higher risk of MetS in females.[39] In this study, despite the similar mean age of males and females, a higher proportion of younger and lower proportion of older female officers compared to male officers could also explain the gender difference in MetS prevalence (Table 1) attributing to lower percentage of menopause in this study group. The higher percentage of male police officers indulging in alcohol drinking and smoking could also lead to higher MetS pervasiveness in this population.

The most frequent component of MetS found in this study was high BP which was more prevalent in males than females. The study by Thayyil and colleagues also showed that the high prevalence of the police officers with MetS had high BP[34], which was in line with similar findings of another study by Yates and colleagues.[43] High perceived and oxidative stress are instrumental in promoting hypertension in police officers.[44] Arterial stiffening and aging being the crucial factors leading to hypertension[41], a relatively higher percentage of older male police officers in this study could have caused the gender based difference. Other than high BP, high TG in male police officers could be attributed to a higher prevalence of alcohol drinking and extrapolating it to other unhealthy habits and lesser physical activity in this population. High impaired FG or dysglycemia are commonly observed in the male population compared to females like in this study. However, the reason for such a pattern is yet to be elucidated. Gender differences in visceral adiposity, lean muscle mass, altered susceptibility to free fatty acid-induced peripheral insulin resistance, the influence of menopausal transition, and other factor may play a role.[41,45] In the same context, high WC and low HDL percentages among female police officers seen in this study were also observed in a study conducted by Hartley and colleagues[12], which can be explained by sexual dysmorphism in body fat distribution and adipocyte size that correlate with a measure of alteration in lipid metabolism and insulin resistance. Premenopausal females have less visceral fat despite having a higher total body fat, BMI, and abdominal subcutaneous adipose tissue, which is gained with age, whereas visceral adipose tissue increase only in post menopause females.[45,46] That explains higher WC, linear correlation between WC, and transitional age increasing towards menopause among female police officers.

Older age, wider pulse pressure, male sex, low rank, non-metropolitan location, alcohol consumption and smoking are associated with increased risk of MetS in this study, which was consistent with the findings of previous studies.[15,19,35,47-50] Male police officers had significantly increased odds of metabolic syndrome.

Low police rank (Non-commissioned) had a higher risk of developing MetS compared to high police rank (Commissioned). A similar finding was also seen in a study by Fontes and colleagues[51] among corporals who were at higher MetS risk. In a NHES 4 study, women with an education level lower than high school had a 60% additional risk of MetS compared to those with a higher educational level.[6] A study conducted in China showed that compared with those with no education, every category of attained academic level lowered the risk of developing MetS.[52] Lower education levels might contribute to inappropriate lifestyle choices, poor eating habits, and physical inactivity.[6,53,54] A lower level of education and socioeconomic status play a decisive role in increasing cardiometabolic risk factors.[41] This could be related to increased MetS risk among low-rank police officers whose required education level is high school, whereas higher rank requires at least a college degree. A higher percentage of smoking and alcohol drinking among lower ranked or non-commissioned police officers may also lead to the difference observed. Another factor influencing higher prevalence in lower ranked police officers could be job descriptions related to shift work disorder leading to sleep deprivation, unhealthy food intake, limited physical activity, drinking, and smoking. The available data suggest that hormonal profiles, coagulatory mediators and autonomic functioning inflammatory are all altered during uncontrolled sleep deprivation, which contributes to the development of atherosclerosis and cardiovascular disease.[10,55,56] However, further investigation regarding the causal relationship is warranted.

The findings of this study reveal that this occupational group of people experiences an increased risk for cardiovascular disease and diabetes, challenging the general notion that police officers constitute a physically fit and healthy population.

There were 19% female and 30.6% male police officers that had "pre-MetS" (2 components of MetS), implicating a significant population at risk of MetS. More than 60% of police officers indulged in drinking alcohol. These findings call for more precise interventions focusing on the detection and treatment of MetS, including health promotion strategies to reduce the incidence of CVD, disability, and death among police officers.

Although the advantage of this study was its large sample size, there were some limitations. Considering the low recruitment percentage of police officers (60%), the study population may not represent the entire Thai police officer population. Data regarding alcohol consumption and smoking was self-reported and may be bias. There were missing or incomplete sociodemographic data, therefore, had to be removed from the study. Other contributors to MetS, including physical inactivity, sleep deprivation, work stress, and low job satisfaction were not assessed. A further in-depth study on associated lifestyle and environmental factors is recommended. This cross-sectional study compared the different burden of MetS using data from the health examination survey of police personnel and the general population during the different year. MetS has become more common among the general Thai community over time therefore this could be one of the limitations of the study findings.

CONCLUSION

The prevalence of MetS is high among Thai police officers compared to the general Population, and the syndrome affects a substantial proportion of male than female police officers in the same age group. Increased risk of MetS among police officers included lower rank, being in a non-metropolitan city, smoking, and drinking alcohol. Therefore, preventive health promotion and treatment programs should be established along with strategic goals to improve the overall health of individual police officers to alleviate adverse cardiovascular outcomes.

ACKNOWLEDGEMENTS

First, I would like to thank the Director General of Police General Hospital for permitting me to conduct this research. Second, my sincere thanks to Dr. Supalerk Pattanaprichakul and Dr. Robert DuBroff for their valuable time and guidance in completing this research paper. Last but not the least, I would like to thank Somchai Wutinarongtrakul MD, Ucharamont Buranasompop , Police Health Care Center, and Pornpen Bunnag MD, Family Medicine Department Police General Hospital Bangkok, Thailand, for their assistance in this research.

REFERENCES

1. Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, et al. The metabolic syndrome and cardiovascular risk: A systematic review and meta-analysis. J Am Coll Cardiol 2010;56:1113-32. https://doi.org/10.1016/j.jacc.2010.05.034.
2. Grundy SM. Metabolic syndrome pandemic. Arterioscler Thromb Vasc Biol 2008;28:629-36. https://doi.org/10.1161/ATVBAHA.107.151092.
3. Moore JX, Chaudhary N, Akinyemiju T. Metabolic syndrome prevalence by race/ ethnicity and sex in the united states, national health and nutrition examination survey, 1988-2012. Prev Chronic Dis 2017;14:E24. https://doi.org/10.5888/pcd14.160287.
4. Ranasinghe P, Mathangasinghe Y, Jayawardena R, Hills AP, Misra A. Prevalence and trends of metabolic syndrome among adults in the asia-pacific region: a systematic review. BMC Public Health 2017;17:101. https://doi.org/10.1186/s12889-017-4041-1.
5. Scuteri A, Laurent S, Cucca F, Cockcroft J, Cunha PG, Mañas LR, et al. Metabolic syndrome across Europe: Different clusters of risk factors. Eur J Prev Cardiol 2015;22:486-91. https://doi.org/10.1177/2047487314525529.
6. Aekplakorn W, Kessomboon P, Sangthong R, Chariyalertsak S, Putwatana P, Inthawong R, et al. Urban and rural variation in clustering of metabolic syndrome components in the Thai population: Results from the fourth National Health Examination Survey 2009. BMC Public Health 2011;11:854. https://doi.org/10.1186/1471-2458-11-854.
7. Aekplakorn W, Chongsuvivatwong V, Tatsanavivat P, Suriyawongpaisal P. Prevalence of metabolic syndrome defined by the International Diabetes Federation and National Cholesterol Education Program criteria among Thai adults. Asia-Pacific J Public Heal 2011;23:792-800. https://doi.org/10.1177/1010539511424482.
8. Suriyawongpaisal P. Cardiovascular risk factor levels in urban and rural Thailand - The International Collaborative

Study of Cardiovascular Disease in Asia (InterASIA). Eur J Cardiovasc Prev Rehabil 2003;10:249-57. https://doi.org/10.1097/00149831-200308000-00006.
9. Garbarino S, Magnavita N. Sleep problems are a strong predictor of stress-related metabolic changes in police officers. A prospective study. PLoS One 2019;14(10):e0224259. https://doi.org/10.1371/journal.pone.0224259.
10. Violanti JM, Burchfiel CM, Hartley TA, Mnatsakanova A, Fekedulegn D, Andrew ME, et al. Atypical work hours and metabolic syndrome among police officers. Arch Environ Occup Heal 2009;64:194-201. https://doi.org/10.1080/19338240903241259.
11. da-Silva CAC, Leite AL, Moreira JA, Abreu DDC, Oliveira PE de A, Nunes DP, et al. Association of dyslipidemia, hypertension and overweight/obesity with work shift and duration of employment among police officers in a small town in Northeastern Brazil. Rev Bras Med Do Trab Publicacao Of Da Assoc Nac Med Do Trab 2019;17:537-44. https://doi.org/10.5327/Z1679443520190401.
12. Hartley TA, Fekedulegn D, Knox SS, Burchfiel CM, Andrew ME, Violanti JM. Associations between police officer stress and the metabolic syndrome. Int J Emerg Ment Health 2011;13:243-56.
13. Chang JH, Huang PT, Lin YK, Lin CE, Lin CM, Shieh YH, et al. Association between sleep duration and sleep quality, and metabolic syndrome in Taiwanese police officers. Int J Occup Med Environ Health 2015;28:1011-23. https://doi.org/10.13075/ijomeh.1896.00359.
14. Violanti JM, Hartley TA, Charles LE, Fekedulegn D, Andrew ME, Mnatsakanova A, et al. Police trauma and cardiovascular disease: Association between PTSD symptoms and metabolic syndrome. Int J Emerg Ment Health 2006;8:227-38.
15. Rostami H, Tavakoli HR, Rahimi MH, Mohammadi M. Metabolic Syndrome Prevalence among Armed Forces Personnel (Military Personnel and Police Officers): A Systematic Review and Meta-Analysis. Mil Med 2019;184:e417-25. https://doi.org/10.1093/milmed/usz144.
16. Chandramohan P, Mohan V. High prevalence of diabetes and metabolic syndrome among policemen. J Assoc Physicians India 2008;56:837-8.
17. Tharkar S, Kumpatla S, Muthukumaran P, Viswanathan V. High prevalence of metabolic syndrome and cardiovascular risk among police personnel compared to general population in India. J Assoc Physicians India 2008;56:845-9.
18. Hong JH, Lee S-W. A Study on the Prevalence and Influencing Factors of Metabolic Syndrome among Police Officers. J Korean Soc Occup Environ Hyg 2014;24:566-77. https://doi.org/10.15269/jksoeh.2014.24.4.566.
19. Zhang J, Liu Q, Long S, Guo C, Tan H. Prevalence of metabolic syndrome and its risk factors among 10,348 police officers in a large city of China: A cross-sectional study. Med (United States) 2019;98(40):e17306. https://doi.org/10.1097/MD.0000000000017306.
20. Nillakupt K, Viravathana N. A survey of metabolic syndrome and its components in Thai medical cadets. J Med Assoc Thai 2010;93 Suppl 6:S179-85.
21. Napradit P, Pantaewan P, Nimit-arnun N, Souvannakitti D, Rangsin R. Prevalence of overweight and obesity in Royal Thai Army personnel. J Med Assoc Thai 2007;90:335-40.
22. National Statistical office, National Statistical Office Thailand, Technology M of I and communication. Key Statistics of Thailand 2012. 2012.
23. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Heal Organ - Tech Rep Ser 2000;894:253.
24. Weber MA, Schiffrin EL, White WB, Mann S, Lindholm LH, Kenerson JG, et al. Clinical Practice Guidelines for the Management of Hypertension in the Community: A Statement by the American Society of Hypertension and the International Society of Hypertension. J Clin Hypertens (Greenwich) 2014; 16(1):14-26. https://doi.org/10.1111/jch.12237.
25. Knopfholz J, Disserol CCD, Pierin AJ, Schirr FL, Streisky L, Takito LL, et al. Validation of the friedewald formula in patients with metabolic syndrome. Cholesterol 2014;2014:261878. https://doi.org/10.1155/2014/261878.
26. Certified AS, Cholesterol LDL. Cholesterol Reference Method Laboratory Network List of Analytical Systems Certified for LDL Cholesterol 2010;186:1-9.
27. Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: A joint interim statement of the international diabetes federation task force on epidemiology and prevention; National heart, lung, and blood institute; American heart association; World heart federation; International atherosclerosis society; And international association for the study of obesity. Circulation 2009;120:1640-5. https://doi.org/10.1161/CIRCULATIONAHA.109.192644.
28. Shiwaku K, Nogi A, Kitajima K, Anuurad E, Enkhmaa B, Yamasaki M, et al. Prevalence of the metabolic syndrome using the modified ATP III definitions for workers in Japan, Korea and Mongolia. J Occup Health 2005;47:126-35. https://doi.org/10.1539/joh.47.126.
29. Lohsoonthorn V, Lertmaharit S, Williams MA. Prevalence of metabolic syndrome among professional and office workers in Bangkok, Thailand. J Med Assoc Thai 2007;90:1908-15.
30. National Statistical Office Thailand, Ministry of Information and communication technology. The Population And Housing Census 2010. http://service.nso.go.th/nso/nso_center/project/search_center/23project-en.htm.
31. Filho RTB, D'Oliveira A. The Prevalence of Metabolic Syndrome Among Soldiers of the Military Police of Bahia State, Brazil. Am J Mens Health 2014;8:310–5. https://doi.org/10.1177/1557988313510928.
32. Yoo H. Assessment of contributors to the metabolic syndrome among law enforcement officers. ProQuest Diss Theses 2011:116.
33. Shiozaki M, Miyai N, Morioka I, Utsumi M, Koike H, Arita M, et al. Assessment of the risk of ischemic heart disease and its relevant factors among Japanese police officers. Sangyo Eiseigaku Zasshi 2013;55(4):115-24. https://doi.org/10.1539/sangyoeisei.B12007.
34. Thayyil J, Jayakrishnan TT, Raja M, Cherumanalil JM. Metabolic syndrome and other cardiovascular risk factors among police officers. N Am J Med Sci 2012;4:630-5. https://doi.org/10.4103/1947-2714.104313.
35. Tharkar S, Kumpatla S, Muthukumaran P, Viswanathan V. High prevalence of metabolic syndrome and cardiovascular risk among police personnel compared to general population in India. J Assoc Physicians India 2008;56:845-9.
36. Febyan ., Wetarini K, Rendy ., Thintarso CS, Suastika K. A Clinical Profile of Metabolic Syndrome and Its Determinants among Police Officers in Bali. Eur J Med Heal Sci 2020;2. https://doi.org/10.24018/ejmed.2020.2.6.619.

