E-ISSN 2228-8082

Volume 74, Number 10, October 2022

SMJ

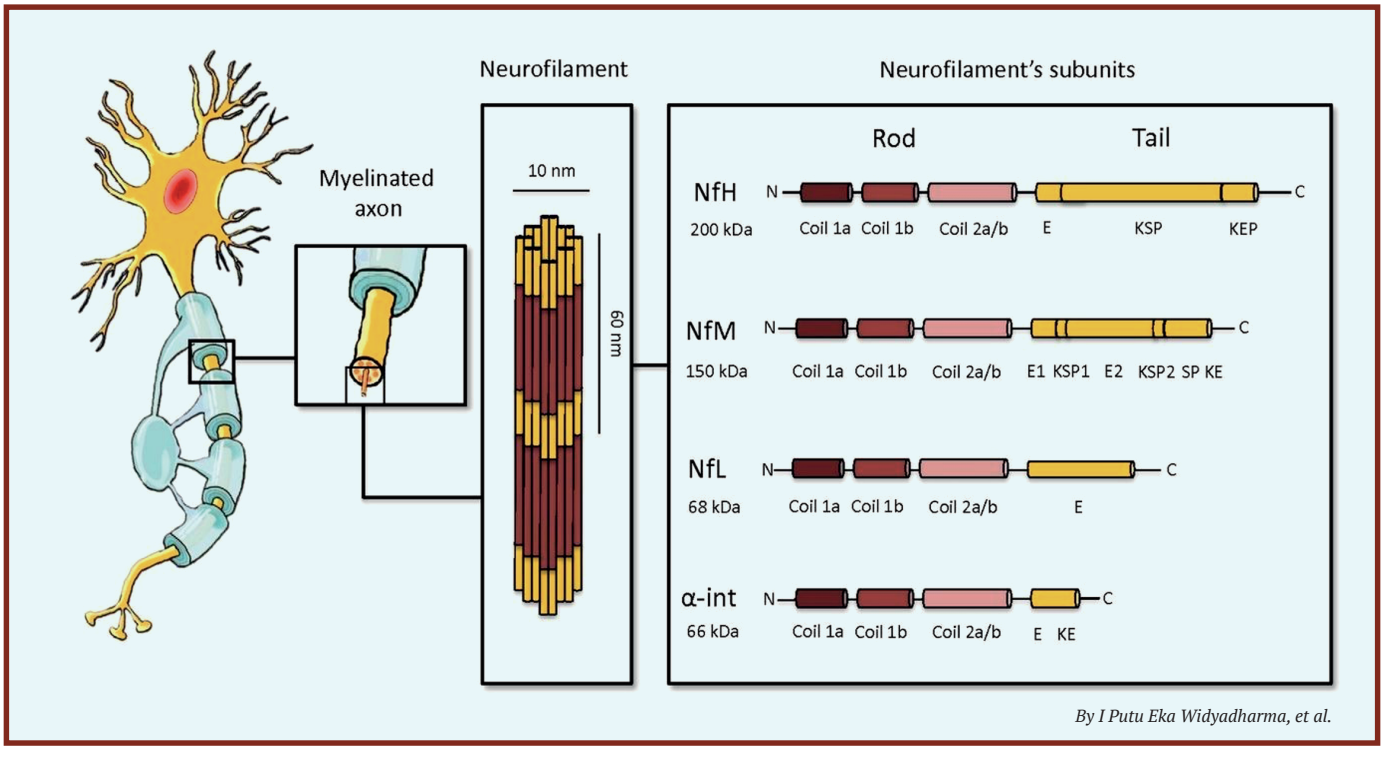
Siriraj Medical Journal

The world-leading biomedical science of Thailand

MONTHLY

ORIGINAL ARTICLE

REVIEW ARTICLE



By I Putu Eka Widyadharma, et al.

Indexed by Scopus®

THAILAND SECTION 1954

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

E-mail: sijournal92@gmail.com

Siriraj Medical Journal SMJ

Volume 74, Number 10, October 2022

ORIGINAL ARTICLE

REVIEW 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].

Paradigm Shift from Open Surgery to Minimally Invasive Surgery in Three Approaches for Radical Prostatectomy: Comparing Outcomes and Learning Curves

Sittiporn Srinualnad, M.D.*, Thitipat Hansomwong, M.D.*, Pubordee Aussavavirojekul, M.D.*, Pat Saksirisampant, M.D.**

**Division of Urology, Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, **Division of Urology, Department of Surgery, Somdech Phra Pinklao Hospital, Naval Medical Department, Royal Thai Navy, Bangkok, Thailand.*

ABSTRACT

Objective: Radical prostatectomy (RP) can be performed by several approaches, such as open retropubic radical prostatectomy (RRP), laparoscopic radical prostatectomy (LRP), and robotic-assisted laparoscopic prostatectomy (RALP). This study investigated and shared the differences in the surgical techniques, learning curves, and outcomes of each approach of RP.

Materials and Methods: The data of patients who received RP given by one of the authors between January 2002 to June 2016 were retrospectively reviewed. We compared perioperative and postoperative outcomes among approaches, searched for predictors of a positive surgical margin (PSM), and assess the learning curves of the two minimally invasive approaches.

Results: 527 patients underwent RP during January 2002 to June 2016 including 42 RRP, 198 LRP, and 327 RALP. RALP had the highest negative surgical margin (68.8%) and lowest multifocal positive surgical margin (10.7%). PSM predictors were the Gleason score and pathological T staging. The learning curve showed that RALP needed one-hundred-cases experience to achieve the lowest PSM rate and 200 cases to master bleeding control. In the first 100 cases in each group, the PSM rate in LRP was lower than in RALP.

Conclusion: Minimally invasive approach in radical prostatectomy showed significant improvements over RRP, especially the RALP approach. RALP would take a surgeon 100 and 200 cases to reach the plateau on the learning curve for achieving the desired oncologic and perioperative outcome efficiencies, respectively. However, LRP and RRP are still feasible in a service setting and for training purposes.

Keywords: Radical prostatectomy; Laparoscopic; Robotic; Localized prostate cancer; Learning curve (Siriraj Med J 2022; 74: 618-626)

INTRODUCTION

Prostate cancer is one of the most commonly diagnosed cancers in male populations worldwide.[1] Back in the day, Thai patients usually developed lower urinary tract symptoms such as frequency, intermittent voiding, and a sense of residual urine as the tumor progressed extensively before their first visit to the hospital.[2] Therefore, initial radical prostatectomy (RP) had significant morbidity and mortality due to the late detection of cancer and poor understanding of the procedure. After early detection

Corresponding author: Sittiporn Srinualnad
E-mail: sitsrinualnad@gmail.com
Received 6 September 2021 Revised 28 January 2022 Accepted 31 January 2022
ORCID ID: http://orcid.org/0000-0002-5118-7675
http://dx.doi.org/10.33192/Smj.2022.73

with PSA screening was adopted, it provided a window of opportunity to eradicate cancer while the patient was still in a curable state and encouraged surgeons to improve both their surgical technique and the operative outcome.[3]

Currently, RP is a standard treatment for localized prostate cancer. The procedure can proceed via several approaches: open retropubic radical prostatectomy (RRP), laparoscopic radical prostatectomy (LRP), and robotic-assisted laparoscopic prostatectomy (RALP).[3-9] Minimally invasive approaches (LRP and RALP) provide several benefits to both the patient and surgeon by helping minimize blood loss, shortening the hospital stay, and improving the oncological outcome in patients.[10,11] From the point of view of the surgeon, minimally invasive approaches allow a better-magnified vision, 3D perception, more comfortable ergonomic posture during operation, and intracorporal wrist rotation in RALP.[12,13]

The present study aimed to investigate and share the differences in the surgical technique, surgeon learning curve for RALP, and oncological outcome of each approach for RP.

MATERIALS AND METHODS

We retrospectively reviewed the RP cases from January 2002 to June 2016 performed by one of the authors (Sittiporn Srinualnad) since he was a novice in the field. In total, 567 cases were eligible for inclusion in the study, comprising 42 RRP cases, 198 LRP cases, and 327 RALP cases. The surgeon started to perform a high number of cases of RRP, LRP, and RALP in 2002, 2006, and 2007, respectively.

All patients were placed in low lithotomy with Trendelenburg position and were under general anesthesia. In RRP setting, the surgeon made a lower midline incision, while in LRP and RALP, the surgeon had laparoscopic ports set in position with one 12 mm camera port, three 8 mm working ports, a 12 mm assistant port, and a 5 mm assistant port. In RALP, the Intuitive da Vinci robotic surgical system (model S, Si, or Xi) was docked into position and its robotic arms were installed to all 8 mm working ports. For visualization, LRP was the only approach performed without depth perception as only 2D monitors were used. On the contrary, RALP had the consoles which provide 3D footage and RRP exposed the surgeon directly to the surgical field. After the initialization had been done, the retropubic radical prostatectomy was similarly commenced. Notably, in the dorsal venous complex (DVC) controlling step, the surgeon used barbed suture ligation to secure the vascular structures.

Preoperative, perioperative, and postoperative data consisting of the patient's demographics, stage of the disease, operative choice decision, operation details, pathological report of the excised gland, and follow-up information were reviewed and analyzed. The positive surgical margin (PSM), operative time, complication rate, perioperative blood loss, and blood transfusion rate were measured and compared among the groups.

The primary objective was defined to compare perioperative and postoperative outcomes of three different surgical approaches of radical prostatectomy. The secondary objectives were the identification of factors associated with PSM and the learning curves of two minimally invasive approaches.