37. Strauss M, Foshag P, Brzek A, Vollenberg R, Jehn U, Leischik R. Metabolic syndrome in female police officers and female office workers: A cross-sectional study in occupations with different physical activities. Diabetes, Metab Syndr Obes Targets Ther 2020;13:3487-95. https://doi.org/10.2147/DMSO.S267948.
38. Wang WY, Li CH, Wu YS, Chien WC, Wang KY, Tzeng WC. Gender Differences in the Prevalence of Metabolic Syndrome Among Taiwanese Air Force Personnel: A Population-Based Study. J Cardiovasc Nurs 2020;35:502-11. https://doi.org/10.1097/JCN.0000000000000714.
39. Vona R, Gambardella L, Straface E. Gender-Associated Biomarkers in Metabolic Syndrome. Carotid Artery - Gend. Heal. [Working Title], IntechOpen; 2018. https://doi.org/10.5772/intechopen.81103.
40. Ventura-Clapier R, Dworatzek E, Seeland U, Kararigas G, Arnal J-F, Brunelleschi S, et al. Sex in basic research: concepts in the cardiovascular field. Cardiovasc Res 2017;113:711-24. https://doi.org/10.1093/cvr/cvx066.
41. Gerdts E, Regitz-Zagrosek V. Sex differences in cardiometabolic disorders. Nat Med 2019;25:1657-66. https://doi.org/10.1038/s41591-019-0643-8.
42. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature 2015;518:187-96. https://doi.org/10.1038/nature14132.
43. Yates JD, Aldous JWF, Bailey DP, Chater AM, Mitchell ACS, Richards JC. The prevalence and predictors of hypertension and the metabolic syndrome in police personnel. Int J Environ Res Public Health 2021;18(13):6728. https://doi.org/10.3390/ijerph18136728.
44. Janczura M, Rosa R, Dropinski J, Gielicz A, Stanisz A, Kotula-Horowitz K, et al. The associations of perceived and oxidative stress with hypertension in a cohort of police officers. Clin Cosmet Investig Dermatol 2021;14:1783-97. https://doi.org/10.2147/DMSO.S298596.
45. Pradhan AD. Sex Differences in the Metabolic Syndrome: Implications for Cardiovascular Health in Women 2013. https://doi.org/10.1373/clinchem.2013.202549.
46. Lovejoy JC, Champagne CM, De Jonge L, Xie H, Smith SR. Increased visceral fat and decreased energy expenditure during the menopausal transition. Int J Obes 2008;32:949-58. https://doi.org/10.1038/ijo.2008.25.
47. Trivedi T, Liu J, Probst JC, Martin AB. The Metabolic Syndrome: Are Rural Residents at Increased Risk? J Rural Heal 2013;29:188-97. https://doi.org/10.1111/J.1748-0361.2012.00422.X.
48. Atanasova G, Yordanova - Laleva P, Marinov MA. Metabolomics and biomarkers. Clin Chem Lab Med 2017;55:22-5. https://doi.org/10.1515/cclm-2017-7062.
49. Kwon YJ, Chung TH, Shim JY, Lee YJ. The association of pulse pressure with metabolic syndrome in Korean elderly: A nationwide population-based study. Diabetes Res Clin Pract 2017;123:75-81. https://doi.org/10.1016/J.DIABRES.2016.11.027.
50. Xin S, Feng-shi T, Song Y, Xin S, Feng-shi T, Song Y. Prevalence and influencing factors of metabolic syndrome among police. Chinese Journal of Public Health 2010;26:1376-8. https://doi.org/10.11847/zgggws2010-26-11-16.
51. Santos Fontes IS, Sampaio CG, Dos Santos CG, Sá Barros AMM, Fontes DS, Dos Santos IMC, et al. Síndrome metabólica em militares de uma unidade da polícia de Aracaju, Sergipe. Sci Med (Porto Alegre) 2016;26:24160. https://doi.org/10.15448/1980-6108.2016.3.24160.
52. Li Y-Q, Zhao L-Q, Liu X-Y, Wang H-L, Wang X-H, Li B, et al. Prevalence and distribution of metabolic syndrome in a southern Chinese population. Relation to exercise, smoking, and educational level. Saudi Med J 2013;34:929–36.
53. (US) NC for HS. Health, United States, 2011. Heal United States, 2011 With Spec Featur Socioecon Status Heal 2012.
54. Cathal Doyle, Paul Kavanagh, Owen Metcalfe TL. Health Impacts of Employment a review Prepared by. 2005.
55. Mullington JM, Haack M, Toth M, Serrador JM, Meier-Ewert HK. Cardiovascular, Inflammatory, and Metabolic Consequences of Sleep Deprivation. Prog Cardiovasc Dis 2009;51:294-302. https://doi.org/10.1016/j.pcad.2008.10.003.
56. van Leeuwen WMA, Lehto M, Karisola P, Lindholm H, Luukkonen R, Sallinen M, et al. Sleep restriction increases the risk of developing cardiovascular diseases by augmenting proinflammatory responses through IL-17 and CRP. PLoS One 2009;4(2):e4589. https://doi.org/10.1371/journal.pone.0004589.

Emotion Regulation and Behavioral Problems in Children with Autism Spectrum Disorder: A University Hospital Based Cross-Sectional Study

Tikumporn Hosiri, M.D.*, Napat Sittanomai, M.D., Thanchanok Bumrungtrakul, M.Sc.*, Vitharon Boonyasidhi, M.D.****

**Department of Psychiatry, **Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Objective: The objectives of this study were to investigate behavioral problems and emotion regulation in children with autism spectrum disorder (ASD) with and without psychiatric comorbidities, and to assess the correlation between emotion regulation and behavioral problems.

Materials and Methods: Participants included 65 children with ASD (ASD group) and 65 children with ASD and psychiatric comorbidities (ASD+CPD group), aged 6 to 12, and their caretakers. Emotion regulation was assessed by the Emotion Regulation Checklist (ERC). Behavioral problems were assessed by the Strengths and Difficulties Questionnaire (SDQ).

Results: An independent sample T-test revealed that the ASD+CPD group had poorer emotion regulation, greater emotion lability/negativity, and a higher total difficulty score when compared to the ASD group. The Pearson's Correlation Coefficient indicated that emotion regulation (r = -0.47, p<0.05) and lability/negativity (r = 0.65, p<0.05) correlated with the total difficulties score of the SDQ. A multiple regression analysis revealed that lower emotion regulation and higher emotion lability/negativity predicted an increase in behavioral problems ($R^2 = 0.44$, $p<0.05$).

Conclusion: We found that children with ASD and psychiatric comorbidities had poorer emotion regulation and more behavioral problems than those without comorbidities, and that poorer emotion regulation predicts behavioral problems. These results indicate that children with ASD should be assessed for psychiatric comorbidities, particularly those who have emotional dysregulation and behavioral problems.

Keywords: Autism spectrum disorder; behavioral problems; emotion regulation; psychiatric comorbidities (Siriraj Med J 2023; 75: 218-223)

INTRODUCTION

Autism spectrum disorder (ASD) is a complicated neurodevelopmental disorder characterized by early onset of social communication deficits, repetitive patterns of behavior and restricted interests.[1] Previous studies have reported the prevalence of comorbid psychiatric disorders (CPD) in ASD at around 50-70% (e.g., attention deficit/ hyperactivity disorder (ADHD), intellectual disability (ID), and anxiety disorders), with around 40-90% having two or more comorbid disorders.[1-4] Based on the high prevalence of comorbidities in ASD, research has started to focus on identifying the consequences of the comorbidities and discovered that the presence of additional psychiatric disorders leads to more severe impairment due to behavioral and emotional problems that interfere with daily life activities, including social, academic, or work

Corresponding author: Napat Sittanomai
E-mail: napat.sittanomai@gmail.com
Received 3 December 2022 Revised 10 January 2023 Accepted 11 January 2023
ORCID ID:http://orcid.org/0000-0001-7169-3041
https://doi.org/10.33192/smj.v75i3.260870

performance.[5-11] Research in this field has demonstrated that behavioral problems in children with ASD are related to their emotional problems. For instance, Mazefsky and Herrington proposed that behavioral problems in this population stem from underlying impairment in emotion regulation making them react negatively and impulsively to emotional stimuli.[12] However, the relationship between comorbid psychiatric disorders and emotional and behavioral difficulties have been limited.[12-14] The primary objective of this study was to investigate emotion regulation and behavioral problems in children with autism, with and without psychiatric comorbidities. The secondary objective was to assess the relationship between emotion regulation and behavioral problems. The results of this study have a potential benefit in emphasizing clinicians who take care of children with ASD to look for psychiatric comorbidities, especially in cases that presented with emotional or behavioral problems.

MATERIALS AND METHODS

A cross-sectional study was conducted between 2020 and 2021. Participants were recruited from the Child and Adolescent Psychiatric Clinics at the Department of Psychiatry and the Department of Pediatrics, Siriraj Hospital, Mahidol University, Thailand. The study protocol was approved by the Siriraj Institutional Review Board (COA no. Si 666/2020).

Participants

The sample size was calculated based on the results from the study of Berenguer et al.[15] which indicated that children with ASD had SDQ mean scores 30% lower than children with ADHD and comorbidities (17.43±6.04 vs. 22.63±5.24). In this study, it was hypothesized that children with ASD would have similar SDQ mean scores of 17.43 and children with ASD and psychiatric comorbidities would have 20% higher SDQ mean scores which would be 20.92 [120% of 17.43]. Standard deviation (SD) of the two groups is estimated to be 6.04, with Type I error (a)= 0.05 and Type II error (b)= 0.1 [power of the test=90% in order to increase the probability of detecting a difference]. The calculated sample per group of participants with ASD and those with ASD and comorbidities was 64. During the study period, primary caretakers of 150 children aged 6-12 years with the diagnosis of ASD (75 children with ASD and 75 children with ASD and psychiatric comorbidities) were invited to participate in the study. The diagnosis of ASD and comorbid psychiatric disorders was made by the treating child and adolescent psychiatrists, based on the Diagnostic and Statistical Manual of Mental Disorder, 5th edition; DSM-5 criteria.[1] Exclusion criteria included children with serious medical conditions and caretakers who were not able to complete questionnaires. Of the invited 150 caretakers, 130 agreed to participate in the study. This included 65 caretakers of children with ASD without comorbidities (ASD group) and 65 of children with ASD and psychiatric comorbidities (ASD+ CPD group). The caretakers were asked to complete questionnaires to evaluate the children's emotion regulation and behavioral problems (Fig 1).

Instruments

Emotion Regulation Checklist (ERC)

ERC is a parent-reported screening measure comprising of 24 items used to assess emotion regulation capacity in children aged between 6 to12. It is divided into 2 subscales as follows: 1) Emotion regulation subscale (ERC–ER) and 2) Lability/Negativity subscale (ERC-LN). The ERC-ER evaluates a child's expression of emotions,

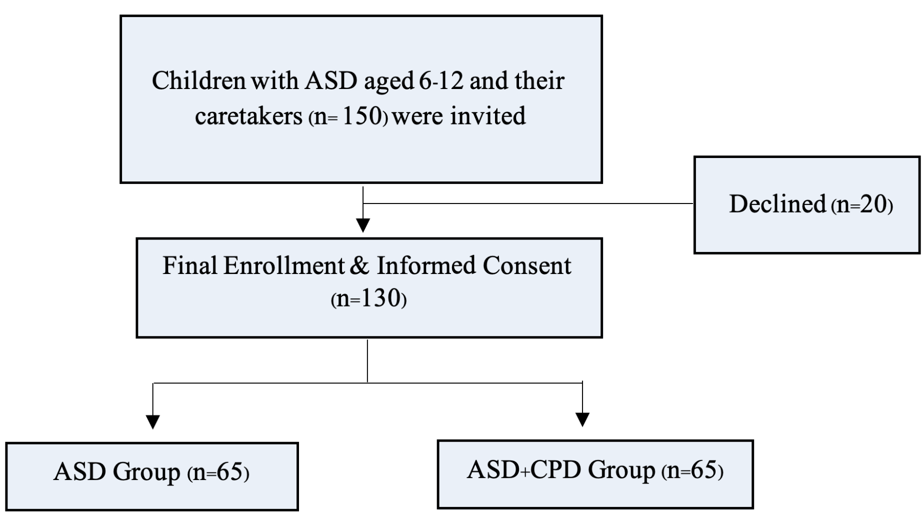


Fig 1. Flow diagram for participant recruitment and retention.
Abbreviations: ASD, autism spectrum disorder; CPD, comorbid psychiatric disorders

emotional self–awareness, and ability to display appropriate emotions. Higher scores on the ERC–ER reflect higher levels of emotion regulation ability. The Lability/Negativity Subscale (ERC–LN) assesses a child's inflexibility, variation in mood and emotional states, and dysregulation in negative effects. Higher scores on the ERC–LN reflect higher levels of emotional dysregulation. The internal consistency for both subscale is high; α= 0.83 for ERC-ER and α= 0.96 for ERC–LN.[16] Permission to use the ERC in the study was obtained from Dante Cicchetti, Ph.D. The ERC was translated into Thai by the investigator (TB) and then back translated by a qualified translator from the Translation and Interpretation Center, Faculty of Liberal Arts, Mahidol University for validation.