Shapiro-Wilk test, D'Agostino K2 test, and Anderson-Darling test to evaluate the distribution of continuous data. Normally distributed data with mean and standard deviation were assessed using the Student's t-test. Non-normally distributed continuous data with the median and the interquartile range (IQR) were assessed using the Kruskal–Wallis test among groups. Categorical data are reported as the frequency and percentage and were assessed using the chi-square test. Each type of PSM, such as multifocal positive surgical margin (MPSM) and single focal positive surgical margin (SPSM), and the site of the PSM (i.e., apex, anterior, posterior, bladder neck) are reported as the frequency in each group and were assessed using the chi-square test. We also constructed univariate and multivariate analyses to find the predictors responsible for PSM.

Statistical analysis was performed using the Python statistical packages: SciPy version 1.5.2, lifelines version 0.25.10, and Matplotlib version 3.3.2. Univariate and multivariate analyses were performed using linear regression and multiple linear regression via the package statsmodels version 0.12.0. A p-value less than 0.05 was considered significant.

RESULTS

Demographics of the patients' group by surgical approach

Between January 2002 to June 2016, the author performed RP on 527 patients, of whom 42 (7.9%) were assigned to the RRP group, 198 (37.6%) to the LRP group, and 327 (62.0%) to the RALP group. The median ages of the patients in each group were 69.5 years old (IQR 63.3, 73.8), 68 years old (IQR 63.3, 73.0), and 66.4 years old (IQR 60.0, 73.0) for the RRP, LRP, and RALP groups. The median preoperative serum prostate-specific antigen levels (PSA) were 15 ng/ml (IQR 5.9, 33.6), 8.5 ng/ml (IQR 5.6, 18.8), and 8.9 ng/ml (IQR 5.9, 16.0), respectively, for the RRP, LRP, and RALP groups. The

median prostate weights in each group were 46.0 g (32.0, 55.0), 38.8 g (31.0, 53.5), and 39.2 g (30.4, 49.6), respectively, for the RRP, LRP, and RALP groups. We found no significant differences in age, preoperative PSA, or prostate weight among the three groups of surgical approaches (Table 1).

Perioperative blood loss, as one of the evaluation indexes, was significantly different among the three groups with a p-value less than 0.01. The median blood loss in the RRP group was 1200 ml (900, 2000), which was the worst loss compared to its minimally invasive surgery (MIS) counterpart. Between the MIS approaches, LRP had a median blood loss of 800 ml (500, 1300) and RALP had a loss of 300 ml (200, 550) (Table 1).

The operative times were remarkably different with a p-value less than 0.01. The most time-consuming operation was LRP with a median duration of 265 minutes (215, 310). RALP had a median operative time of 180 minutes (155,220), while RRP still involved the lowest operative time of 165 minutes (145, 200) (Table 1).

Nerve-sparing attempts were significantly performed more frequently in RALP (26.6%) than RRP and LRP (7.1% and 9.6%, respectively) with a p-value < 0.01. We found that RRP had a significantly higher postoperative persistent PSA rate of 14.3% compared to the other approaches at 6.6% and 4.3% for LRP and RALP, respectively. Patients with seminal vesicle invasion proportionally underwent the RRP approach more frequently than any other approach with a p-value of 0.02, but this proportion between the two minimally invasive approaches was not significantly different (Table 2).

Adjuvant therapy was given proportionally the most frequently in the RRP group (42.9%) compared to LRP (17.2%) and RALP (28.8%), to cover any residual tumor at the PSM. Androgen deprivation therapy (ADT), which comprises bilateral orchidectomy or intramuscular injection of GnRH agonist, was the most popular selection, with 42.9%, 16.7%, and 27.8% of patients from the RRP, LRP, and RALP groups receiving ADT as either a single treatment or as a part of combination treatment as their adjuvant therapy. Novel androgen receptor inhibitors (AR) were rarely given after RRP (4.8%) and LRP (5.1%) but were given in 11.9% of patients after RALP. Radiation therapy was only given in combination treatment with ADT and/or AR, which happened in 0.5% and 7.7% of patients after LRP and RALP, respectively (Table 2).

TABLE 1. Characteristics of chronic low back pain patients.

	RRP (n = 42)		LRP (n = 198)		RALP (n = 327)		P-value
	Median	*IQR*	*Median*	*IQR*	*Median*	*IQR*	
Age (years old)	69.5	63.3 - 73.8	68	63.3 - 73.0	67	60.0 - 73.0	0.16
PSA (ng/mL)	15	5.9 - 33.6	8.5	5.6 - 18.8	8.9	5.9 - 16.0	0.21
post-op PSA (ng/mL)	0.01	0.00 - 0.06	0	0.00 - 0.03	0	0.00 - 0.01	0.08
Op-time (min)	165	145.0 - 200.0	265	215.0 - 310.0	180	155.0 - 220.0	< 0.01
Hospital stays after surgery (days)	8	7.0 - 10.0	8	7.0 - 9.0	7	7.0 - 8.0	< 0.01
Highest PSA after surgery (ng/mL)	0.1	0.00 - 0.90	0.1	0.01 - 0.56	0.03	0.00 - 0.40	0.39
Prostate (g)	46	32.0 - 55.0	38.8	31.0 - 53.5	39.2	30.4 - 49.6	0.33
F/u time (years)	6.8	1.3 - 12.2	6.9	2.6 - 9.3	5.49	2.3 - 8.2	0.01
Blood loss (ml)	1200	900.0 - 2000.0	800	500.0 - 1300.0	300	200.0 - 550.0	< 0.01
node gain	4.5	3.0 - 7.0	7	4.0 - 10.0	6	4.0 - 8.0	0.06
positive node	0	0.0 - 0.0	0	0.0 - 0.0	0	0.0 - 0.0	0.52

TABLE 2. Categorical data of patients categorized by the three surgical approaches.

	RRP (n = 42)		LRP (n = 198)		RALP (n = 327)		P-value
	n	%	n	%	n	%	
NVB sparing							
None	39	92.9%	179	90.4%	240	73.4%	<0.01
Unilateral	2	4.8%	4	2.0%	31	9.5%	
Bilateral	1	2.4%	15	7.6%	56	17.1%	
Pathological T stage							
0	0	0.0%	10	5.1%	4	1.2%	0.12
2	24	57.1%	103	52.0%	200	61.2%	
3a	5	11.9%	55	27.8%	76	23.2%	
3b	13	31.0%	30	15.2%	47	14.4%	
Pathological N stage							
pN0	39	92.9%	190	96.0%	315	96.3%	0.56
pN1	3	7.1%	8	4.0%	12	3.7%	
Seminal vesical invasion	13	31.0%	30	15.2%	46	14.1%	0.02
Extra prostatic extension	17	40.5%	81	40.9%	112	34.3%	0.28
Perineural invasion	25	59.5%	134	67.7%	243	74.3%	0.06
Persistent PSA	6	14.3%	13	6.6%	14	4.3%	0.03
Adjuvant therapy							
None	24	57.1%	164	82.8%	233	71.3%	<0.01
ADT	16	38.1%	23	11.6%	30	9.2%	
AR	0	0.0%	1	0.5%	2	0.6%	
ADT AR	2	4.8%	9	4.6%	37	11.3%	
ADT RT	0	0.0%	0	0.0%	7	2.1%	
AR RT	0	0.0%	0	0.0%	1	0.3%	
ADT AR RT	0	0.0%	1	0.5%	17	5.2%	
Immediate complication							
Surgical site infection	2	4.76%	0	0.0%	0	0.0%	0.22
Bowel injury	0	0.00%	1	0.5%	1	0.3%	
Ureter injury	0	0.00%	1	0.5%	0	0.0%	
Hematoma	0	0.00%	1	0.5%	1	0.3%	
Other	0	0.00%	0	0.0%	2	0.6%	
Delay complication							
inguinal hernia	5	11.9%	8	4.0%	27	8.3%	0.05
urethral stricture	3	7.1%	5	2.5%	4	1.2%	
incontinence	1	2.4%	5	2.5%	15	4.6%	
Gleason score from surgery							
6	9	21.4%	47	24.0%	86	27.5%	0.94
3+4	8	19.1%	68	34.7%	116	37.1%	
4+3	8	19.1%	34	17.4%	55	17.6%	
8	6	14.3%	22	11.2%	23	7.4%	
9	11	26.2%	24	12.2%	33	10.5%	
10	0	0.0%	1	0.5%	0	0.0%	

Abbreviations: ADT; Androgen deprivation therapy: including bilateral orchidectomy and GnRH-agonist injection, AR; Novel anti-androgen receptor therapy, RT; Radiation therapy.

Factors responsible for a positive surgical margin rate

According to Table 3, only the negative surgical margin rate (NSM) and MPSM rate were significantly different among the surgical approaches, with a p-value less than 0.05. RALP had the highest NSM rate (68.8%) and the lowest MSPM rate (10.7%), while RRP remained on the worst side with a 50.0% NSM and 28.6% MPSM.

Having categorized SPSM into specific sites where the PSM resides, we found that the statistical analysis did not find a significant difference for the PSM sites among the three approaches, and the most common site of SPSM was at the apex of the prostate for all approaches (11.9% in RRP, 20.2% in LRP, 20.2% in RALP).

While the frequency of patients with each pathological T staging (pT) was not significantly different among the three approaches (Table 2), the PSM rates were. We detected significantly higher odds ratios of PSM in the pT3a and pT3b stages compared to the pT2 stage in both the univariate and multivariate models, with a p-value less than 0.001 (Table 4).

The nerve-sparing technique was attempted significantly more frequently in RALP than in the other surgical approaches, with a p-value less than 0.01 (Table 2). The data on the univariate analysis showed that the bilateral nerve-sparing technique might lower the PSM rate, with an odd ratio of 0.40, 95% confidence interval -1.53 to -0.30, and p-value of less than 0.01, but the multivariate analysis proved otherwise (Table 4).

Learning curve

The split method was used to analyze the learning curves of the surgeon for all approaches of RP, as presented in Table 5. The 567 patients in total were split into two groups chronologically. Perioperative blood loss, transfusion rate, and NSM all improved over time. Delay complications were the only group that did not seem to be related to the surgeon experience. According to Table 2, the most common delay complication encountered was an inguinal hernia, which is common in 50–69-year-old men.[14]

Furthermore, subgroup learning curve analysis was performed, especially for the minimally invasive surgery platform, as shown in Fig 1. The factors focused on were perioperative blood loss, number of cases that received blood transfusion, surgical margin status, hospital stay after surgery, and number of dissected lymph nodes.