Strengths and Difficulties Questionnaire (SDQ)

SDQ is a parent-reported questionnaire designed to assess behavioral problems in children aged 4 to 16. It comprises of 25 items with five subscales, including hyperactivity/inattention scale, emotional problems scale, peer relationship problems scale, conduct problems scale, and prosocial behavior scale. Summation of all the subscales, except the prosocial behavioral scale, yields a total difficulties score. Higher scores indicate greater behavioral difficulties, except for the prosocial behavioral scale where higher scores reflect positive behavior attributes. This study used the Thai version of SDQ which demonstrated good reliability at α= 0.76.[17-19]

Data analyses

All statistical analyses were done based on a 95% confidence interval, p<0.05. Descriptive statistics were used to explain demographic data. Pearson's Chi-Square test was performed to examine categorical variables such as gender, education type, education level, and medication. An independent sample T-test was performed to compare emotion regulation, emotion lability/negativity, and behavioral problems between children with and without comorbidities. The association between emotion regulation, emotion lability/negativity, and behavioral problems were examined using the Pearson's Correlation Coefficient. A multiple linear regression analysis for the predictive variable of behavioral problems was also conducted.

RESULTS

Participant characteristics

The mean age of the children was 8.7 years (SD=2.97), with the range of 6 to 12 years. Eighty three percent of the children (n=108) were males. In the ASD+CPD group, 44 children (67.7%) had one comorbidity and 21 children (33.3%) had more than one comorbidity. Fifty three children (81.5%) had attention deficit/hyperactivity disorder (ADHD), 6 children (9.25%) had intellectual disability (ID), and 6 children (9.25%) had other comorbid diagnoses.

There were no significant group differences in gender, education type, or education level (p > 0.05) between the ASD group and the ASD+CPD group. Children in ASD+CPD group were on medication at a significantly higher rate compared to those in the ASD group (p <0.05) (Table 1).

Emotion regulation

Children in the ASD+CPD group had significantly lower Emotion Regulation Subscale (ERC-ER) scores and significantly higher Lability/Negativity Subscale (ERC-LN) scores (p<0.01) than children in the ASD group (Table 2).

Behavioral problems

Children in the ASD+CPD group had significantly higher scores on all SDQ subscales and the total difficulties score (p<0.01), and significantly lower prosocial behavior scores (p<0.05) than children in the ASD group (Table 2).

Correlation between the ERC and the SDQ

ERC-ER correlated negatively with the total difficulties scores, emotional problems scores, hyperactivity scores, and peer problems scores of the SDQ, and correlated positively with prosocial behavior scores. There was no correlation between ERC-ER and conduct problems scores (Table 3).

ERC-LN correlated positively with the total difficulties scores, emotion problems scores, conduct problem scores, hyperactivity scores, and peer problems scores of the SDQ. There was a negative correlation between ERC-LN and prosocial behavior scores (Table 3).

Predictors of behavioral problems

A multiple linear regression analysis was used to predict behavioral problems of a child's emotion regulation and emotion lability/negativity. The collinearity of emotion regulation and emotion lability/negativity was assessed. Results showed that Tolerance= 0.861 and VIF= 1.162. Therefore, emotion regulation and emotion lability/negativity have no multicollinearity problems. A significant regression equation was found to predict behavioral problems (R^2=0.44, F = 50.87, p<0.01), thus explaining the 44% variance in behavioral problems. When ERC-ER increases 10.54 scores, the Total Difficulties Score decreases one score. When ERC-LN increases

TABLE 1. Demographic characteristics of children with ASD (ASD group) and children with ASD and psychiatric comorbidities (ASD+CPD group).

	ASD Group (N=65)	ASD+CPD Group (N=65)	*P-value*
Age: Mean (SD)	8.5 (2.4)	9.0 (2.2)	0.18
Gender: n (%)			
Male	55 (84.6)	53 (81.5)	0.64
Education type: n (%)			
Mainstream	42 (64.6)	38 (58.5)	0.70
Inclusive classroom	5 (7.7)	9 (13.8)	
Special education	11 (16.9)	10 (15.4)	
Not in school	7 (10.8)	8 (12.3)	
Education level: n (%)			
Kindergarten	13 (20.0)	10 (15.4)	0.77
Primary school	31 (47.7)	35 (53.8)	
Secondary school	3 (4.6)	4 (6.2)	
Special education	11 (16.9)	8 (12.3)	
Not in school	7(10.8)	8 (12.3)	
Medication: n (%)			
On medication	25 (38.5)	51 (78.5)	<0.05*
Not on medication	40 (61.5)	14 (21.5)	

* $p < 0.05$

TABLE 2. Emotion Regulation Checklist and Strengths and Difficulties Questionnaire scores in children in the ASD and ASD+CPD groups.

	ASD Group	ASD + CPD Group	*P-value*
ERC: mean (SD)			
Emotion Regulation Subscale (ERC–ER)	23.7 (3.65)	21.4 (3.63)	<0.01**
Lability/Negativity Subscale (ERC–LN)	30.5 (5.12)	36.3 (5.71)	<0.01**
SDQ: mean (SD)			
Total difficulties score	14.70 (5.05)	20.0(4.40)	<0.01**
Emotional problems	2.8 (1.87)	4.3(2.14)	<0.01**
Conduct problems	2.1 (1.46)	3.2 (1.86)	<0.01**
Hyperactivity	5.4 (2.0)	7.01 (1.58)	<0.01**
Peer problems	4.4 (1.85)	5.4 (1.94)	<0.01**
Prosocial behavior	5.7 (2.25)	4.9 (2.23)	<0.05*

* $p < 0.05$, ** $p < 0.01$

TABLE 3. Correlation between ERC and SDQ.

	Total difficulties	Emotional problems	Conduct problems	Hyperactivity	Peer problems	Prosocial behavior
ERC-ER	–.47*	–.45*	–.12	–.35*	–.39*	.64*
ERC-LN	.65*	.42*	.56*	.47*	.36*	–.34*

*$p < 0.05$

16.46 scores, the Total Difficulties Score increases one score.

The predictive equation is Total Difficulties Score = 18.17 – 10.54 (ERC–ER) + 16.46 (ERC–LN).

DISCUSSION

The results of this study showed that children with ASD who had comorbid psychiatric disorders had more behavioral problems and poorer emotion regulation abilities, as well as greater emotion lability/negativity than those without comorbidities. These results are in line with existing research highlighting the impact of psychiatric comorbidities on additional impairment in psychological profile and overall function in children with ASD.[3,5,6,8]

The high percentage of ADHD diagnosis in this study is congruent with previous studies exploring the rate of ADHD in the ASD population. Epidemiology data from the United States and Europe indicates that 37% to 85% of children with ASD present with core symptoms of ADHD.[3,5,6,20-22] Children with ASD and ADHD symptoms often exhibit greater behavioral problems (including temper tantrums, conduct problems, physical aggression, self-injury), severe social impairment and delays in adaptive functioning than children with ASD without ADHD.[22,23] This may explain why children in the ASD+ CPD group in this study had more behavioral problems. The higher rate of emotional difficulties in the ASD+CPD group is also supported by previous studies which show greater both behavioral and emotional problems in children who were diagnosed with ASD and ADHD.[5,20,22-24] Since children with ADHD alone can have emotion regulation problem, the high percentage of ADHD comorbidity in this population could be a confounding factor. The further study investigating emotion regulation and behavioral problems among children with ASD, ASD and ADHD, and ASD and other psychiatric comorbidities might be warranted.

We also found a significant relationship between emotion regulation, emotion lability/negativity, and behavioral problems. Multiple regression analysis indicated that both emotion regulation and emotional lability/negativity were predictors of variance in behavioral problems. This supports our hypothesis and is congruent with previous studies which also found a relatively strong association between emotional dysregulation, emotion lability and externalizing behaviors.[14,25] This suggests that children diagnosed with ASD who have lower emotion regulation abilities or more emotional labile will generally exhibit more behavioral problems. However, we did not see a correlation between emotion regulation abilities and conduct problems. This result is in contrast to the result from a study by Guttmann-Steinmetz et al. that showed a link between emotional dysregulation and aggressive behavior, including conduct problems.[25] This might be due to the difference in conduct problems measured in each study.

There were some limitations in this study, firstly, we did not control the children's IQ levels and the intervention program, which could potentially affect the intensity and frequency of reported behavior and emotional problems. Secondly, not all children were given standardized IQ tests, therefore comorbid intellectual disability might be under diagnosed. Third, the population in this study was tertiary hospital-based that clinical presentations may be more severe. Then, generalization of the results to other clinical settings might be limited. Finally, 11.5% of children in our study were not currently in school, therefore certain questions about school behaviors were not be applicable to them. Despite these limitations, our study demonstrated higher rate of behavioral and emotional problems in children with ASD who had psychiatric comorbidities and a correlation between emotion regulation and behavioral problems. We recommend that clinicians should pay more attention to psychiatric comorbidities in children with ASD who presented with emotional difficulties or behavioral problems.

CONCLUSION

Children with ASD and comorbid psychiatric disorders had poorer ability to regulate emotions and more behavioral problems than those without comorbidities. We also found that poor emotion regulation predicts behavioral problems. Therefore, we suggest that all children with ASD should be evaluated for psychiatric comorbidities, especially those who show extensive behavioral problems and poor emotional regulation that exceeds what is typically seen in children with ASD alone. Interventions targeting emotion regulation in these children should also be further studied.

ACKNOWLEDGEMENTS

This study received funding from the Graduate Scholarship (Type II), Faculty of Medicine Siriraj Hospital, Mahidol University. The authors gratefully acknowledge to Assistant Professor Thanayot Sumalrot of the Department of Psychiatry and to Suthipol Udompunturak of the Division of Clinical Epidemiology, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University for the assistance with statistical analysis.

Conflicts of interest: No potential conflict of interest was reported by the author(s).

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders : DSM-5. Arlington, VA: American Psychiatric Association; 2013.
2. Van Steensel FJ, Bögels SM, de Bruin EI. Psychiatric Comorbidity in Children with Autism Spectrum Disorders: A Comparison with Children with ADHD. J Child Fam Stud. 2013;22(3):368-76.
3. Joshi G, Petty C, Wozniak J, Henin A, Fried R, Galdo M, et al. The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: a large comparative study of a psychiatrically referred population. J Autism Dev Disord. 2010;40(11):1361-70.
4. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric Disorders in Children With Autism Spectrum Disorders: Prevalence, Comorbidity, and Associated Factors in a Population-Derived Sample. J Am Acad Child Adolesc Psychiatry. 2008;47(8):921-9.
5. Yerys BE, Wallace GL, Sokoloff JL, Shook DA, James JD, Kenworthy L. Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. Autism Res. 2009;2(6):322-33.
6. Belardinelli C, Raza M, Taneli T. Comorbid Behavioral Problems and Psychiatric Disorders in Autism Spectrum Disorder. J Child Dev Disord. 2016;2:11. doi: 10.4172/2472-1786.100019
7. Kerns CM, Kendall PC, Zickgraf H, Franklin ME, Miller J, Herrington J. Not to be overshadowed or overlooked: functional impairments associated with comorbid anxiety disorders in youth with ASD. Behav Ther. 2015;46(1):29-39.
8. Gonzalez MR, Lopez IPV. Impact Of Emotional Disorders In The Functionality Of Children And Adolescents With Autism Spectrum Disorders Review. J Child Dev Disord. 2016;2:1-11.
9. Avni E, Ben-Itzchak E, Zachor DA. The Presence of Comorbid ADHD and Anxiety Symptoms in Autism Spectrum Disorder: Clinical Presentation and Predictors. Front Psychiatry. 2018;9:717.
10. Gupta S, Caskey A, Soares N, Augustyn M. Autism Spectrum Disorder and Mental Health Comorbidity Leading to Prolonged Inpatient Admission. J Dev Behav Pediatr. 2018;39(6):523-5.
11. Farrugia S, Hudson J. Anxiety in Adolescents With Asperger Syndrome: Negative Thoughts, Behavioral Problems, and Life Interference. Focus Autism Other Dev Disabl. 2006;21(1):25-35.
12. Mazefsky CA, Herrington J, Siegel M, Scarpa A, Maddox BB, Scahill L, et al. The role of emotion regulation in autism spectrum disorder. J Am Acad Child Adolesc Psychiatry. 2013;52(7): 679-88.
13. Maskey M, Warnell F. Parr JR, Couteur AL, McConachie H. Emotional and Behavioral Problems in Children with Autism Spectrum Disorder. J Autism Dev Disord. 2013;43:851-9.
14. Berkovits L, Eisenhower A, Blacher J. Emotion Regulation in Young Children with Autism Spectrum Disorders. J Autism Dev Disord. 2017;47(1):68-79.
15. Berenguer C, Rosello-Miranda B, Colomer C, Baixauli I, Miranda A. Children with autism and attention deficit hyperactivity disorder. Relationships between symptoms and executive function, theory of mind, and behavioral problems. Res Dev Disabil.2018; 83: 260-9.
16. Shields A, Cicchetti D. Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. Dev Psychol. 1997;33(6):906-16.
17. Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry. 1997;38(5):581-6.
18. Wongpiromsarn Y, Wipulakorn P, Nuanmanee S, Woener W, Mongkol A. Strengths and Difficulties Questionnaire (SDQ) Thai improved version: change and administration. J Ment Health Thai. 2011;19:128-34.
19. Woerner W, Nuanmanee S, Becker A, Wongpiromsarn Y, Mongkol A. Normative data and psychometric properties of the Thai version of the Strengths and Difficulties Questionnaire (SDQ). J Ment Health Thai. 2011; 19:42-57.
20. Craig F, Lamanna AL, Margari F, Matera E, Simone M, Margari L. Overlap Between Autism Spectrum Disorders and Attention Deficit Hyperactivity Disorder: Searching for Distinctive/ Common Clinical Features. Autism Res. 2015;8(3):328-37.
21. Gargaro BA, Rinehart NJ, Bradshaw JL, Tonge BJ, Sheppard DM. Autism and ADHD: How far have we come in the comorbidity debate? Neuroscience & Biobehavioral Reviews. Neurosci Biobehav Rev.2011;35(5):1081-8.
22. Leitner Y. The co-occurrence of autism and attention deficit hyperactivity disorder in children - what do we know? Front Hum Neurosci. 2014;8:268.
23. Jang J, Matson JL, Williams LW, Tureck K, Goldin RL, Cervantes PE. Rates of comorbid symptoms in children with ASD, ADHD, and comorbid ASD and ADHD. Res Dev Disabil. 2013;34(8): 2369-78.
24. Waddington F, Hartman C, de Bruijn Y, Lappenschaar M, Oerlemans A, Buitelaar J, et al. An emotion recognition subtyping approach to studying the heterogeneity and comorbidity of autism spectrum disorders and attention-deficit/hyperactivity disorder. J Neurodev Disord. 2018;10(1):31.
25. Guttmann-Steinmetz S, Gadow KD, Devincent CJ. Oppositional defiant and conduct disorder behaviors in boys with autism spectrum disorder with and without attention-deficit hyperactivity disorder versus several comparison samples. J Autism Dev Disord. 2009;39(7):976-85.