In Fig 1, it can be seen that both the perioperative blood loss and transfusion rates showed a relative decrease with the increasing experience of the surgeon in the RALP approach, whereas the LRP approach group still had relatively high perioperative blood loss and needed transfusions more often than for RALP. By the time the surgeon's RALP experience had reached the 90th case, the transfusion rate was almost reduced in half. LRP still needed experience from over 100 cases to show a reduced transfusion rate.

For the positive surgical margin, the RALP learning curve suggested it would take around 90-120 cases to achieve a plateau, which was faster than for LRP. However, LRP was able to achieve a lower PSM rate than RALP in the first 100 cases (Fig 1).

In terms of the operative time, the learning curve had already taken place at 30–60 cases of RALP and continued to reduce to a plateau at the 200th case. LRP also could reduce the operative time within the first 100 cases, but it could not match RALP's slope, as shown in Fig 1.

TABLE 3. Frequency of PSM among the three surgical approaches.

	RRP (n = 42)		LRP (n = 198)		RALP (n = 327)		P-value
	n	%	n	%	n	%	
Negative SM	21	50.0%	127	64.1%	225	68.8%	0.04
SPSM	10	23.8%	67	33.8%	129	39.5%	0.09
Apical	5	11.9%	40	20.2%	66	20.2%	0.43
Posterior	3	7.1%	10	5.1%	31	9.5%	0.18
Posterolateral	0	0.0%	1	0.5%	2	0.6%	0.87
Anterior	0	0.0%	1	0.5%	7	2.1%	0.22
Bladder neck	1	2.4%	15	7.6%	23	7.0%	0.47
Vas Deferens + Seminal Vesicle	1	2.4%	0	0.0%	0	0.0%	
MPSM	12	28.6%	29	14.7%	35	10.7%	< 0.01

TABLE 4. Univariate and multivariate analyses of the risk factors for PSM.

	Univariate analysis Crude OR (95% CI)	P-value	Multivariate analysis Adjusted OR (95% CI)	P-value
Surgical approach				
RRP	1(ref)		1(ref)	
LRP	0.6(-1.3, 0.1)	0.089	0.7(-1.1, 0.5)	0.399
RALP	0.5(-1.4, -0.1)	0.017	0.7(-1.2, 0.4)	0.343
Gleason score				
6	1(ref)		1(ref)	
3+4	3.8(0.7, 1.9)	<0.001	2.3(0.2, 1.5)	0.012
4+3	3.9(0.7, 2.0)	<0.001	2.2(0.04, 1.5)	0.037
8	11.3(1.7, 3.2)	<0.001	5.9(0.9, 2.6)	<0.001
9	20.3(2.3, 3.8)	<0.001	6.9(1.1, 2.8)	<0.001
NVB sparing				
None	1(ref)		1(ref)	
Unilateral	0.5(-1.6, 0.03)	0.058	0.5(-1.6, 0.4)	0.211
Bilateral	0.4(-1.5, -0.3)	0.004	0.6(-1.2, 0.2)	0.148
Pathological T				
pT2	1(ref)		1(ref)	
pT3a	4.4(1.1, 1.9)	<0.001	3.1(0.6, 1.6)	<0.001
pT3b	8.8(1.7, 2.7)	<0.001	3.6(0.6, 1.9)	<0.001
PSA (ng/ml)				
<10	1(ref)		1(ref)	
10 to 20	1.3(-0.2, 0.7)	0.280	0.9(-0.6, 0.5)	0.791
>=20	3.8(0.9, 1.8)	<0.001	1.3(-0.3, 0.8)	0.356
Prostate volume (g)				
<25	1(ref)			
25 to 50	0.9(-0.8, 0.5)	0.638		
50 to 75	0.6(-1.3, 0.2)	0.150		
>=75	0.8(-1.1, 0.6)	0.602		

TABLE 5. Split method analysis of the learning curves

Outcome	First half (n = 284)		Later half (n = 283)		P-value
NSM (n, %)	157	55.5%	216	76.1%	<0.05
SPSM (n, %)	109	38.5%	97	34.2%	0.31
MSPM (n, %)	43	15.2%	33	11.6%	0.22
Immediate complication (n, %)	5	1.8%	4	1.4%	0.74
Delay complication (n, %)	25	8.8%	49	17.3%	<0.05
Blood transfusion (n, %)	104	36.8%	25	8.8%	<0.05
Blood loss (median, IQR)	700	400-1200	350	200-600	<0.05

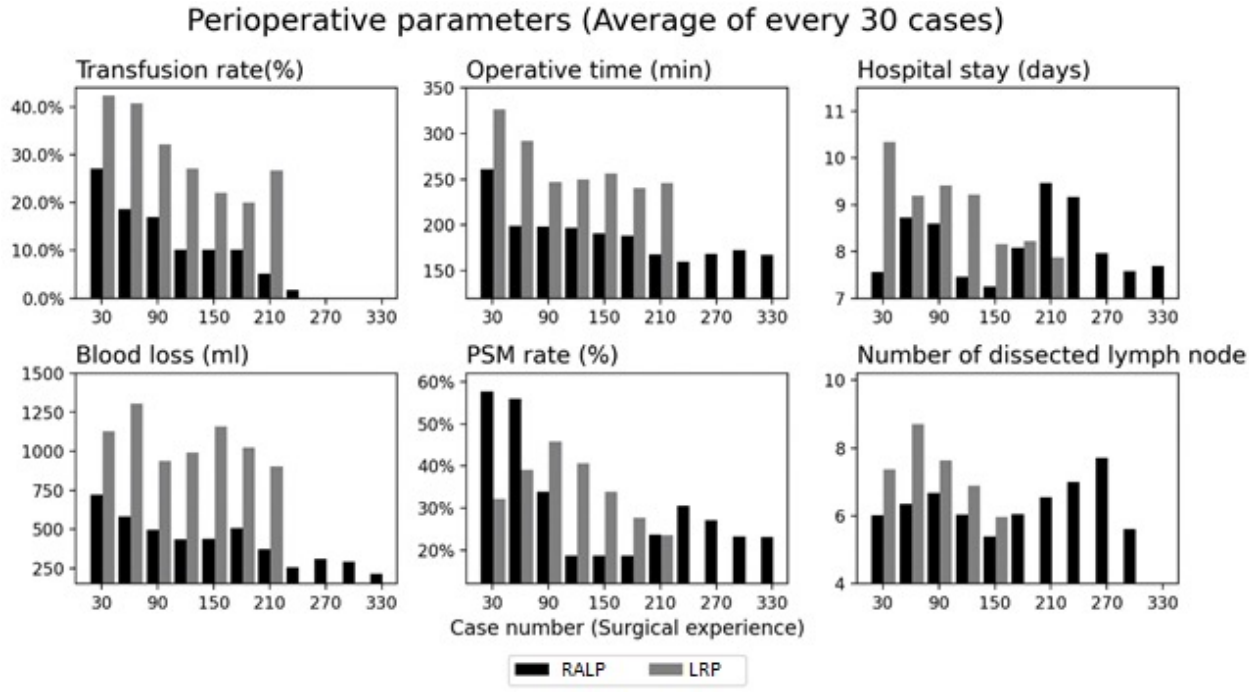


Fig 1. Learning curve analyses using perioperative parameters and a subgrouping of every 30 cases

DISCUSSION

Originally when RP was first developed, patients were usually diagnosed as having prostate cancer with a higher clinical and pathological stage due to the lack of screening serum PSA level. The outcome of the procedure was described as very poor and played very little part in the management. After the serum PSA level was widely used as a screening tool, patients could have prostate cancer diagnosed 7–9 years earlier than previously was possible, which meant it was more likely to be found in a curable stage of the disease. Later, Patrick Walsh made dramatic contributions to the RP field by sharing a better understanding of the anatomy of the prostate, which he described by a cavernous nerve-sparing technique, and he proposed DVC ligation for early bleeding control.[3,15] The procedure then achieved improvements in both safety and functional outcome. After that, the laparoscopic approach came into play for pelvic lymph node dissection early on. Then, LRP was first described in 2000.[3,16]. The procedure could clearly improve patient recovery but still could not achieve a comparable functional outcome, i.e., continence. In the 21st century, RALP was introduced to the field. The procedure drove another revolution in outcome expectations by reducing the transfusion rate to near zero while requiring only 24 hours hospital stay, and achieving better functional outcomes in terms of both continence and sexual function.[3]

Although the average blood loss of RALP is similar to the previous publication with a mean of 300 and IQR of 200-500 mL, LRP is more complicated.[10] For LRP, the step that is majorly responsible for bleeding is controlling DVC, which could be done in two manners: suture ligation or endoscopic stapler. In this study, the surgeon favors the suture ligation technique with barbed suture material over another technique. Even though the two techniques shouldn't make difference in estimate blood loss[17], two potential reasons might entitle for more blood loss in LRP group. Firstly, as the surgeon aimed for preserving urethral length to secure long-term functional outcomes of the patients, he made fewer ligations on DVC. With the cruder movement of LRP, sometimes the suture was slipped and led to more bleeding. The second explanation associate with tumor location as apical tumor did increase the difficulty of the dissection step and needed wider excision around DVC, which directly related to more bleeding.

Regarding Saksirisampant et al, although PSM status could be influenced by several preoperative factors, which were serum PSA level, small prostate, percentage of tumor volume, pathological T stage, and ISUP Gleason grade group, it is one of the outcomes that directly demonstrate one's surgical skills.[18] The surgeon's always responsible for the decision making whether to perform wider excision on which border to compensate for the aggressiveness of the tumor. Having prostate MRI preoperatively might help specify tumor's locations and boundaries, which aware the surgeon of those areas and eventually reduce PSM.

While the oncological outcome result showed little to no difference among the three approaches of RP[10,11], outcomes in the perioperative and immediate postoperative periods, such as blood loss, transfusion rate, and postoperative hospital stay, were improved significantly in RALP. Recent research in the field suggests that a single night's stay or same-day discharge might be possible for selected patients regarding safety.[19,20] Compared to our study, RALP did show a lower transfusion rate and minimized the immediate complications, while our data did not show an extremely shortened hospital stay due to various reasons, such as institute policy and the patients' request. Implementing enhanced recovery after surgery (ERAS) protocols might be one solution that could contribute toward this outcome.[21-23]

Between the two minimally invasive approaches, there are many differences in the way the operations are performed. LRP uses surgical instruments that need holding and control directly by the surgeon's hand in the surgical field themselves, whereas RALP has a surgical robot to hold and move the instruments, which the surgeon operates from a console in a sitting posture. The time required to initialize for each operation is obviously longer in RALP, including the cleaning and docking time, but the RALP console time could be reduced lower than in LRP with a more experienced surgeon. While RALP may lack haptic feedback from the operating field, the platform allows the surgeon to operate with steady tool movement, at ease and in a relaxed manner. Therefore, RALP helps reduce fatigue and minimize the surgeon's stress. With a good magnifying view and precise movement, robotic-assisted surgery eliminates the need for excellent human eye vision and perfectly steady human hand movement. Therefore, it could extend the surgeon's retirement age in terms of overcoming the physical limitations due to their increasing age.

According to our study, RALP would take around 100 cases to master full PSM control and another 100 cases to achieve minimal blood loss and operative time, which was similar to in Song et al. and Kongchareonsombat's studies that marked 230 operations experience as needed for the best bleeding control performance in RALP.[24,25] Compared to LRP in the same evaluating matrix, LRP would take a much longer time to improve the outcome and would achieve even less. According to Secin et al., it would take 200 to 250 cases to find the plateau for the PSM rate in the learning curve of LRP.[26] Due to the limitations of our dataset, which consisted of only 198 cases of LRP, our findings on this issue remain inconclusive in our study.

Interestingly, the oncological outcome (i.e., PSM) was poorer in RALP at the beginning of the surgeon's expertise. However, after 100 cases of experience were gained, the surgeon's performance in RALP eventually surpassed LRP with a very surprising steep slope for the PSM rate. The surgeon's familiarity with the platform may play an important role in this finding, because the loss of tactile sensation, hand-eye coordination process, and tools handling may not play along with the surgeon's expectations in the early phase of their learning experience. However, after the surgeon has become more familiar with the robotic platform and can handle it well, then the PSM rate is improved.

With the advancement of technology and surgical options, there might be some questions regarding whether RRP should be performed in the modern-day. Despite the improvement in outcomes that RALP may give, we believe that there is still some role for RRP to play. For example, barriers to financial affordability and accessibility to a minimally invasive surgery platform should allow RRP to continue thriving in this century. In some countries where surgical fees may also increase by the time under anesthesia, a shorter operative time might have a considerable beneficial impact for patients with financial problems. Therefore, many surgeons still see benefits in the RRP approach and will continue to refine their skills in this area. Recently, they can even match some of RRP's beneficial outcomes to the MIS platform, such as sexual function and continence function.[27-29]

There are limitations in learning curve interpretation including a possible confounder as the author started performing these three approaches of RP at different times in his career. We believe some surgical experience could be used as cross-platform knowledge, whereby the outcome of learning a later proposed surgical approach, such as RALP, could be confounded by prior experience of RRP and LRP.[30] Therefore, the number of cases required for a surgeon who is a complete novice in minimally invasive surgery to learn RALP could not be determined by this study and needs further research.

CONCLUSION

Nowadays, urologists in academic centers have a high interest in the MIS approach of RP and prefer RALP most of the time. RALP has clear advantages in disease control and hospital resource utilization. Here, learning curve analysis determined that 100 cases of RALP were the minimal requirement for a surgeon to master the procedure. However, not all patients can afford such an operation and cost is a major concern. Therefore, LRP and RRP still play important roles in both service hospital settings and academic training centers.

ACKNOWLEDGEMENTS

The authors would like to thank Ms. Jitsiri Chaiyatho, Ms. Julaporn Pooliam, and all the coordinators at Siriraj Hospital for their important contributions to this study.

REFERENCES

1. 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. doi: 10.1002/ijc.33588.
2. Sivilaikul S, Soontrapa S. Epidemiology of Genitourinary Tract Carcinoma in Siriraj Hospital During 1990-1995. Siriraj Med J. 2020;53(3):130-7.
3. Costello AJ. Considering the role of radical prostatectomy in 21st century prostate cancer care. Nat Rev Urol. 2020;17(3):177-88.
4. Arenas-Gallo C, Shoag JE, Hu JC. Optimizing Surgical Techniques in Robot-Assisted Radical Prostatectomy. Urol Clin North Am. 2021;48(1):1-9.
5. Chopra S, Srivastava A, Tewari A. Robotic radical prostatectomy: The new gold standard. Arab J Urol. 2012;10(1):23-31.
6. Faria EF, Maciel CVM, Berger A, Mitre A, Dauster B, Freitas CH, Jr., et al. Recommendations on robotic-assisted radical prostatectomy: a Brazilian experts' consensus. J Robot Surg. 2021.
7. Huynh LM, Ahlering TE. Robot-Assisted Radical Prostatectomy: A Step-by-Step Guide. J Endourol. 2018;32(S1):S28-S32.
8. Martini A, Tewari AK. Anatomic robotic prostatectomy: current best practice. Ther Adv Urol. 2019;11:1756287218813789.
9. Minafra P, Carbonara U, Vitarelli A, Lucarelli G, Battaglia M, Ditonno P. Robotic radical perineal prostatectomy: tradition and evolution in the robotic era. Curr Opin Urol. 2021;31(1):11-7.
10. Cao L, Yang Z, Qi L, Chen M. Robot-assisted and laparoscopic vs open radical prostatectomy in clinically localized prostate cancer: perioperative, functional, and oncological outcomes: A Systematic review and meta-analysis. Medicine (Baltimore). 2019;98(22):e15770.
11. Ilic D, Evans SM, Allan CA, Jung JH, Murphy D, Frydenberg M. Laparoscopic and robotic-assisted versus open radical prostatectomy for the treatment of localised prostate cancer. Cochrane Database Syst Rev. 2017;9:CD009625.
12. Hung S-C, Ou Y-C, Cheng C-L, Hung S-W, Ho H-C, Chiu K-Y, et al. Standardized procedure of robotic assisted laparoscopic radical prostatectomy from case 1 to case 1200. Urological Science. 2016;27(4):199-207.
13. Haga N, Takinami R, Tanji R, Onagi A, Matsuoka K, Koguchi T, et al. Comprehensive approach for post-prostatectomy incontinence in the era of robot-assisted radical prostatectomy. Fukushima J Med Sci. 2017;63(2):46-56.
14. Kark AE, Kurzer M. Groin hernias in women. Hernia. 2008;12(3):267-70.
15. Walsh PC, Donker PJ. Impotence Following Radical Prostatectomy: Insight Into Etiology and Prevention. J Urol. 1982;128(3):492-7.
16. Ballantyne GH, Moll F. The da Vinci telerobotic surgical system: the virtual operative field and telepresence surgery. Surg Clin North Am. 2003;83(6):1293-304.
17. Feng T, Heulitt G, Lee JJ, Liao M, Li HF, Porter JR. Randomised comparison of techniques for control of the dorsal venous complex during robot-assisted laparoscopic radical prostatectomy. BJU Int. 2020;126(5):586-94.
18. Saksirisampant P, Nualyong C, Srinualnad S, Leewansangtong S, Taweemonkongsap T, Jitpraphai S, et al. Positive surgical margins after radical prostatectomy: Associated risk factors in Thai prostate cancer patients. J Med Assoc Thai. 2020;103(5):68-74.
19. Kotamarti S, Williams T, Silver M, Silver DA, Schulman AA. Rethinking the need for overnight admission after robotic-assisted laparoscopic prostatectomy. J Robot Surg. 2020;14(6):913-5.
20. Nason GJ, Kim JK, HeeTan G, Ajib K, Nam RK. Single-night stay for open radical prostatectomy. Can Urol Assoc J. 2021;15(3):E130-E4.
21. Lin C, Wan F, Lu Y, Li G, Yu L, Wang M. Enhanced recovery after surgery protocol for prostate cancer patients undergoing laparoscopic radical prostatectomy. J Int Med Res. 2019;47(1):114-21.
22. Xu Y, Liu A, Chen L, Huang H, Gao Y, Zhang C, et al. Enhanced recovery after surgery (ERAS) pathway optimizes outcomes and costs for minimally invasive radical prostatectomy. J Int Med Res. 2020;48(6):300060520920072.
23. Zhao Y, Zhang S, Liu B, Li J, Hong H. Clinical efficacy of enhanced recovery after surgery (ERAS) program in patients undergoing radical prostatectomy: a systematic review and meta-analysis. World J Surg Oncol. 2020;18(1):131.
24. Song W, Lee SW, Chung JH, Kang M, Sung HH, Jeon HG, et al. Relationship between robotic-assisted radical prostatectomy and retropubic radical prostatectomy in the learning curve of a single surgeon as a novice in radical prostatectomy: A retrospective cohort study. Int J Surg. 2020;81:74-9.
25. Kongchareonsombat W. Perioperative Outcomes and the Learning Curve for Robotic-Assisted Laparoscopic Radical Prostatectomy in Thailand by a Single Surgeon: Six Years' Experience in Ramathibodi Hospital. J Med Assoc Thai. 2019;102(9):951-6.
26. Secin FP, Savage C, Abbou C, de La Taille A, Salomon L, Rassweiler J, et al. The Learning Curve for Laparoscopic Radical Prostatectomy: An International Multicenter Study. J Urol. 2010;184(6):2291-6.
27. Coughlin GD, Yaxley JW, Chambers SK, Occhipinti S, Samaratunga H, Zajdlewicz L, et al. Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: 24-month outcomes from a randomised controlled study. Lancet Oncol. 2018;19(8):1051-60.
28. Pereira R, Joshi A, Roberts M, Yaxley J, Vela I. Open retropubic radical prostatectomy. Transl Androl Urol. 2020;9(6):3025-35.
29. Yaxley JW, Coughlin GD, Chambers SK, Occhipinti S, Samaratunga H, Zajdlewicz L, et al. Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: early outcomes from a randomised controlled phase 3 study. Lancet. 2016;388(10049):1057-66.
30. Tobias-Machado M, Mitre AI, Rubinstein M, Costa EF, Hidaka AK. Robotic-assisted radical prostatectomy learning curve for experienced laparoscopic surgeons: does it really exist? Int Braz J Urol. 2016;42(1):83-9.