Assessment of Volume Status in Chronic Hemodialysis: Comparison of Lung Ultrasound to Clinical Practice and Bioimpedance

Kornchanok Vareesangthip, M.D.*, Banthita Thanapattaraborisuth, M.D.*, Kullanuch Chanchairujira, M.D., Suwimon Wonglaksanapimon, M.D.**, Thawee Chanchairujira, M.D.***

**Renal Division, Department of Medicine, **Department of Radiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*

ABSTRACT

Objective: Lung ultrasonography (LUS) has recently been used to evaluate extravascular lung water, and shown to be able to predict all-cause mortality in hemodialysis (HD) patients. This study aim to compare LUS with other volume assessment methods, and to verify the prognostic value of LUS in Thai chronic HD patients.

Materials and Methods: We conducted a prospective cohort study in 36 chronic HD patients. Volume status before the HD session was evaluated by physical examinations, bioimpedance analysis (BIA), and ultrasound lung comets (ULCs). Mortality and morbidities were recorded during a 1-year follow-up period.

Results: The degree of lung fluid accumulation was assessed by summation of the number of ULCs, and was classified into 3 groups: mild-to-moderate (ULC<15–29), severe (ULC=30–59), and very severe (ULC≥60) in 11.1%, 77.8%, and 11.1% of the patients, respectively. Either clinical edema or lung crackle had low sensitivity (20-32%) to detect extravascular lung water excess in patient with mild-to-moderate ULC and severe ULC. Overhydration assessed by BIA was found in 75% and 64.3% of patients with mild-to-moderate and severe ULC, respecively. In patients with very severe ULC, the admission rate due to volume overload was significantly higher, there was also a trend of increased mortality, as well as intradialytic complications.

Conclusion: Clinical assessment and BIA have limited value in determining extravascular fluid excess in the lung. Lung ultrasound is a useful tool to detect subclinical pulmonary congestion. The long-term outcome by using LUS-guided fluid management needs larger population studies.

Keywords: Lung ultrasonography; extravascular lung water; hemodialysis; all-cause mortality; morbidities; intradialytic complications (Siriraj Med J 2023; 75: 224-233)

INTRODUCTION

The prevalence of end-stage kidney disease (ESKD) patients requiring renal replacement therapy (RRT) is rising annually. From the annual report of renal replacement therapy (RRT) in Thailand in 2019, the prevalence of RRT is almost 2,300 persons per million population, with hemodialysis (HD) as the most chosen mode of RRT.[1] Fluid retention and high inter-dialytic weight gain in HD patients were reported in previous studies as among the important predictive factors for mortality and cardiovascular morbidity.[2,3] The accumulation of fluid in the lung is considered a major consequence of fluid overload and cardiovascular complications. Routine examinations for the assessment of fluid overload, including by history taking, blood pressure measurement, peripheral edema, and lung auscultation, have been shown to have poor diagnostic accuracy in the detection of interstitial lung edema or total body fluid accumulation.[4]

Corresponding author: Thawee Chanchairujira
E-mail: thaweechan@hotmail.com
Received 22 October 2022 Revised 15 January 2023 Accepted 16 January 2023
ORCID ID:http://orcid.org/0000-0001-7692-2560
https://doi.org/ 10.33192/smj.v75i3.261016

Several new objective methods, including bioimpedance analysis (BIA) and lung ultrasound (LUS), have been introduced and validated in chronic HD patients for the assistance of the volume status evaluation along with the routine physical examinations. Bioimpedance analysis is a noninvasive tool that has been applied in the setting of dialysis patients to evaluate the patient's body composition for nutritional assessment and quantitative determination of the total body water and extracellular fluid volume.[5,6] Water in the lungs can be inversely correlated with the systolic and diastolic functions from echocardiography but shows only a weak correlation with the fluid status from BIA.[10] Lung ultrasonography was initially used for extravascular lung water evaluation, mostly in critical care situations.[7,8] The degree of accumulation of fluid in the lung interstitial was quantified by assessing the ultrasound B-lines or ultrasound lung comets (ULCs), characterized by a comet-tail-like hyperechoic line continuing from the ultrasound transducer to the visceral pleural line. This was later adapted in a chronic HD setting, and was found to be associated with fluid loss during HD.[9] In the LUST study, Torino *et al.* found that the detection of extravascular lung water accumulation in chronic HD patients by using the ultrasound lung comet score had a higher sensitivity than by routine physical examination, which included lung auscultation and peripheral edema.[4] Fluid assessment by using lung ultrasound was also shown to be able to predict mortality in HD patients.[11,12] However, lung ultrasonography has not yet been studied in Thai chronic HD patients despite its many advantages seen in other studies worldwide. The aim of this study was to evaluate abilities of clinical assessment and bioimpedance for determining extracellular fluid excess comparing with lung ultrasonography, and to evaluate the prognostic value significance of the presence of ultrasound lung comets in Thai chronic HD populations.

MATERIALS AND METHODS

Study population

This single-center prospective cohort study was conducted in the HD units of Siriraj Hospital, Bangkok from July 2017 to March 2018. The inclusion criteria were patients who had been on three-times weekly chronic HD for more than 3 months with a stable dry weight, defined as target dry weight changes within ±0.5 kg in the past 3 months. The exclusion criteria were age under 18 years old, patients with current systemic infections, terminal cancers, lung diseases that may interfere with lung ultrasonography interpretation (e.g., lung fibrosis, interstitial lung disease, pleural effusion, patchy infiltration and consolidation), recent hospitalization within the past 3 months, patients with a cardiac pacemaker or defibrillator implantation, and a history of limb amputation. Of the total of 70 patients screened for enrollment, 36 patients were included in this study, while 29 patients denied informed consent, and 5 patients were excluded because of pleural effusion (n = 1), current systemic infection (n = 1), and current hospitalization (n = 3).

Study design

The patients' baseline characteristics, including age, sex, body mass index (BMI), comorbidities, dialysis vintage, blood pressure, ultrafiltration rate (UFR), and laboratory data, were collected. In each participant, all 3 methods of volume status assessment, namely clinical evaluation, lung ultrasonography, and BIA, were evaluated pre-HD in the same day at the beginning of the week of the HD session (after the longest interdialytic interval). The participants were followed up for 1 year for hospitalizations, all-cause mortality, and intradialytic complications.

Clinical evaluation of the volume status

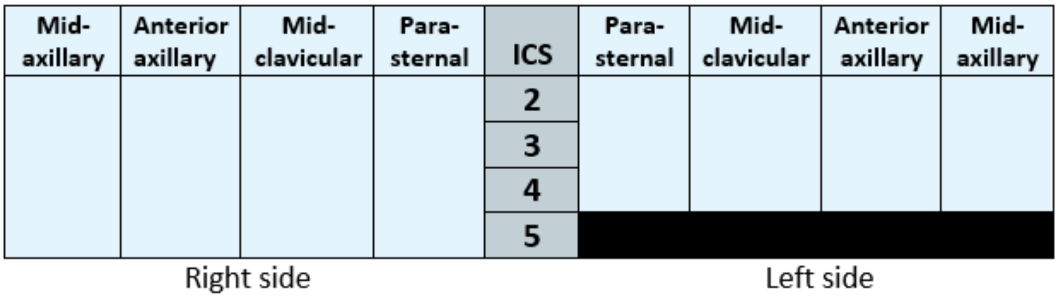
Clinical evaluation of the volume status included physical examination of clinical edema and lung auscultation. Clinical edema was examined at the mid-portion of both legs and classified according to the following scale: 1, no clinical edema; 2, more than 0 to 2 mm depth pitting; 3, more than 2 to 4 mm depth pitting; 4, more than 4 to 6 mm depth pitting; 5, more than 6 mm depth pitting.[4]

Lung auscultation was examined at the anterior and posterior sites of each hemithorax in a sitting position. Participants were asked to perform slow and deep inhalations during the evaluation for crackles. Lung crackle findings was classified according to the following scale: 1, no crackles; 2, uncertain about crackles; 3, fine crackles at the basal lungs; 4, half of the hemithorax crackles; 5, bilateral diffuse crackles.[4,13]

Lung ultrasonography

Lung ultrasonography (using an ultrasound machine [GE Logiq E9] with a cardiac probe 3–9 MHz sector scan probe; "S4–10" was performed by 2 radiologists who were blinded from the results of the other volume status measurement methods. The ultrasound scanning was performed with the patient in a supine position in 8 areas of the chest wall (including the mid-axillary, anterior axillary, mid-clavicular, and parasternal areas) in the 2^{nd}, 3^{rd}, 4^{th}, and 5^{th} intercostal spaces at both the right and left hemithorax, as shown in Fig 1A. The degree of lung fluid accumulation was measured as the number of ultrasound lung comets (ULCs) (Fig 1B). The ULC score for each patient was quantified by summation of the ULCs found in all 8 zones.[14] The ULC score was divided into 4 categories: mild (< 15 ULCs), moderate

A



B

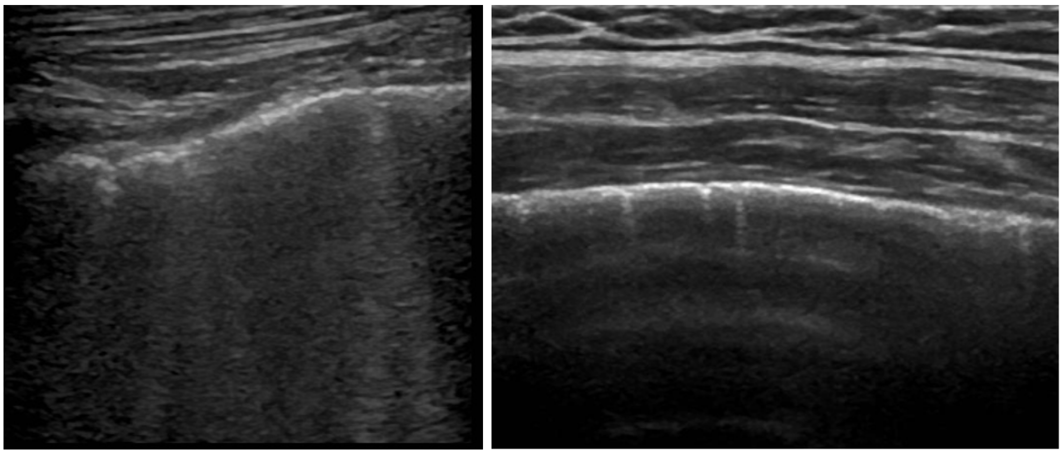


Fig 1A. Division of the chest wall into 8 areas for lung ultrasonography evaluation.
Fig 1B Ultrasound B-lines or ultrasound lung comets (ULCs), characterized by a comet-tail-like hyperechoic line continuing from the ultrasound transducer to the visceral pleural line *(Figure provided by Asso.Prof.Suwimon Wonglaksanapimon, Radiology Department, Siriraj Hospital).*

(15–29 ULCs), severe (30–59 ULCs), and very severe (> 60 ULCs).[4,11,12,14]

Bioimpedance analysis

Whole-body bioimpedance analysis (BIA) was used to measure the volume status by multi-frequency bioimpedance spectroscopy. Electrodes were attached to the participant's wrist on the non-vascular access side of the body and ipsilateral ankle in the supine position. Extracellular fluid volume was calculated as the percentage of overhydration compared with the relative hydration status in normal populations (ΔHS). The patients' hydration status were classified into 2 groups: 1) normohydration, ΔHS < 15%, and 2) hyperhydration, ΔHS > 15%.[12]

Echocardiography

Echocardiography, as performed according to the recommendations of the American Society of Echocardiography[15], was evaluated in all available patients within 3 months of the lung ultrasound and BIA. The staff performing the echocardiography were unaware of the lung ultrasound and BIA results.

Outcomes

The primary outcome was to compare ultrasound lung comets with physical examinations (clinical edema and lung crackle) and BIA in the evaluation of fluid overload. The secondary outcomes were: 1) to determine the sensitivity and specificity of physical examinations of clinical edema with BIA, 2) to study the associations of extravascular lung water evaluated by lung ultrasonography and 1-year all-cause mortality and cardiovascular events, as well as intradialytic complications, 3) to compare lung ultrasonography findings with echocardiography findings.

Intradialytic hypotension and intradialytic hypertension episodes were collected as the percentage of the complicated sessions to all hemodialysis sessions. Definitions from KDIGO were used: intradialytic hypotension was defined as any symptomatic decrease in systolic blood pressure (SBP) or a nadir intradialytic SBP < 90 mmHg, while intradialytic hypertension was defined as any rise of >10 mmHg from pre- to post-dialysis in the hypertensive range (>140 mmHg).[16]

Statistical analysis

We estimate the minimum sample size required, based on the prevalence of fluid overload in chronic hemodialysis patients and the sensitivity of physical examinations (clinical edema and lung crackles) with the lung ultrasonography as the gold standard (p-value, is set to be less than 0.05). The sensitivity of the physical examinations from the LUST study was 26%[4], we set the maximum marginal error of estimate to be 15% for constructing confidence interval of true value of sensitivity with the prevalence of fluid overload in our population of 50%, the total sample size of the sensitivity determines 972 patients. Continuous data were presented as the mean ± SD for normally distributed data and as the median with interquartile range (IQR) for non-normally distributed data. Categorical data were presented as the percent frequency. Comparisons among groups were made by p-value using one-way ANOVA analysis. The sensitivity and specificity of lung auscultation and clinical edema were also calculated. Statistical analysis was performed using SPSS software version 20.

This study was approved by the Human Research Protection Unit, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (COA no. Si 412/2017), and was conducted according to the principles of the Declaration of Helsinki. All participants received a detailed explanation of the study and gave their written informed consent.

RESULTS

We report this preliminary result because the prolonged COVID-19 pandemic halted further enrollment in our study. The patients' baseline characteristics and biochemical data are shown in Table 1. The mean age of the patients was 63 years and 55.6% were male. The comorbidities included diabetes mellitus (41.7%), myocardial infarction (22.2%), and a history of coronary artery bypass graft (CABG) or stent (25%). Since there were only 4 patients in the mild and moderate group, we divided the patients into 3 groups according to the severity of ULC score as per the following; mild-to-moderate 11.1% (n = 4), severe 77.8% (n = 28), and very severe 11.1% (n = 4). A history of myocardial infarction and history of CABG or stent was found more frequently in the very severe and severe ULC groups, but the difference did not reach statistical significance. Patients in the very severe ULC group had a higher New York Heart Association (NYHA) functional class than in the other less severe ULC groups. No differences in the baseline characteristics data were found, apart from for gender, in which the proportion of males was found to be higher in the moderate and severe ULC groups.

Comparison of fluid overload assessement by lung ultrasonography with clinical evaluation and BIA

Fluid overload was defined by using the following criteria: BIA overhydration defined as ΔHS > 15%, clinical edema positive score ≥ 2, and lung auscultation positive score ≥ 3. The severity of lung congestion assessed by ULC score was compared with physical examinations and BIA (Table 2). Clinical edema was found in 75% of the patients in the very severe ULC group, whereas it was found in only 25% and 21.4% of the patients in the mild-to-moderate and severe ULC groups, respectively. Lung crackle was positive in 100% of the patients with very severe ULC, whereas it was positive in only 25% and 32.1% of the patients with mild-to-moderate and severe ULC (p = 0.03), respectively. When compared with lung ultrasound, lung crackles, either alone or in combination with clinical edema, poorly reflected lung congestion in patients with mild-to-moderate ULC and severe ULC. When comparing lung ultrasonography with BIA, hyperhydration (ΔHS >15%) was found in all patients with very severe ULC, whereas it was found in 75% and 64.3% of patients with mild-to-moderate and severe ULC, respecively. Overall, the vast majority of clinical assessments of fluid overload were not sensitive for detecting lung congesiton, especially in patients with mild-to-moderate and severe ULC. When lung ultrasonography was used as the gold standard for the assessment of extravascular fluid excess in the lung, it was found that clinical assessment and BIA had a limited value for determing pulmonary congestion.

Comparison of fluid overload assessment by clinical evaluation with BIA

The comparison of the assessment of fluid overload by clinical evaluation with BIA is reported in Table 3. Overall, 25 patients (69.4%) were classified by BIA as overhydration. When using BIA as gold standard for diagnosed fluid overload, the sensitivity of detection of fluid overload by clinical edema and by lung crackle was 25% and 45.8%, respectively. The detection of fluid overload by either clinical edema or lung crackle increased the sensitivity to only 50%. Whereas, the presence of both clinical edema and lung crackle had a specificity of 83.3% to detect fluid overload, but the sensitivity was decreased to only 20.8%. Compared with BIA assessment, lung crackles and clinical edema, either alone or in combination, poorly reflected fluid overload, with an accuracy of 39%–55.6%.

Comparison between the ultrasound lung comet score and echocardiographic findings

Echocardiogarphic findings in 25 patients were classified into 3 groups according to the severity of the ULC score, and the results were compared between the

TABLE 1. Patients' baseline characteristics and biochemical data according to the number of ultrasound lung comets.

Participant characteristics	ULC numbers				
	All	Mild-to-Moderate 15–29	Severe 30–59	Very severe ≥ 60	P-value
Number of patients	36	4	28	4	
Age, years	63±13	53±4	64±13.8	66±9	0.242
Male sex, %	55.6	75.0	60.7	0.0	0.031*
BMI, kg/m^2	24.9±6.2	26.1±7.2	24.7±6.8	24.8±3.3	0.916
Dialysis vintage, mos	105.5±74.2	101.3±45.9	106.7±80.1	101.5±66.3	0.985
Diabetes, %	41.7	25.0	42.9	50.0	0.861
Ex-Smoker, %	27.8	25.0	32.1	0.0	0.581
Myocardial infarction %	22.2	0	21.4	50.0	0.23
CABG/ stent %	25	0	25	50	0.26
NYHA class I, %	36.1	25.0	39.2	25.0	0.537
II, %	44.4	75.0	42.9	25.0	
III, %	19.4	0.0	17.9	50.0	
SBP, mmHg	152±30	156±15	152±28	142±53	0.857
DBP, mmHg	77±22	88±15	76±22	78±25	0.600
PP, mmHg	74±20	68±20	76±18	64±30	0.480
HR, bpm	79±14	68±14	78±12	90±17	0.062
Hemoglobin, g/dL	10.5±1.2	10.5±0.6	10.6±1.3	10.3±0.9	0.866
Serum albumin, g/dL	4.1±0.3	4.4±0.3	4.1±0.3	4.1±0.6	0.141
Serum BUN, mg/dL	65±22	62±15	67±22	54±21	0.53
Serum Cr, mg/dL	9.6±3.3	10.9±1.8	9.7±3.4	8.0±3.4	0.46
Serum [Na^+], mmol/L	138±4	138±4	138±4	135±4	0.37
Serum [K^+], mmol/L	4.4±0.5	4.4±0.3	4.3±0.5	4.7±0.4	0.33
Ferritin, µg/L	649 (365, 929)	806 (595, 964)	649 (311, 929)	448 (379,1014)	0.68
Calcium, mg/dL	8.9±0.7	9.2±0.6	8.8±0.7	9.1±1.2	0.474
Phosphate, mg/dL	4.5±1.3	4.7±0.9	4.6±1.3	4.0±1.7	0.705
iPTH, ng/mL	213 [71–435]	361[149–646]	203 [83–395]	198 [33–534]	0.789
LDL–C, mg/dL	81 [64–104]	77 [59–83]	81 [65–104]	94 [43–131]	0.762
Equilibrated Kt/V	2.0±0.4	1.8±0.5	2.0±0.4	1.9±0.3	0.562

Abbreviations: body mass index (BMI), coronary artery bypass graft (CABG), New York Heart Association (NYHA), pulse pressure (PP), heart rate (HR), low density lipoprotein–cholesterol (LDL–C).
* Statistically significant: P-Value compared between group.

TABLE 2. Comparison of lung ultrasonography with clinical assessment and BIA for fluid overload.

Parameters compared with lung ultrasonography	Ultrasound lung comet numbers Mild-to-Moderate 15–29 (n=4)	Severe 30–59 (n=28)	Very severe ≥ 60 (n=4)	P-value
Clinical edema positive[††]	25.0%	21.4%	75.0%	0.11
Lung auscultation positive[†††]	25.0%	32.1%	100.0%	0.03*
Clinical edema and lung auscultation positive	25.0%	10.7%	75.0%	0.02*
BIA hyperhydration[†]	75.0%	64.3%	100.0%	0.568

[†]BIA hyperhydration, ΔHS > 15%, [††]Clinical edema positive score > 2, [†††]Lung auscultation positive, score > 3.
*Statistically significant: P-Value compared between group.

TABLE 3. Comparison of the assessment of fluid overload by clinical evaluation with BIA.

Compared parameters with BIA hyperhydration[†]	Accuracy*	Sensitivity*	Specificity*
Clinical edema positive[††]	38.9 (23.0-54.8)	25.0 (7.7-42.3)	66.7 (40.0-93.3)
Lung auscultation positive[†††]	55.6 (39.3-71.8)	45.8 (25.9-65.8)	75.0 (50.5-99.5)
Clinical edema or lung auscultation positive	52.8 (36.5-69.1)	50.0 (30.0-70.0)	58.3 (30.4-86.2)
Clinical edema and lung auscultation positive	41.7 (25.6-57.8)	20.8 (4.6-37.1)	83.3 (62.2-104.4)

[†]BIA hyperhydration, ΔHS > 15%, [††]Clinical edema positive score > 2, [†††]Lung auscultation positive, score > 3.
*Presented as percentage and 95% confidence interval

3 groups, as shown in Table 4. The number of ultrasound lung comets showed a moderate negative correlation with the left ventricular ejection fraction (LVEF) ($r = -0.48$, $P = 0.02$) (Table 5). The mean LVEF in patients with very severe ULC was 37.6±23.4%, which was significantly lower than those in the patients with mild-to-moderate and severe ULC ($p = 0.03$). Other echocardiographic parameters such as the mean left ventricular mass index (LVMI), mean left atrial volume index, and mean pulmonary arterial pressure (mPAP) were not correlated with the number of ultrasound lung comets.

All-cause mortality, hospitalization, and intradialytic complications

After a 1 year follow-up period, 3 deaths occurred (one patient in each ULC group), and the causes of all the deaths were infections (Table 6). Patients with very severe ULC had a significantly higher admision rate due to fluid overload. There was no significant difference in all-cause admission, admission due to infection, and admission due to cardiovascular disease (CVD) between the 3 groups. Intradialytic blood pressure, interdialytic weight gain, and ultrafiltration rate were not statistically

TABLE 4. Comparison of the cardiac echocardiography findings in each ULC group.

Variables	ULC numbers			P-value
	Mild-to-moderate 15–29 (n=2)	Severe 30–59 (n=20)	Very severe ≥ 60 (n=3)	
LV end-diastolic volume index, mL/m^2	50.7±16.1	63.5±20.9	53.1±19.4	0.55
LV end-systolic volume index, mL/m^2	22.6±12.7	25.0±14.3	29.0±25.9	0.89
LVEF, %	57.1±11.3	62.0±12.5	37.6±23.4	0.03*
LVMI, g/m^2	130.0±14.1	133.1±36.9	142.0±46.9	0.92
LA volume index, mL/m^2	34.9±12.9	46.3±15.7	52.7±23.9	0.51
mPAP, mmHg	27.6±17.6	28.8±11.3	36.7±1.8	0.65
RAP, mmHg	10.0±7.1	7.8±3.1	11.67±2.9	0.17
RVSP, mmHg	36.9±14.8	50.4±19.2	52.2±17.9	0.62

Abbreviations: left ventricle (LV), LV ejection fraction (LVEF), LV mass index (LVMI), left atrium (LA), mean pulmonary arterial pressure (mPAP), right atrial pressure (RAP), right ventricular systolic pressure (RVSP).
*Statistically significant: P-Value compared between group.

TABLE 5. Correlation between ultrasound lung comets and echocardiographic parameters.

Variables	Pearson's correlation coefficient	P-value
LV end-diastolic volume index, mL/m^2	0.09	0.65
LV end-systolic volume index, mL/m^2	0.24	0.27
LVEF, %	-0.48	0.02*
LVMI, g/m^2	0.26	0.22
LA volume index, mL/m^2	0.21	0.32
mPAP, mmHg	0.37	0.16
RAP, mmHg	0.27	0.23
RVSP, mmHg	0.25	0.33

Abbreviations: left ventricle (LV), LV ejection fraction (LVEF), LV mass index (LVMI), left atrium (LA), mean pulmonary arterial pressure (mPAP), right atrial pressure (RAP), right ventricular systolic pressure (RVSP).
*Statistically significant: P-Value compared between group.

different between the 3 groups. Patients in the very severe and severe ULC groups tended to have more frequent episodes of intradialytic hypotension and intradialytic hypertension; however, this did not reach statistical significance. (supplement Table 1)

DISCUSSION

The present study showed that when BIA was used as the gold standard, routine physical examinations, either clinical edema or lung auscultation, had low sensitivity (25% and 45.8%, respectively) to detect extracellular

TABLE 6. All-cause mortality and morbidities according to the severity of ultrasound lung comets.