Risk Factor of Proximal Lag Screw Cut-Out After Cephalomedullary Nail Fixation in Trochanteric Femoral Fractures: A Retrospective Analytic Study

Chantas Mahaisavariya, M.D.*, Nomina Pradhan, M.D., Kongkhet Riansuwan, M.D.**, Theerawoot Tharmviboonsri, M.D.**, Likit Rugpolmuang, M.D.**, Banchong Mahaisavariya, M.D.**, Direk Tantigate, M.D.****

**Golden Jubilee Medical Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, **Department of Orthopedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.*

ABSTRACT

Objective: A cephalomedullary nail is the treatment of choice for trochanteric fractures; however, a lag screw cut-out is one of the most devastating complications. The lag screw cut-out rate was reported to be around 2.5%–8.3%. This study aimed to evaluate the prevalence of lag screw cut-outs and identify the associated risk factors.

Materials and Methods: A retrospective review of 267 trochanteric fracture patients treated with cephalomedullary nail fixation from January 2007 to December 2017 was conducted. The demographic variables were documented, comprising age, gender, fracture pattern, and AO/OTA classification. Immediate postoperative radiographs were assessed for quality of reduction and implant position. Lag screw cut-outs or radiographic union were determined using the final follow-up radiograph. Prognostic factors associated with lag screw cut-out were determined using univariate and multivariate logistic regression analyses.

Results: Of the 175 patients, 154 were successfully treated, and 21 had a lag screw cut-out. There were no significant differences in mean ages or genders of the union and cut-out groups. No lag screw cut-outs were observed in patients with AO/OTA 31-A1. Patients with AO/OTA 31-B2.1 had a higher rate of screw cut-out (OR 10.5, [3.22, 34.25] $p < .001$). The disintegration of basicervical fragments was significantly associated with lag screw cut-out (OR 5.51, [2.01, 15.12] $p = .001$). The highest cut-out rate was found in the superoanterior and superoposterior positions of the lag screw. However, the screw position did not reach the significance level in a multivariate analysis ($p = .094$).

Conclusion: The prevalence of lag screw cut-out after cephalomedullary nail fixation for trochanteric fractures was 12%. A simple, two-part, basicervical trochanteric fracture hads a significantly higher risk of lag screw cut-out.

Keywords: Trochanteric fracture; pertrochanteric fracture; hip fracture; cut-out; cephalomedullary nail (Siriraj Med J 2022; 74: 627-633)

INTRODUCTION

Hip fracture is one of the common osteoporotic fractures in the geriatric population. Surgical treatment can reduce mortality and improve patients' quality of life.[1-5] A cephalomedullary nail is the treatment of choice for trochanteric fractures, offering superior biomechanical stability to a sliding hip screw.[2,6] However, a lag screw cut-out is one of the most devastating complications, causing difficulties for reoperations or hip replacements.[7,8]

The risk factors associated with lag screw cut-outs have been reported in recent studies. As well as patient factors-age, bone mineral density (BMD) and fracture

Corresponding author: Direk Tantigate
E-mail: dtantigate@gmail.com
Received 22 June 2021 Revised 29 June 2022 Accepted 4 July 2022
ORCID ID: http://orcid.org/0000-0002-3502-8252
http://dx.doi.org/10.33192/Smj.2022.74

configuration-the fixation technique strongly influences the outcomes.[7,9-11] The tip-apex distance (TAD) is the most commonly used parameter to determine the prognosis of a lag screw cut-out.[12]

Even though implant development has tended to match the configuration of a proximal femur, the reported lag screw cut-out rate is still around 2.5% - 8.3%, according to the previous study.[2] This study aimed to evaluate the prevalence of lag screw cut-outs and identify the associated risk factors.

MATERIALS AND METHODS

A retrospective review was conducted of all patients admitted with a trochanteric fracture between January 2007 and December 2017. Only those patients treated with a short cephalomedullary nail (Natural Nail system cephalomedullary nails, Zimmer) with no additional fixation were enrolled. The exclusion criteria were: (1) a patient was treated with other implants; (2) a patient had a pathologic fracture; (3) the pre-or postoperative radiograph was not present in a patient's medical records; (4) a patient was lost to follow-up before the fracture union or screw cut-out; and (5) a patient had a peri-implant fracture.

The demographic data recorded were age, gender, and fracture classification. Based on the AO/OTA classification system[13], the patients were grouped into 31-A1, 31-A2, 31-A3, and 31-B2.1. The preoperative radiographs were reviewed by two investigators (MC and PN), who recorded the following data: (1) lateral buttress fragment-anatomically defined as the lateral femoral cortex distal to the vastus ridge in the trochanteric region[10]; (2) posteromedial fragment-represented by the calcar femorale. The integrity of the lesser trochanter was used as evidence of the presence or absence of the calcar femorale[14]; and (3) basicervical fragment-the fracture line separating the proximal fragment at the intertrochanteric line, corresponding with AO/OTA classification 31-B2.1 (Fig 1). Data from both investigators were compared for inter-observer reliability. Disagreements regarding the result of the measurements were resolved by a third-person review.

The following parameters were measured in the immediate postoperative radiographs: (1) TAD in the anteroposterior (AP) and lateral view; (2) the medial gap after reduction; (3) the postoperative neck-shaft angle; (4) the position of the lag screw in the AP view-defined as inferior, centered and superior; (5) the position of the lag screw in the lateral view-defined as posterior, centered and anterior; (6) the number of distal locking screws; and (7) the mode of distal locking. The TAD

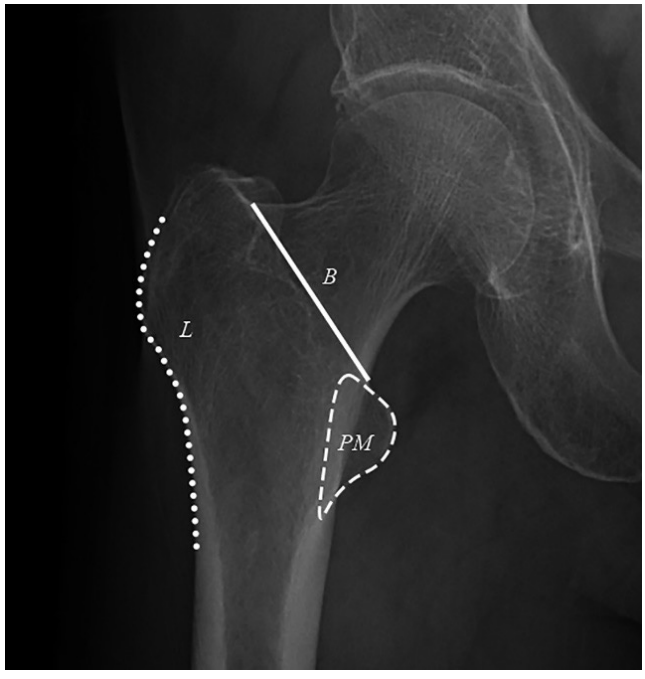


Fig 1. Integrity of specific fragment: Lateral buttress fragment (L), Posteromedial fragment (PM), Basicervical fragment (B)

measurement was obtained using the Baumgaertner method.[1] In this, an imaginary line was drawn parallel to the femur neck, dividing it into superior and inferior portions. The position of the lag screw was recorded using this line as a reference. If the line bisected the screw, the position was recorded as centered.[1] The lateral radiograph used the same method but labeled the positions as posterior, centered, and anterior.

The radiograph with a minimum of 3 months follow-up was measured again when the patient met the endpoint of fixation-union or cut-out. The fracture was labeled as "union" if the bone bridge was present in both the AP and lateral views. If the last follow-up film did not have a lateral view, the presence of both the medial and lateral bone bridges in the AP view was labeled as a union.[4,5,7,15] A *lag screw cut-out* was defined as a protrusion of the screw past the subchondral bone on either the AP or lateral projection (Fig 2).[8,16] Medial migration of the screw past the subchondral bone-recently called "axial migration"-was also labeled as a cut-out.[8]

Statistical analysis

The sample size calculation was based on a report on lag screw cut-outs, which indicated that they represented about 10% of the overall population.[7] A minimum of 139 patients was needed to demonstrate a 10% lag screw cut-out prevalence with a 95% confidence interval and 5% allowable error. Statistical analyses were performed

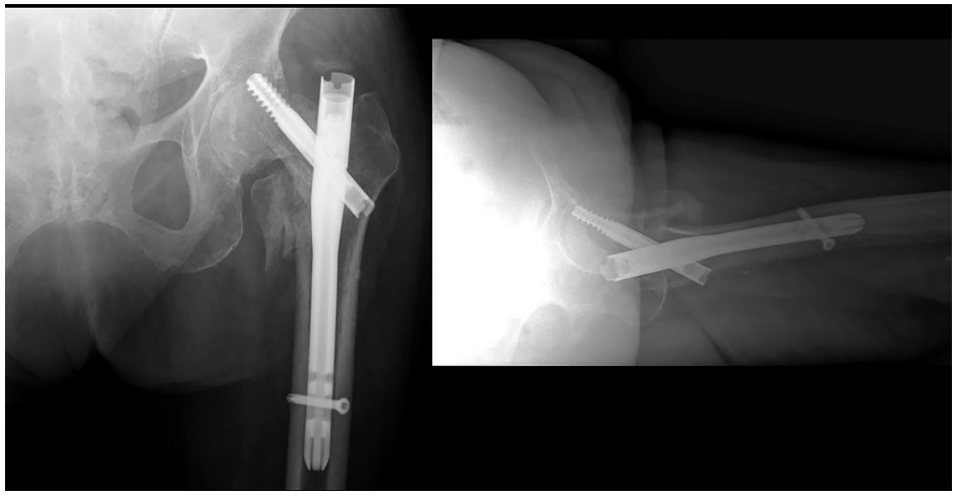


Fig 2. *Left* The radiograph of the affected hip in AP view shows proximal screw cut-out. The proximal screw cut through superolateral aspect of the femoral neck. *Right* The radiograph of the affected hip in the Lateral cross-table view shows proximal screw protrusion past the anterior aspect of the subchondral bone of the femoral head.

using IBM SPSS Statistics for Windows, version 20 (IBM Corp., Armonk, N.Y., USA). The chi-square and independent t-test were used for univariate analyses of the categorical and continuous parameters, respectively, with the significance level (p) < .05. A multivariate analysis was performed for all parameters with $p < .2$.

The inter-observer reliability was calculated for the parameters identified as significant in the univariate analysis. A two-way random-effects model with a 95% confidence interval was used to determine the measurements' intra-class correlation coefficient (ICC). The ICC interpretation scale used was poor to fair (< 0.4), moderate (0.41 to 0.60), excellent (0.61 to 0.80) and almost perfect (0.81 to 1). For categorical data, the κ coefficients were calculated using the same interpretation scale as for the ICC.

RESULTS

Between January 2007 and December 2017, 267 patients diagnosed with a trochanteric fracture treated at Siriraj Hospital, Bangkok, Thailand, with cephalomedullary nails were enrolled. Of those, 92 were excluded: the preoperative radiograph was missing for 7 patients; 6 had a pathologic fracture; the endpoint was unable to be determined for 10 patients due to incomplete radiographs, and 69 were lost to follow-up after discharge from the hospital. The total number of patients included in this study was 175; 154 were successfully treated with a radiographic union of the trochanteric fracture, while 21 (12%) had a lag screw cut-out.

The demographic data are summarized in Table 1. There were no significant differences in the mean ages or the gender proportions of the union and cut-out groups. However, the fracture classifications of the groups were significantly different, with all patients with AO/OTA 31-A1 having been successfully treated without a lag screw cut-out, whereas one-third of the patients with AO/OTA 31-B2.1 experienced a lag screw cut-out. The T-scores of BMD were only available for 40 patients due to the surgeons' judgment to forego a further investigation. Concerned that a large amount of missing data would reduce the study's statistical power, the authors excluded this factor from the multivariate analysis, notwithstanding that it reached a statistically significant level ($p = .048$).

All recorded parameters were evaluated in univariate analysis for an association with lag screw cut-out; the results are in Table 2. The integrities of the basicervical fragment fractures of the union and cut-out groups were statistically significantly different ($p = <.001$). However, the lateral buttress and posteromedial fragment fractures of the two groups showed no significant difference ($p = 1.0$ and .064, respectively).

The postoperative lag screw positions were evaluated in the AP and lateral projections. The screw placement was divided into nine zones (Fig 3). The screws were most frequently placed in the inferoposterior zone. The lowest rate of cut-outs occurred in the center-inferior zone, while the highest cut-out rate was found in the peripheral area, in the superoanterior and superoposterior positions. However, after comparing the cut-out rate of each zone, no statistical significance was found ($p = .053$).

An inter-observer reliability test was performed on the factors with $p < .2$ before a multivariate analysis was carried out; all factors demonstrated good inter-observer reliability (Table 3). The results of the multivariate regression analysis are in Table 4. The integrity of the basicervical

TABLE 1. Summary of demographic data.

Factor	All (n=175)	Union (n=154)	Cut-out (n=21)	*P*-value
Age	80.45 ± 10.17	80.21 ± 10.01	82.24 ± 11.35	.39
Gender				.62
Male	31.4% (55)	32.5% (50)	23.8% (5)	
Female	68.6% (120)	67.5% (104)	76.2% (16)	
Fracture classification				<.0001
31-A1	8.6% (15)	9.7% (15)	0%	
31-A2	75.4% (132)	77.3% (119)	61.9% (13)	
31-A3	8% (14)	8.4% (13)	4.8% (1)	
31-B2.1	8% (14)	4.5% (7)	33.3% (7)	
Total T-score of hip BMD*	-2.6 ± 1.0	-2.4 + 1	-3.3 + 0.7	.048

*Data were available for only 40 out of the total of 175 patients.

TABLE 2. Summary of associated factors evaluated in univariate analysis.

Factor	All (n=175)	Union (n=154)	Cut-out (n=21)	*P*-value
Lateral buttress fracture	18.3% (32)	18.2% (28)	19% (4)	1.0
Posteromedial fragment fracture	80% (140)	81.8% (126)	66.7% (14)	.142
Basicervical fragment fracture	20% (35)	15.6% (24)	52.4% (11)	<.001
TAD*	19.62±5.53	19.53±5.48	20.3±5.97	.56
Neck-shaft angle	137.25±7	137.15±6.75	137.95±8.79	.62
Medial gap	2.07±2.4	2.04±2.35	2.27±2.71	.68
Distal locking screw				.62
Single screw	93.7% (164)	94.2% (145)	90.5% (19)	
Two screws	6.3% (11)	5.8% (9)	9.5% (2)	
Distal locking mode				.49
Static	54.3% (95)	53.2% (82)	61.9% (13)	
Dynamic	45.7% (80)	46.8% (72)	38.1% (8)	

*Post-operative lateral films were retrieved for 143 patients. The missing data were not included in the analysis.

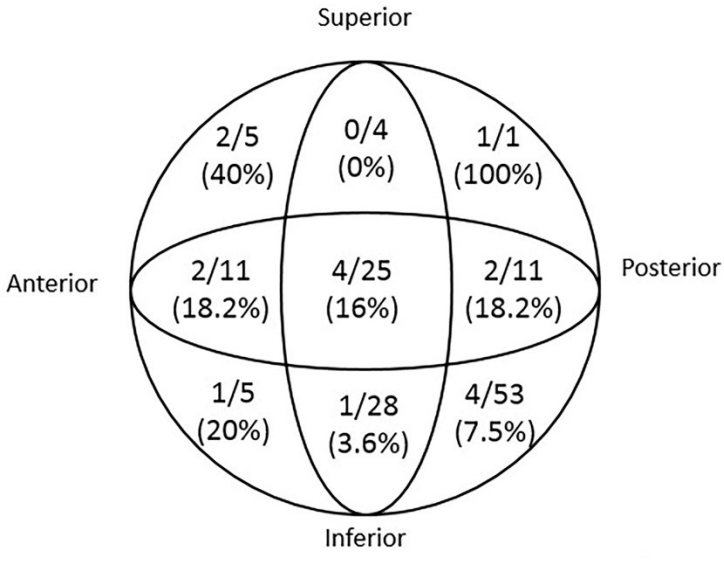


Fig 3. Summary of screw position: The total number of screws in each position was represented by the denominator, and the number of the screw(s) cut-out was represented by the numerator. Post-operative lateral films were retrieved for 143 patients. The missing data were not included in the analysis.

TABLE 3. Summary of inter-observer reliability tests.

Factor	Measure of agreement (κ)
Fracture classification	0.932
Posteromedial fragment fracture	0.888
Basicervical fragment fracture	0.982
Screw position	0.867

TABLE 4. Summary of multivariate analyses by odds ratio.

Factor	*P*-value	Odds ratio	95% confidence interval for odds ratio
Fracture classification	<.001	10.5	[3.22, 34.25]
Posteromedial fragment fracture	.62		
Basicervical fragment fracture	.001	5.51	[2.01, 15.12]
Screw position	.094		

fragment fractures and the fracture classification reached a significant level. Those patients with AO/OTA 31-B2.1 had a significantly high rate of lag screw cut-out (OR 10.5, [3.22, 34.25] $p < .001$). The disintegration of the basicervical fragments was associated with lag screw cut-out (OR 5.51, [2.01, 15.12] p = .001), whereas neither posteromedial fragments nor the lag screw position showed a significant association.

DISCUSSION

The fixation of choice for trochanteric fractures has shown a paradigm shift toward cephalomedullary nails (CMNs) from sliding hip screws (SHSs).[8] The advantages of CMNs include; more biomechanical stability, less invasive, and safer for early weight-bearing ambulation than SHSs.[2] Another factor encouraging using CMNs instead of SHSs is the fracture configuration.[2,4] To date,

the fragments associated with fixation failure in SHS are posteromedial and lateral buttress fragments.[10] With the size of the geriatric population steadily growing, patients with trochanteric fractures increasingly present with osteoporotic bones and comminuted configurations. With such complex fracture configurations, fixation with an SHS has a higher failure rate than with a CMN.

Although the design of CMNs has improved over several generations, the fixation failure rate remains nearly the same. Our study showed a cut-out rate of 12%, which was similar to that reported by another study. Aside from the patient factors, surgical technique plays an important role in predicting the outcomes. The tip-apex distance (TAD) has been proposed to predict a failed SHS or CMN fixation.[1,4,9] With this concerning factor of TAD, careful placement of lag screw during surgery was attempted in all cases of this series; patients in both groups treated in our institute had a mean TAD value of 20 mm. Thus, our study revealed no significant association between the TAD and lag screw cut-outs. Reduction quality also plays a vital role in achieving bone union. Varus mal-reduction or wide posteromedial gap resulted in unstable fixation and increased the cut-out rate. This study's overall reduction quality was good to excellent, with minimal medial gap and slightly valgus neck-shaft angle. From this result, the authors conclude that the surgical technique was appropriate in both the union and cut-out groups, with an insignificant difference in the reduction quality.