Outcomes	ULC numbers			P-value
	Mild-to-moderate 15–29 (n=4)	Severe 30–59 (n=28)	Very severe ≥ 60 (n=4)	
All admissions	1 (25%)	14 (50%)	2 (50%)	0.64
Admission due to infection	1 (25%)	7 (25%)	2 (50%)	0.58
Admission due to CVD	0	1 (3.6%)	1 (25%)	0.19
Admission due to fluid overload	0	1 (3.6%)	1 (25%)	0.02*
Dead	1 (25%)	1 (3.6%)	1 (25%)	0.15

Abbreviations: ultrasound lung comet (ULC), cardiovascular disease (CVD).
*Statistically significant: P-Value compared between group.

volume excess, and had specificities of 66.7% and 75%, respectively. Although the presence of both clinical edema and lung auscultation had high specificity (83%) in detecting fluid overload, the sensitivity was only 20.8%. When lung ultrasonography was used as the gold standard, the sensitivity of lung auscultation to detect extravascular lung fluid excess was 32% and 25% in the patients with severe ULC and mild-to-moderate ULC, respectively. Lung crackle was found to be associated with high accuracy for evaluating pulmonary congestion only in patients with very severe ULC. Similary, clinical edema was found in only 25% and 21.4% of the patients in the mild-to-moderate and severe ULC groups, respectively. Our findings were consistent with Torino *et al.*'s study, which reported a sensitivity of 10%–30% for peripheral edema and/or pulmonary crackles in the detection of extravascular water in ESKD patients as compared with lung ultrasonography.[4] Likewise, another meta-analysis compared these two clinical parameters in the diagnosis of volume overload in patients presenting with dyspnea, compared with radio-isotropic study, which revealed a sensitivity of 50%–60%.[17] Although lung auscultation is a simple method, when compared with lung ultrasound it is evidently insensitive to detect extravascular lung water in many conditions, such as acute or chronic heart failure, critical care, and ESKD settings.[4,8,18,19] Lung ultrasound, however, may help in detecting subclinical extravascular lung water excess and may provide a useful monitoring tool for fluid management in HD patients.

Many studies have shown a correlation between extracellular fluid volume status assessed by BIA and extravascular lung water evaluated with lung ultrasound. Siriopol *et al.* found a significant correlation between the lung ultrasound congestion score and BIA-derived parameters[12], while Ngoh *et al.* found that more than 60% of chronic HD patients with a normohydration status had moderate or severe lung congestion on lung ultrasound.[20] Consistent with these previous studies, our results showed that all the patients with very severe ULC had a hyperhydration status as assessed by BIA; whereas, 25% and 35.7% of HD patients with a normohydration status had mild-to-moderate ULC (15–29) and severe ULC (30–59) as assessed by lung ultrasound, respectively. This difference may be due to the principles of the fluid compartmental volume assessment of each technique; whereby, whole-body BIA evaluates the extracellular fluid component, while lung ultrasound evaluates only extravascular lung water, which is largely attributable to the severity of the underlying cardiac dysfunction. This could explain why clinical and BIA assessment have limited value in determining extravascular fluid excess in the lung. The presence of excess extravascular lung water is an important contributor to a patient's symptoms, and is a major predictor of hospitalization and mortality in HD patients[21]; whereas, BIA-guided fluid management may not be associated with long-term survival.[12]

The prevalence of pulmonary congestion determined as severe or very severe ULCs in the present study was 88%, which was higher than the prevalent reported in previous studies (27%–32%)[4,12] The dissimilarity in these results might be associated with the longer dialysis vintage in the present study (mean 105.5±74.2 months) compared to the other studies (30–85 months). In patients with very severe ULC, LVEF was significantly lower than in the other less severe ULC groups, while the LVMI, left atrial volume index, mPAP, RAP, and RVSP tended to be higher. These echocardiographic findings are similar to previous studies[12,22] and it may be due to the direct effect of volume overload on the heart chambers, which may be associated with an increased risk of cardiovascular event-related mortality.

A previous study[24] demonstrated an independent predictive value of the severe lung ultrasound comet score for mortality and cardiac events. Our result showed that in patients with very severe ULC, the admission rate due to volume overload was significantly higher, and there was a trend of an increased mortality and all-cause admission rate at the 3-month and 1-year follow-up periods, as well as intradialytic complications. Our results did not demonstrate a strong correlation between the lung ultrasound comet score with mortality and morbidity as in previous studies[4,12], which could be explained by the small sample size of our study and hence the limited preliminary data.

There are some limitations in this study to note. First, we did not assess the inter- and intraobserver agreement in lung ultrasound findings. However, the lung ultrasound scan in this present study was performed by 2 radiologists with 15 years' experience of ultrasound procedure, who were blinded from the results of the other volume status measurement methods. Second, the measurement of ultrasound lung comets has some patient-dependent limitations (such as obesity, interstitial lung diseases), which may cause inaccuracy or false–positive ULC findings. We tried to combat this by excluding all possible known conditions that may interfere with image interpretation. Third, this preliminary result is limited by the small sample size, it needs a larger population study for provide more conclusive evidence of the benefit of fluid assessment by lung ultrasound and prognostic value significance of the presence of ULC in Thai chronic HD patients. However, this preliminary result shows the promising approach by using lung ultrasound to detect subclinical pulmonary congestion, and it may be a useful tool to guide management of extravascular lung water excess, which is a major factor for mortality and morbidities in HD patients.

CONCLUSION

Fluid overload assessed by physical examinations and BIA have limited value for early detecting extravascular lung water in HD patients. Lung ultrasound can be used as a noninvasive point-of-care tool for detecting subclinical pulmonary congestion and may provide semi-quantitative guided fluid management in HD patients. The benefits for the long-term outcomes regarding morbidity and mortality by using lung ultrasound-guided fluid management to adjust patients' dry weight need to be studied in a larger population.

ACKNOWLEDGEMENTS

We thank the physicians and nurses in both the Siriraj hemodialysis units of the Nephrology Department and Imaging Center of the Radiology Department for facilitating the participant inclusion, performing BIA and lung ultrasonography, and for support with the data collection during the study period.

REFERENCES

1. Lumpaopong A, Chuasuwan A. Thailand Renal Replacement Therapy Year 2016-2019. The Nephrology Society of Thailand. 2019:1-64.
2. Kalantar-Zadeh K, Regidor DL, Kovesdy CP, Van Wyck D, Bunnapradist S, Horwich TB, et al. Fluid retention is associated with cardiovascular mortality in patients undergoing long-term hemodialysis. Circulation. 2009;119(5):671-9.
3. Lee MJ, Doh FM, Kim CH, Koo HM, Oh HJ, Park JT, et al. Interdialytic weight gain and cardiovascular outcome in incident hemodialysis patients. Am J Nephrol. 2014;39(5):427-35.
4. Torino C, Gargani L, Sicari R, Letachowicz K, Ekart R, Fliser D, et al. The Agreement between Auscultation and Lung Ultrasound in Hemodialysis Patients: The LUST Study. Clin J Am Soc Nephrol. 2016;11(11):2005-11.
5. Kotanko P, Levin NW, Zhu F. Current state of bioimpedance technologies in dialysis. Nephrol Dial Transplant. 2008;23(3):808-12.
6. Yilmaz Z, Yildirim Y, Aydin FY, Aydin E, Kadiroglu AK, Yilmaz ME, et al. Evaluation of fluid status related parameters in hemodialysis and peritoneal dialysis patients: Clinical usefulness of bioimpedance analysis. Medicina. 2014;50(5):269-74.
7. Lichtenstein DA. Lung ultrasound in the critically ill. Ann Intensive Care. 2014;4(1):1.
8. Enghard P, Rademacher S, Nee J, Hasper D, Engert U, Jorres A, et al. Simplified lung ultrasound protocol shows excellent prediction of extravascular lung water in ventilated intensive care patients. Crit Care. 2015;19:36.
9. Vitturi N, Dugo M, Soattin M, Simoni F, Maresca L, Zagatti R, et al. Lung ultrasound during hemodialysis: the role in the assessment of volume status. Int Urol Nephrol. 2014;46(1):169-74.
10. Zoccali C, Tripepi R, Torino C, Bellantoni M, Tripepi G, Mallamaci F. Lung congestion as a risk factor in end-stage renal disease. Blood Purif. 2013;36(3-4):184-91.
11. Zoccali C, Torino C, Tripepi R, Tripepi G, D'Arrigo G, Postorino M, et al. Pulmonary congestion predicts cardiac events and mortality in ESRD. J Am Soc Nephrol. 2013;24(4):639-46.
12. Siriopol D, Hogas S, Voroneanu L, Onofriescu M, Apetrii M, Oleniuc M, et al. Predicting mortality in haemodialysis patients: a comparison between lung ultrasonography, bioimpedance data and echocardiography parameters. Nephrol Dial Transplant. 2013;28(11):2851-9.
13. Kataoka H, Matsuno O. Age-related pulmonary crackles (rales)

in asymptomatic cardiovascular patients. Ann Fam Med. 2008;6(3):239-45.
14. Frassi F, Gargani L, Tesorio P, Raciti M, Mottola G, Picano E. Prognostic value of extravascular lung water assessed with ultrasound lung comets by chest sonography in patients with dyspnea and/or chest pain. J Card Fail. 2007;13(10):830-5.
15. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, et al. Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. J Am Soc Echocardiogr. 2005;18(12):1440-63.
16. Flythe JE, Chang TI, Gallagher MP, Lindley E, Madero M, Sarafidis PA, et al. Blood pressure and volume management in dialysis: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int. 2020;97(5):861-76.
17. Wang CS, FitzGerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? JAMA. 2005;294(15):1944-56.
18. Gheorghiade M, Follath F, Ponikowski P, Barsuk JH, Blair JE, Cleland JG, et al. Assessing and grading congestion in acute heart failure: a scientific statement from the acute heart failure committee of the heart failure association of the European Society of Cardiology and endorsed by the European Society of Intensive Care Medicine. Eur J Heart Fail. 2010;12(5):423-33.
19. Platz E, Lewis EF, Uno H, Peck J, Pivetta E, Merz AA, et al. Detection and prognostic value of pulmonary congestion by lung ultrasound in ambulatory heart failure patients. Eur Heart J. 2016;37(15):1244-51.
20. Ngoh CLY, Teng HL, Chua YT, Leo CCH, Wong WK. Comparison between lung ultrasonography and current methods for volume assessment in Asian chronic hemodialysis patients. Hemodial Int. 2020;24(4):516-27.
21. Banerjee D, Ma JZ, Collins AJ, Herzog CA. Long-term survival of incident hemodialysis patients who are hospitalized for congestive heart failure, pulmonary edema, or fluid overload. Clin J Am Soc Nephrol. 2007;2(6):1186-90.
22. Mallamaci F, Benedetto FA, Tripepi R, Rastelli S, Castellino P, Tripepi G, et al. Detection of pulmonary congestion by chest ultrasound in dialysis patients. JACC Cardiovasc Imaging. 2010;3(6):586-94.
23. Loutradis C, Sarafidis PA, Ferro CJ, Zoccali C. Volume overload in hemodialysis: diagnosis, cardiovascular consequences, and management. Nephrol Dial Transplant. 20121; 36(12):2182-93.
24. Saad MM, Kamal J, Moussaly E, Karam B, Mansour W, Gobran E, et al. Relevance of B-lines on lung ultrasound in volume overload and pulmonary congestion: Clinical correlations and outcomes in patients on hemodialysis. Cardiorenal Med. 2018; 8(2): 83-91.

Predictive Factors of Intravesical Recurrence after Ureteroscopy in Upper Urinary Tract Urothelial Carcinoma Followed by Radical Nephroureterectomy

Sittiporn Srinualnad, M.D., Atichet Sawangchareon, M.D., Kantima Jongjitaree, M.D., Kittipong Phinthusophon, M.D., Tawatchai Taweemonkongsap, M.D., Sunai Leewansangtong, M.D., Chaiyong Nualyong, M.D., Karn Liangkobkit, M.D., Ekkarin Chotikawanich, M.D.

Division of Urology, Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

ABSTRACT

Objective: To investigate the risk factors of developing intravesical recurrence (IVR) in patients with upper urinary tract urothelial carcinoma (UTUC) who underwent ureterorenoscopy (URS) before radical nephroureterectomy with bladder cuff excision (RNU).

Materials and Methods: This retrospective study collected data from the medical records of patients diagnosed with UTUC between January 2012 and December 2019. All the patients underwent ureteroscopy before radical surgery. Patients previously diagnosed with bladder cancer were excluded. A total of 63 patients were included in the study. Tumour factors, such as multiplicity, location, size, histologic grade, pathologic T-stage, and lymphovascular invasion status, were evaluated. The type of endoscopic procedure and time interval between URS and RNU were analysed to determine the factors affecting IVR.

Results: The associated factors with IVR included multifocal tumours (HR = 4.8 (1.9–11.9)), large size tumours greater than or equal to 4 cm (HR = 3.3 (1.5–7.0)), and time interval greater than or equal to 5 weeks between URS and RNU (HR = 2.6 (1.2–5.5)). Factors including tumour location (kidney or ureter), size, grading, T-stage, and lymphovascular invasion as well as the type of endoscopic procedure were not at high risk for IVR.

Conclusion: The predictive factors of IVR for UTUC patients who underwent URS before RNU included a multiplicity of primary tumours and a tumour size greater than or equal to 4 cm, while a time interval between URS and RNU greater than or equal to 5 weeks increased the risk of IVR.

Keywords: Upper tract urothelial carcinoma; intravesical recurrent; bladder cancer; ureteroscopy; biopsy (Siriraj Med J 2023; 75: 234-240)

INTRODUCTION

Upper urinary tract urothelial carcinoma (UTUC) is an uncommon malignancy accounting for about 5–10% of urothelial carcinomas.[1-3] Radical nephroureterectomy with bladder cuff excision (RNU) is considered as the gold standard procedure for localised UTUC because it provides effective control and improves cancer-specific survival.[4,5]

A unique feature of urothelial carcinoma is the formation of multifocal tumours simultaneously and subsequently. Thus, the development of a tumour on the other side throughout the collecting system is possible in

Corresponding author: Ekkarin Chotikawanich
E-mail: ekkarinc@yahoo.com
Received 26 October 2022 Revised 27 January 2023 Accepted 1 February 2023
ORCID ID:http://orcid.org/0000-0003-3914-2770
https://doi.org/10.33192/smj.v75i3.261149

patients with primary UTUC.[6-8] The most common site of recurrence is the bladder. The intravesical recurrence (IVR) rate in patients who underwent RNU was found to be as high as 15–50%, while 5–10% of patients with IVR developed muscle invasive disease.[9,10] Thus, IVR is related to the risk of disease progression as well as an increased overall cost of treatment and decreased quality of life resulting from surveillance cystoscopy and bladder tumour surgery.[11]

Due to recent advancements in endoscopic technology, ureterorenoscopy (URS) is now increasingly used for the diagnosis and treatment of UTUC. This procedure is considered minimally invasive and pain-killers are not extensively required.[12] Several previous studies have reported that URS improved the accuracy of definitive diagnosis, staging, and histopathologic grading for UTUC.[13-15] URS has now become the method of choice to evaluate UTUC before performing definite surgery or for treating the disease following kidney sparing surgery.

However, there are many concerns that URS affects the long-term outcome of UTUC. Several studies have reported an association between URS and an increased risk of IVR but did not specify the risk factors involved.[13,16-21] Intravesical chemotherapy has been recommended for patients who underwent URS before RNU to prevent IVR, but this may not be universally suitable.[11,22,23] Therefore, here, the risk factors were evaluated for IVR in UTUC patients who underwent URS before RNU. The results may elucidate the patient choices available for future IVR prevention.

MATERIALS AND METHODS

This retrospective study reviewed the medical records of UTUC patients who underwent URS before RNU at a university hospital between January 2012 and December 2019. The study was approved by the ethics committee of the institute. (COA no. Si 648/2019) All the patients participated in a surveillance program that included physical and laboratory examination, cystoscopy, and imaging studies. None of the patients received intravesical chemotherapy before the first IVR. Patients who were previously diagnosed with bladder cancer, locally advanced UTUC, or metastatic disease at the first diagnosis and those who had absences in the surveillance program were excluded. Data from the medical records were collected as case record forms and classified as follows:

1. **Patient characteristics,** including gender, age at the first diagnosis, and follow-up interval.
2. **Disease factors** (final pathologic report of the RNU specimen), including:
   - i Multiplicity of the tumour (unifocal, multifocal)
   - ii Location of the tumour (renal calyx or pelvis, upper ureter, and lower ureter)
   - iii Tumour size; cumulative (mm)
   - iv Histologic grading
   - v Pathologic T-stage
   - vi Lymphovascular invasion.
3. **Procedural factors,** including:
   - i The endoscopic procedure (URS alone, URS with biopsy, URS with laser ablation, URS immediately followed by RNU)
   - ii Time interval between URS and RNU (weeks).
4. **Outcomes,** including the incidence of IVR, time interval between URS and the first IVR, and total number of IVR during surveillance.

The endoscopic procedure was performed utilising semi-rigid or flexible ureteroscopy. Tumour biopsy was performed using a basket or cup biopsy forceps, while laser ablation was performed utilising a holmium:YAG or thulium:YAG laser depending on the surgeon preference.

Radical nephroureterectomy was performed as different approaches, including open surgery (retroperitoneal through flank incision), laparoscopy (transperitoneal and retroperitoneal), and transperitoneal robotic-assisted. Distal ureter and bladder cuff excision for all the nephroureterectomy approaches was performed in one fashion through a low small transverse incision as a similar incision to extract the kidney. All the bladder cuff excisions were performed using an extravesical approach.

The patients' demographics, disease factors, and procedural factors were compared between the IVR and non-IVR patients. The independent t-test (normality) or Mann–Whitney U-test (non-normality) were used to assess quantitative data, while the chi-square test or Fisher's exact test were utilised for qualitative data. All the continuous data are shown herein as the median and interquartile range and an ROC curve was used to determine the appropriate cut-off value that showed a significant relationship with IVR. Factors with a p-value < 0.05 were considered statistically significant. Kaplan–Meier analysis was utilised to calculate the recurrence-free survival, while Cox-regression analysis was employed to predict the factors affecting IVR. Factors with a p-value < 0.05 in the univariate analyses were enrolled into the multivariate analyses.

RESULTS

Out of the 63 patients included in this study, 29 (46.0%) developed IVR during surveillance. The median time to develop the first IVR after the endoscopic procedure was 8.0 (4.5–11.5) months. The median number of total

IVR during the follow-up period was 2.0 (1.0–3.0).

The patient characteristics for both the non-IVR and IVR groups are shown in Table 1. There were no differences in the proportions between the two groups in terms of the gender, age at the first diagnosis, and follow-up interval (from the first diagnosis to the last visit).

Disease factors

Tumour multiplicity showed a significant difference between the non-IVR and IVR groups (p-value = 0.002) (Table 1). Comparing tumour multiplicity, 100% of patients who had multifocal primary tumours developed IVR, while the unifocal tumours did not significantly lead to the development of IVR among the various tumour sites (renal calyx or pelvis, proximal ureter, and distal ureter). Univariate analyses showed an increased risk of IVR for multifocal tumours (HR, 4.8; 95%CI, 1.9–11.9; p-value = 0.001) as well as in the multivariate analyses (HR, 4.3; 95%CI, 1.7–10.9; p-value = 0.002) (Table 2). Kaplan–Meier analysis revealed that the patients with unifocal tumours had a significantly higher rate of recurrence-free survival compared to patients with multifocal tumours (62.4 months vs 9.0 months, p-value < 0.005) (Fig 1).

Utilising the ROC curve to identify an appropriate cut-off value showed that a tumour size of 40 mm or larger had a higher rate of IVR than smaller tumours in both univariate analyses (HR, 3.3; 95%CI, 1.5–7.0; p-value = 0.002) and multivariate analyses (HR, 3.0; 95%CI, 1.4–6.5; p-value = 0.005). Patients with tumours smaller than 40 mm also had a significantly higher rate of recurrence-free survival compared to patients with tumours larger than 40 mm (68.8 months vs. 31.3 months, p-value < 0.001) (Fig 2).

Histologic grading, pathologic T-stage, and presenting with LVI did not affect the IVR (p-value = 0.72, 0.38, and 0.65, respectively).

Procedural factors

The time interval between URS and RNU demonstrated a significant difference between the non-IVR and IVR patients (p-value = 0.03). Utilising the ROC curve, a cut-off value of 5 weeks showed an association with IVR (p-value = 0.01). A time interval of 5 weeks or longer was related to an increased risk of IVR in the univariate analyses (HR, 2.6; 95%CI, 1.2–5.5; p-value = 0.01); however, the multivariate analyses did not show a significant association. Patients who underwent RNU less than 5 weeks after the first endoscopic intervention had a higher rate of recurrence-free survival (69.3 months vs. 35.7 months, p-value < 0.01) (Fig 3).

No difference in developing IVR was shown among the various types of initial endoscopic procedures (p-value = 0.36).

DISCUSSION

Although preoperative URS has been associated with an increased risk of IVR, this endoscopic procedure remains popular in the diagnosis and treatment of UTUC. Preoperative URS is crucial, especially in cases of controversial radiographic findings or in those considering kidney sparing surgery.[24] A previous study reported that 3% of patients who had suspected UTUC and who underwent RNU had benign pathologic findings. The suggestion to perform preoperative ureteroscopy undoubtedly improves precise decision-making before radical surgery.[25]

Many studies have demonstrated that IVR after endoscopic management of UTUC does not affect the long-term outcome and cancer-specific mortality. Gurbuz et al. reported that UTUC patients without preoperative URS had a similar 5-year survival rate compared with UTUC patients who received preoperative URS (77% vs. 73%, p-value = 0.4).[26] Sankin et al. reported a similar result, whereby patients with preoperative URS had a significantly higher IVR compared with patients without preoperative URS, but there were no significant differences in cancer-free survival and metastatic-free survival.[20] Therefore, IVR remains a factor to be considered when opting for URS.

Our results revealed that a multiplicity of tumours was a risk factor in developing IVR. Patients with multifocal tumours had 100% IVR and also a significantly lower recurrence-free survival rate compared with those suffering from unifocal tumours. Sung et al. also reported that multifocal tumours were a predictive factor of IVR after RNU[21], while Kang et al. reported a 3-fold greater risk of IVR in patients with multifocal tumours and recommended that these high-risk patients should be closely followed up.[6] There was no difference in risk of IVR among various tumour locations, including intrarenal, and proximal and distal ureter, which had unifocal tumours.

This study reported a correlation between the tumour size and an increased risk of IVR. Patients who had a tumour size of 40 mm or larger were at a 3-fold greater risk of developing IVR (p-value = 0.005). Shibing et al. reported that a tumour size larger than 3 cm was an adverse prognostic factor for cancer-specific survival, recurrence-free survival, and overall survival in UTUC patients who had undergone RNU with or without URS.[27]

Other previous studies reported that delaying radical surgery according to preoperative URS was associated with an increased risk of IVR. Lee et al. noted that UTUC

TABLE 1. Patient characteristics among the IVR and non-IVR groups.

Characteristic	*Non-IVR group*	*IVR group*	*P-value*
Patient demographics			
Gender ***n*** (%)			0.06
Male	15 (42.9%)	19 (67.9%)	
Female	20 (57.1%)	9 (32.1%)	
Age at the 1st diagnosis (year), median [IQR]	68 [61-76]	71 [64-79]	0.35
Follow-up interval (month), median [IQR]	45 [33-67]	48 [29-72]	0.98
Disease factors			
Multiplicity of the tumour ***n*** (%)			**0.002**
Unifocal tumour	35 (100.0%)	21(75.0%)	
Multifocal tumour	0 (0.0%)	7 (25.0%)	
Location of unifocal tumour ***n*** (%)			0.33
Renal calyx or pelvis	23 (65.7%)	10 (47.6%)	
Upper ureter	6 (17.1%)	4 (19.1%)	
Lower ureter	6 (17.1%)	7 (33.3%)	
Size of the tumour (mm), median [IQR]	28 [18-34]	34 [25-48]	**0.04**
Histologic grade ***n*** (%)			0.72
Low grade	6 (17.1%)	3 (10.7%)	
High grade	29 (82.9%)	25 (89.3%)	
Pathologic T-stage ***n*** (%)			0.38
T1 or lower	22 (62.9%)	14 (50.0%)	
T2	6 (17.1%)	9 (32.1%)	
T3 or higher	7 (20.0%)	5 (17.9%)	
LVI ***n*** (%)			0.65
Not present	33 (94.3%)	25 (89.3%)	
Present	2 (5.7%)	3 (10.7%)	
Procedural factors			
Initial endoscopic procedure ***n*** (%)			0.36
URS alone	5 (14.3%)	2 (7.1%)	
URS with biopsy	12 (37.1%)	17 (57.1%)	
URS with laser ablation	3 (8.6%)	3 (10.7%)	
URS followed by immediate-RNU	14 (40.0%)	7 (25.0%)	
Time interval between URS and RNU (week), median [IQR]	2 [1-7]	6 [2-10]	**0.03**

Abbreviations: IVR = intravesical recurrence; LVI = lymphovascular invasion; URS = ureteroreno-scopy; RNU = radical nephroureterectomy with bladder cuff excision.

TABLE 2. Univariate and multivariate analyses of the factors associated with intravesical recurrence.

Characteristic	*Univariate analyses*		*Multivariate analyses*	
	HR (95% CI)	***P-value***	***HR (95% CI)***	***P-value***
Location of tumour				
Unifocal tumour	1	-	1	-
Multifocal tumour	4.8 (1.9-11.9)	0.001	4.3 (1.7-10.9)	**0.002**
Size of tumour				
< 40 mm	1	-	1	-
≥ 40 mm	3.3 (1.5-7.0)	0.002	3.0 (1.4-6.5)	**0.005**
Time interval between URS and RNU+BCE				
< 5 weeks	1	-	1	-
≥ 5 weeks	2.6 (1.2-5.5)	0.01	1.7 (0.7-3.9)	0.19

Abbreviations: HR = hazard ratio; CI = confidence interval.