Proximal screw position also has been reported to have an association with lag screw cut-outs. The most common direction of the cut-outs in the anterosuperior portion of the femoral head[17,18]; although placing the screw in this position might increase the risk of lag screw cut-outs, the results of the multivariate analysis in this study did not reach the significant level.

Most hip fragility fractures tend to be comminuted, classified as AO/OTA 31-A2. About half of the patients with a lag screw cut-out fall into this classification. Theoretically, a complex fracture configuration tends to give poor results after fixation.[4,7,8] Interestingly, when the proportions of unions to cut-outs are compared, simple basicervical trochanteric fractures-classified as 31-B2.1-have a 50% cut-out rate, while comminuted fractures have only a 9.8% cut-out rate. In our study, simple trochanteric fractures-AO/OTA 31-A1-showed promising results with cephalomedullary nails, with zero cut-outs. After analyzing specific fragment integrity, the authors found an association between basicervical fragment fractures and five times the odds of a lag screw cut-out. Based on this result, the authors believe that a specific fragment has more influence on the results than the comminution of fracture.

Bone mineral density has been reported as a factor associated with lag screw cut-outs.[2] Osteoporosis decreases the density of the cancellous bone in the femoral head and decreases the pull-out strength after fixation, which results in fixation failure. The current study showed a significant association with BMD; nevertheless, due to the large amount of missing data that would have resulted in an underpowered analysis, the authors cannot draw any conclusions about the effects of BMD on lag screw cut-outs.

The dilemma in choosing a fixation implant probably stems from the intraoperative additional fragment fracture. Because lateral buttress disintegration can cause a fixation failure if SHS is chosen, an intraoperative lateral wall fracture in the osteoporotic bone can be a significant drawback, influencing surgeons to choose CMN fixation.

The most suitable treatment for basicervical fractures is a contentious issue. Because the fracture line stays between the femoral neck and the intertrochanteric region, the treatment can be either for a femoral neck or a trochanteric fracture. A previous study described this area as an "extracapsular area," which should be considered for treatment as a trochanteric fracture.[19] Still, as the configuration is comprised mainly of the femoral neck, the motion of the basicervical fragment after fixation behaves as a femoral neck fracture with high rotational instability. Massoud[19] reported an excellent result with no cut-out of the lag screw using SHS combined with a derotation screw as the fixation of choice for this type of fracture as it provides an interfragmentary compression effect along with rotational stability via the additional screw. The results of the present study support this concept since the CMN did not provide a lag effect, and there was no room for an additional screw to improve the rotational stability, resulting in a higher rate of lag screw cut-out after fixation for simple basicervical fractures. The rate of screw cut-out for basicervical fractures was also higher than those for other types of trochanteric fractures in a large case series of trochanteric fracture fixations using a gamma nail, reported by Bojan et al.[7] Recently, Yoo et al. and Yoon et al. have also reported that the basicervical fracture is the most important risk factor for the cut-out problem of cephalomedullary nail fixation in treating the trochanteric fracture in their case series.[20,21] Our findings also support the awareness of using cephalomedullary nails in basicervical fractures. We believe such a gray-zone fracture should be considered

differently from the typical neck or trochanteric fracture. Further investigation for an appropriate fixation device or design should be considered.

The limitations of this study are its retrospective design, the sizeable proportion of excluded patients, and the small number of patients with lag screw cut-outs. The overall trochanteric fracture patient admitted to Siriraj Hospital was about 100 cases per year. Due to the strict criteria, only the patient who received surgical treatment with short CMN were included, resulting in only about 26 cases per year. Those patients who were lost to follow-up possibly moved to another province or may have been inconvenient to visit the hospital.

CONCLUSION

The prevalence of lag screw cut-outs after cephalomedullary nail fixation for trochanteric fractures was 12%-a simple, two-part, basicervical trochanteric fracture has a significantly higher risk of a lag screw cut-out. Using a CMN fixation for basicervical fractures should be avoided.

Disclosure

No benefits in any form related directly or indirectly to the subject of this article have been received or will be received from a commercial party

REFERENCES

1. Baumgaertner MR, Curtin SL, Lindskog DM, Keggi JM. The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip. J Bone Joint Surg Am. 1995;77(7):1058-64.
2. Lorich DG, Geller DS, Nielson JH. Osteoporotic pertrochanteric hip fractures: management and current controversies. Instr Course Lect. 2004;53:441-54.
3. Ninlerd C, Dungkong S, Phuangphay G, Amornsupak C, Narkbunnam R. Effect of home-based rehabilitation exercise program for elderly patients with femoral neck fracture after bipolar hemiarthroplasty. Siriraj Med J. 2020;72(4):307-14.
4. Geller JA, Saifi C, Morrison TA, Macaulay W. Tip-apex distance of intramedullary devices as a predictor of cut-out failure in the treatment of peritrochanteric elderly hip fractures. Int Orthop. 2010;34(5):719-22.
5. Lobo-Escolar A, Joven E, Iglesias D, Herrera A. Predictive factors for cutting-out in femoral intramedullary nailing. Injury. 2010;41(12):1312-6.
6. Riansuwan K, Jivasomboonkul S, Phimolsarnti R, Chandhanayingyong C, Asavamongkolkul A. Outcomes of proximal femoral locking-plate fixation for pathological fractures of the proximal femur. Siriraj Med J. 2021;73(9):603-8.
7. Bojan AJ, Beimel C, Taglang G, Collin D, Ekholm C, Jönsson A. Critical factors in cut-out complication after Gamma Nail treatment of proximal femoral fractures. BMC Musculoskelet Disord. 2013;14:1.
8. Stern LC, Gorczyca JT, Kates S, Ketz J, Soles G, Humphrey CA. Radiographic Review of Helical Blade Versus Lag Screw Fixation for Cephalomedullary Nailing of Low-Energy Peritrochanteric Femur Fractures: There is a Difference in Cutout. J Orthop Trauma. 2017;31(6):305-10.
9. Kashigar A, Vincent A, Gunton MJ, Backstein D, Safir O, Kuzyk PR. Predictors of failure for cephalomedullary nailing of proximal femoral fractures. Bone Joint J. 2014;96-b(8):1029-34.
10. Palm H, Jacobsen S, Sonne-Holm S, Gebuhr P. Integrity of the lateral femoral wall in intertrochanteric hip fractures: an important predictor of a reoperation. J Bone Joint Surg Am. 2007;89(3):470-5.
11. Kane P, Vopat B, Heard W, Thakur N, Paller D, Koruprolu S, et al. Is tip apex distance as important as we think? A biomechanical study examining optimal lag screw placement. Clin Orthop Relat Res. 2014;472(8):2492-8.
12. Davis TR, Sher JL, Horsman A, Simpson M, Porter BB, Checketts RG. Intertrochanteric femoral fractures. Mechanical failure after internal fixation. J Bone Joint Surg Br. 1990;72(1):26-31.
13. Marsh JL, Slongo TF, Agel J, Broderick JS, Creevey W, DeCoster TA, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. J Orthop Trauma. 2007;21(10 Suppl):S1-133.
14. Wroblewski BM, Siney PD, Fleming PA, Bobak P. The calcar femorale in cemented stem fixation in total hip arthroplasty. J Bone Joint Surg Br. 2000;82(6):842-5.
15. De Bruijn K, den Hartog D, Tuinebreijer W, Roukema G. Reliability of predictors for screw cutout in intertrochanteric hip fractures. J Bone Joint Surg Am. 2012;94(14):1266-72.
16. Ciufo DJ, Zaruta DA, Lipof JS, Judd KT, Gorczyca JT, Ketz JP. Risk Factors Associated With Cephalomedullary Nail Cutout in the Treatment of Trochanteric Hip Fractures. J Orthop Trauma. 2017;31(11):583-8.
17. Ehmke LW, Fitzpatrick DC, Krieg JC, Madey SM, Bottlang M. Lag screws for hip fracture fixation: Evaluation of migration resistance under simulated walking. J Orthop Res. 2005;23(6):1329-35.
18. Sommers MB, Roth C, Hall H, Kam BC, Ehmke LW, Krieg JC, et al. A laboratory model to evaluate cutout resistance of implants for pertrochanteric fracture fixation. J Orthop Trauma. 2004;18(6):361-8.
19. Massoud EI. Fixation of basicervical and related fractures. Int Orthop. 2010;34(4):577-82.
20. Yoo J, Chang J, Park C, Hwang J. Risk Factors Associated with Failure of Cephalomedullary Nail Fixation in the Treatment of Trochanteric Hip Fractures. Clin Orthop Surg. 2020;12(1):29-36.
21. Yoon JY, Park S, Kim T, Im GI. Cut-out risk factor analysis after intramedullary nailing for the treatment of extracapsular fractures of the proximal femur: a retrospective study. BMC Musculoskelet Disord. 2022;23(1):107.

Comparative Study Regarding Autonomy of Final-Year Surgical Residents: A Case Study of Perception among Surgical Residents, Surgical Staff, Administrators, and Patients at Siriraj University Hospital

Ravit Ruangtrakool, M.D.*, Oranuch Pruetipibultham**

**Department of Surgery, Faculty of Medicine Siriraj University Hospital, Mahidol University, Bangkok, Thailand, **Graduate School of Human Resource Development, National Institute of Development Administration (NIDA), Bangkok, Thailand.*

ABSTRACT

Objective: To identify barriers towards resident autonomy as perceived through four groups; surgical residents, surgical staff, administrators, and patients.