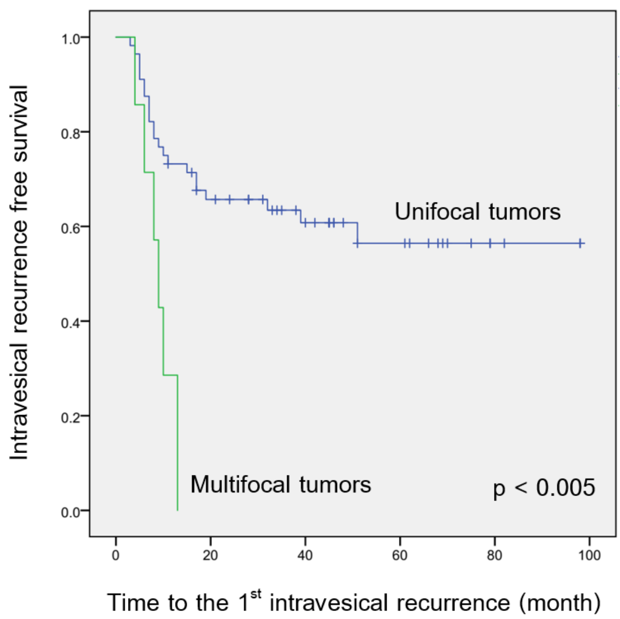


Fig 1. Intravesical free survival according to the location of the tumour

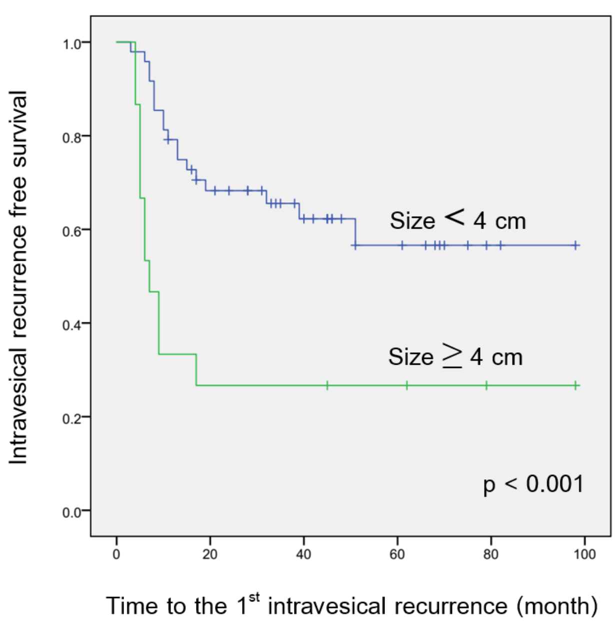


Fig 2. Intravesical free survival according to the size of the tumour

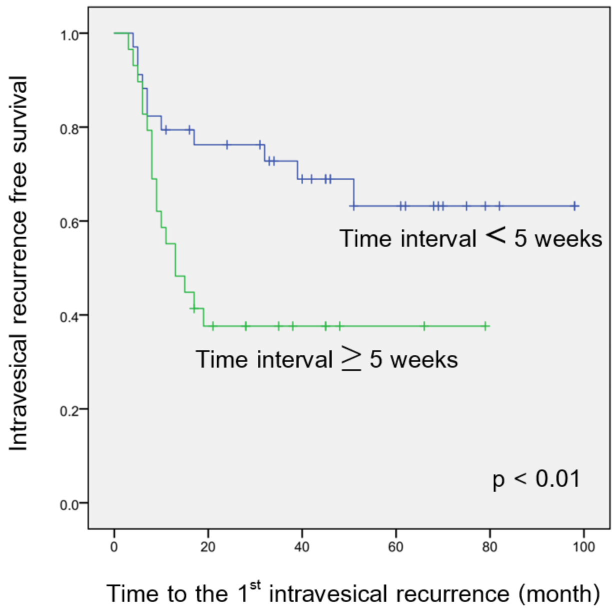


Fig 3. Intravesical free survival according to the time interval between URS and RNU

patients who underwent a 2-session approach (separate sessions of URS and RNU) had a significant risk of IVR compared with patients who followed a 1-session approach (URS and RNU in the same session) or without diagnostic URS (HR, 3.6; 95%CI, 1.0–12.6; p-value = 0.04).[28] However, they did not demonstrate an appropriate time interval between diagnostic URS and RNU that could reduce the risk of IVR. This study showed a significant association between delayed RNU and the risk of IVR at 5 weeks from the ROC curve. Univariate analyses showed that patients who had a time interval between URS and RNU of 5 weeks or longer were at a high risk of IVR (HR, 2.6; 95%CI, 1.2–5.5; p-value = 0.01); however, multivariate analyses did not show a significant association.

An investigation of the role of post-operative intravesical therapy for preventing IVR after RNU using the ODMIT-C trial demonstrated that post-operative single-dose intravesical Mitomycin-C (MMC) following RNU appeared to reduce the relative risk of IVR in the subsequent year by approximately 40%.[11] Wu et al. also reported a significantly lower IVR rate in patients who received either intravesical MMC or doxorubicin after RNU compared with those who did not.[10] These studies reported the efficacy of intravesical chemotherapy in patients who underwent RNU regardless of URS before radical surgery. According to the risk factors identified in this study, the administration of intravesical chemotherapy should be considered for patients who underwent URS before RNU and who had multifocal tumours, a tumour size greater than 40 mm, and a time interval between preoperative URS and RNU of more than 5 weeks. However, further study is required to verify these findings.

There are several limitations of this study to note, including the small number of patients. UTUC is an uncommon cancer and not every patient underwent URS before RNU and other inclusion criteria as aforementioned. This retrospective study also had inherent bias from the different techniques used for the surgical procedure, which depended on the surgeon preference.

CONCLUSION

In this study, the risk factors of developing IVR in UTUC patients who underwent URS before RNU were evaluated as being a multiplicity of primary tumours and a tumour size greater than or equal to 4 cm. A time interval between preoperative URS and RNU of greater than or equal to five weeks was associated with an increased risk of IVR in the univariate analyses.

Potential Conflicts of Interest

The authors declare they have no conflicts of interest.

REFERENCES

1. Munoz JJ, Ellison LM. Upper tract urothelial neoplasms: incidence and survival during the last 2 decades. J Urol. 2000;164(5): 1523-5.
2. Raman JD, Messer J, Sielatycki JA, Hollenbeak CS. Incidence and survival of patients with carcinoma of the ureter and renal pelvis in the USA, 1973-2005. BJU Int. 2011;107(7):1059-64.
3. Ristau BT, Tomaszewski JJ, Ost MC. Upper tract urothelial carcinoma: current treatment and outcomes. Urology. 2012; 79(4):749-56.
4. Roupret M, Babjuk M, Comperat E, Zigeuner R, Sylvester R, Burger M, et al. European guidelines on upper tract urothelial carcinomas: 2013 update. Eur Urol. 2013;63(6):1059-71.
5. Margulis V, Shariat SF, Matin SF, Kamat AM, Zigeuner R, Kikuchi E, et al. Outcomes of radical nephroureterectomy: a

series from the Upper Tract Urothelial Carcinoma Collaboration. Cancer. 2009;115(6):1224-33.
6. Kang CH, Yu TJ, Hsieh HH, Yang JW, Shu K, Huang CC, et al. The development of bladder tumors and contralateral upper urinary tract tumors after primary transitional cell carcinoma of the upper urinary tract. Cancer. 2003;98(8):1620-6.
7. Maneesuwansin S, Suk-ouichai C, Ramart P, Jitpraphai S, Phinthusophon K, Chotikawanich E, et al. Administration of Renin-Angiotensin System Inhibitor Affects Tumor Recurrence and Progression in Non-Muscle Invasive Bladder Cancer Patients. Siriraj Med J. 2019;71(1):31-7.
8. Nualyong C, Woranisarakul V, Tantranont N, Chotikawanich E, Shrestha S, Taweemonkongsap T. Metastatic Malignant Melanoma of the Urinary Bladder: A Case Report and Review of Literature. Siriraj Med J. 2018;70(3):254-9.
9. Azémar MD, Comperat E, Richard F, Cussenot O, Rouprêt M. Bladder recurrence after surgery for upper urinary tract urothelial cell carcinoma: frequency, risk factors, and surveillance. Urol Oncol. 2011;29(2):130-6.
10. Wu WJ, Ke HL, Yang YH, Li CC, Chou YH, Huang CH. Should patients with primary upper urinary tract cancer receive prophylactic intravesical chemotherapy after nephroureterectomy? J Urol. 2010;183(1):56-61.
11. O'Brien T, Ray E, Singh R, Coker B, Beard R, British Association of Urological Surgeons Section of O. Prevention of bladder tumours after nephroureterectomy for primary upper urinary tract urothelial carcinoma: a prospective, multicentre, randomised clinical trial of a single postoperative intravesical dose of mitomycin C (the ODMIT-C Trial). Eur Urol. 2011;60(4):703-10.
12. Boonyapalanant C, Woranisarakul V, Jitpraphai S, Chotikawanich E, Taweemonkongsap T, Kc HB, et al. The Efficacy of Inside-Out Transversus Abdominis Plane Block vs Local Infiltration before Wound Closure in Pain Management after Kidney Transplantation: A Double-blind, Randomized Trial. Siriraj Med J. 2022;74(4): 233-8.
13. Blute ML, Segura JW, Patterson DE, Benson RC, Jr., Zincke H. Impact of endourology on diagnosis and management of upper urinary tract urothelial cancer. J Urol. 1989;141(6):1298-301.
14. Favaretto RL, Shariat SF, Savage C, Godoy G, Chade DC, Kaag M, et al. Combining imaging and ureteroscopy variables in a preoperative multivariable model for prediction of muscle-invasive and non-organ confined disease in patients with upper tract urothelial carcinoma. BJU Int. 2012;109(1):77-82.
15. Rojas CP, Castle SM, Llanos CA, Santos Cortes JA, Bird V, Rodriguez S, et al. Low biopsy volume in ureteroscopy does not affect tumor biopsy grading in upper tract urothelial carcinoma. Urol Oncol. 2013;31(8):1696-700.
16. Chung Y, Lee DH, Lee M, Kim H, Lee S, Hong SK, et al. Impact of diagnostic ureteroscopy before radical nephroureterectomy on intravesical recurrence in patients with upper tract urothelial cancer. Investig Clin Urol. 2020;61(2):158-65.
17. Guo RQ, Hong P, Xiong GY, Zhang L, Fang D, Li XS, et al. Impact of ureteroscopy before radical nephroureterectomy for upper tract urothelial carcinomas on oncological outcomes: a meta-analysis. BJU Int. 2018;121(2):184-93.
18. Luo HL, Kang CH, Chen YT, Chuang YC, Lee WC, Cheng YT, et al. Diagnostic ureteroscopy independently correlates with intravesical recurrence after nephroureterectomy for upper urinary tract urothelial carcinoma. Ann Surg Oncol. 2013;20(9): 3121-6.
19. Marchioni M, Primiceri G, Cindolo L, Hampton LJ, Grob MB, Guruli G, et al. Impact of diagnostic ureteroscopy on intravesical recurrence in patients undergoing radical nephroureterectomy for upper tract urothelial cancer: a systematic review and meta-analysis. BJU Int. 2017;120(3):313-9.
20. Sankin A, Tin AL, Mano R, Chevinsky M, Jakubowski C, Sfakianos JP, et al. Impact of Ureteroscopy Before Nephroureterectomy for Upper Tract Urothelial Carcinoma on Oncologic Outcomes. Urology. 2016;94:148-53.
21. Sung HH, Jeon HG, Han DH, Jeong BC, Seo SI, Lee HM, et al. Diagnostic Ureterorenoscopy Is Associated with Increased Intravesical Recurrence following Radical Nephroureterectomy in Upper Tract Urothelial Carcinoma. PLoS One. 2015;10(11):e0139976.
22. Fang D, Li XS, Xiong GY, Yao L, He ZS, Zhou LQ. Prophylactic intravesical chemotherapy to prevent bladder tumors after nephroureterectomy for primary upper urinary tract urothelial carcinomas: a systematic review and meta-analysis. Urol Int. 2013;91(3):291-6.
23. Ito A, Shintaku I, Satoh M, Ioritani N, Aizawa M, Tochigi T, et al. Prospective randomized phase II trial of a single early intravesical instillation of pirarubicin (THP) in the prevention of bladder recurrence after nephroureterectomy for upper urinary tract urothelial carcinoma: the THP Monotherapy Study Group Trial. J Clin Oncol. 2013;31(11):1422-7.
24. Potretzke AM, Knight BA, Vetter JM, Anderson BG, Hardi AC, Bhayani SB, et al. Diagnostic Utility of Selective Upper Tract Urinary Cytology: A Systematic Review and Meta-analysis of the Literature. Urology. 2016;96:35-43.
25. Hong S, Kwon T, You D, Jeong IG, Hong B, Hong JH, et al. Incidence of benign results after laparoscopic radical nephroureterectomy. JSLS. 2014;18(4):e2014.00335.
26. Gurbuz C, Youssef RF, Shariat SF, Lotan Y, Wood CG, Sagalowsky AI, et al. The impact of previous ureteroscopic tumor ablation on oncologic outcomes after radical nephrouretectomy for upper urinary tract urothelial carcinoma. J Endourol. 2011;25(5): 775-9.
27. Shibing Y, Liangren L, Qiang W, Hong L, Turun S, Junhao L, et al. Impact of tumour size on prognosis of upper urinary tract urothelial carcinoma after radical nephroureterectomy: a multi-institutional analysis of 795 cases. BJU Int. 2016;118(6): 902-10.
28. Lee JK, Kim KB, Park YH, Oh JJ, Lee S, Jeong CW, et al. Correlation Between the Timing of Diagnostic Ureteroscopy and Intravesical Recurrence in Upper Tract Urothelial Cancer. Clin Genitourin Cancer. 2016;14(1):e37-41.