Materials and Methods: Anonymous surveys were distributed to these four groups. Data were thematically analyzed.

Results: 401 responses were collected including 231 patients. The response rate of residents, surgical staff, and administrators was 62.2% (119), 44.8% (26), and 43.1% (25) respectively. Patients had more favorable views of resident participation than administrators and surgical staff. Administrators and surgical staff indicated that residents have a positive effect on overall quality of care provided and so do the patients, however, administrators and surgical staff believed that too much autonomy for a resident deceased patient safety. When resident autonomy increased, increased cost of patient care was considered. Residents and patients have the same opinion that patients should receive a discount on medical expenses, which is opposite to administrators' and surgical staff's opinion. The presence of surgical staff in the operation room had a major impact on resident autonomy and a big influence on patient acceptance of operative complications. Even in complicated operations, most patients felt comfortable having a resident perform on with surgical staff controlling the operation. Surgical staff provided too much direction in either patient care or operation and did not take residents' input as seriously as expected and seldom explained the reasons before changing treatment regimens.

Conclusion: Surgical residents, surgical staff, and patients had discordant perceptions of resident autonomy in many aspects. Self-determination theory should be applied. Scaffolding strategy, mentoring program would be the solutions.

Keywords: Autonomy; independence; resident; surgery (Siriraj Med J 2022; 74: 634-649)

INTRODUCTION

Currently, the problem of a lack of surgical skills, including pre- and post-operative care is being studied. This could possibly be the result of the fact that surgical residents have little opportunity to care for patients requiring surgery from the time they are a medical student.[1] More importantly, during the period in which surgical residents practice in provincial hospitals after completing their MD, they do not have the opportunity to practice surgical skills.[2] Another factor to consider is

Corresponding author: Oranuch Pruetipibultham
E-mail: oranuch.p@nida.ac.th
Received 9 May 2022 Revised 8 September 2022 Accepted 9 September 2022
ORCID ID:http://orcid.org/0000-0003-1037-5730
http://dx.doi.org/10.33192/Smj.2022.75

that the Department of Surgery, the Faculty of Medicine at Siriraj Hospital offers various training opportunities in various sub-disciplines following the completion of the Diplomate of the Thai Board of Surgery. These sub-disciplines take another one to two years to complete. This means surgical residents who have not completed the Diplomate of the Thai Board of Surgery before, lose the opportunity to develop surgical skills and techniques required in pre and post-operative care. Such problems do not occur only in Thailand, but also in the United States as well.[3]

The ability of surgical residents to have autonomy from surgical staff plays a very important role in the skills needed to train capable surgeons. Once surgical residents have autonomy, they are more willing to take on responsibility and work hard to become capable surgeons. Any competent and skilled surgical resident to care for surgical patients will be accepted by surgical staff and then this surgical resident would be given the autonomy to make treatment decisions and operate on patients. However, surgical residents who do not take care of patients or do not have the required surgical skills as expected of a staff member, will not be given autonomy, which can lead to an inability to develop skills in caring for patients and surgical techniques. Without the trust of surgical staff, a resident will not be allowed to operate by himself or herself. Hence, the skills of a surgical resident to care for a patient cannot be developed. *Sterkenburg et al*[4] reported that when surgical residents are given greater autonomy than they expected to make treatment decisions, they are more confident in learning, maintaining, and practicing which leads to increased competency. A review by *Kempenich JW*[5] stated that 47% of surgical residents and 38% of surgical staff commented that surgical residents did not have sufficient independence to make treatment decisions.

Factors affecting the lack of autonomy and/or lack of surgical residents' independence in decision making were reviewed.

1. Perspectives of surgical staff who teach and train[6]

1.1 Surgical treatment. If a surgical resident is capable and has good surgical skills, he/she is given the autonomy to operate on a patient. However, this depends not only his/her surgical skill, but also on the difficulty of the procedure.[5,7] On the other hand, surgical residents who don't have the skills to perform a difficult operation should not be given the autonomy to perform operations on their own.

1.2 Pre and post-operative care. Any surgical resident who doesn't have the skill to handle pre and post-operative care of patients, is not given autonomy by surgical staff, making it impossible to develop the skills needed to care for patients. The more trust surgical staff have in a resident, the greater the likelihood of surgical residents developing adequate skills.[6,8]

2. Administrator's perspective

Administrators were more likely to prevent the degree of autonomy surgical residents expected.[5] It may be related to the following factors:

2.1. Finance and reimbursement. The administrator has to limit payments of surgical treatments. There are many sources of reimbursement, such as the Thailand National Health Services, Social Security Fund, or other measures. If surgical residents have too much autonomy, it will increase cost of treatment, either due to prolonged operative time and costs incurred because of an increase in complications caused by surgical residents. On the contrary, the administrator has the duty to strike a balance between autonomy of surgical residents as well as costs incurred by them. Financial issues are another factor in the regulation of autonomy afforded to surgical residents.[9,10]

2.2. Legal prosecution. If a surgical resident was granted too much autonomy without adequate supervision, a decision-making error or mistake in surgical technique can lead to a civil or criminal case. Therefore, legal prosecution is another factor to consider when regulating the autonomy of surgical residents.[11]

2.3. Regulation of resident training standards. Regulatory systems aim to manipulate the ability of surgical residents during training such as those established by the Royal College of Surgeons of Thailand, Advanced Hospital Accreditation or other standards like the World Federation for Medical Education (WFME). These factors relate to the autonomy of surgical residents and define what they can and cannot do during any operation.[12]

Some regulations of resident training standards also encourage surgical staff to remain in the operating room at all times while the resident is performing the operation on his/her own, which further leads to reduced autonomy of surgical residents.[9,12]

3. Patient and relatives' point of view

3.1. Safety was considered by patients and their relatives. The patient and their relatives view the presence of a surgical resident to have an effect on safety.[13] According to the *Kempenich JW*[5] study, 95% of patients felt good to have a surgical resident taking part in the care process.

3.2. Complexity of the procedure. The complexity and difficulty of the operation was a major consideration. Patients become worried if surgical residents are allowed to perform complex operations.[5] *Kim HN, et al*[14] reported

on patients who received a hysterectomy by a team of obstetrics and gynecologists. The study revealed that up to 80% of patients wanted to know to what extent gynecologist residents would be involved in the operation and how the gynecologist staff selected this extent. Moreover, 61% of patients wanted to know which part of the gynecology operation was performed by gynecologist residents.[14]

Different points of view between patients and surgical residents were realized. Although 80.6% of patients stated that surgical residents should be able to perform surgeries on their own before graduating from resident training, only 73.1% agreed to allow final year residents to perform a surgery on them even if it was a basic operation.[5]

The question of independence mostly focuses on decision-making and to what extent should surgical residents be given freedom during surgery. Therefore, the focus of this research is to compare the viewpoints of surgical decision-making among relevant stakeholders, namely surgical residents, surgical staff, administrators, and patients and relatives. The results of the study look at factors affecting the autonomy of surgical residents and could be used as academic information to formulate a policy for the surgical resident training program at the Faculty of Medicine, Siriraj Hospital, Mahidol University.

MATERIALS AND METHODS

Siriraj Institutional Review Board Approval (Si 812/2021) was obtained before commencement of the study. For this study, four questionnaires were created to determine how much autonomy and independence in regards to decision-making should surgical residents be given during an operation to compare the perceptions of relevant stakeholders, namely surgical residents, surgical staff, administrators, and patients and relatives. The questionnaire was created by modifying the questionnaire in *Kempenich JW*[5] and *Biondi EA*[6] but the researcher also added more questions in this survey (see appendix). Each survey was specific to the appropriate group, but the questions were similar to those given to other groups for comparison purposes. A five-point Likert scale was used for all questions except two which asked the participants to rate the degree of appropriate independence on a scale from 0 to 10 for a second-year resident and a final year resident in the process of completing the Diplomate of the Thai Board of Surgery.

The surgical resident and staff surveys contained 23 questions, while the administrator survey had 14 questions and patient and relative survey had 16 questions. Each survey was piloted and feedback was solicited from representative individuals in each group. Responses from the surgical resident group were collected using a Google form, while responses from the surgical staff and administrator groups were collected by mail. Last but not least, responses from patients and their relatives were collected directly in person at surgical OPD, and no personally identifiable information was collected. All responses of all groups were collected anonymously. There was no compensation or reward for participation.

Inclusion criteria

1. Surgical resident refers to full-time residents training in all disciplines at the Department of Surgery, Faculty of Medicine Siriraj Hospital (190 persons).

2. Surgical staff refers to staff and Clinical Educators in the Department of Surgery, Faculty of Medicine Siriraj Hospital who presently teach and train surgical residents (58 persons). Surgical staff who are members of the Postgraduate Education Committee of the Department of Surgery were counted as administrators.

3. The administrator refers to positions that control, supervise, train surgical residents and regulate patient safety standards. The administrators include: Head of Department of Surgery, Head of all divisions of the Department of Surgery, all members of the Postgraduate Education Committee of the Department of Surgery, Associate Dean for Postgraduate Studies, Assistant Dean for Postgraduate Studies, Director of Siriraj Hospital, Deputy Director of Siriraj Hospital, Deputy Dean for Quality Development, Deputy Dean for Human Resources, Head of Surgical and Orthopedic Nursing, Head of Operating room Nursing, Head of Disciplinary and Legal Affairs Unit. The total number of administrators was 58.

4. The patient refers to the sample of patients and relatives (over 18), and nonmedical persons receiving care at the Department of Surgery, Faculty of Medicine, Siriraj Hospital.

Three experts checked the validity of the content of questionnaires of surgical residents, surgical staff form, administrators' form. The Content Validity Index (CVI) of these questionnaires was 0.80.

For the reliability of questionnaires, the Cronbach's alpha coefficient (a) was used, with the researcher setting an acceptable coefficient of 0.70 or higher (H_0: α=0.70). The reliability of questionnaires using the Cronbach's alpha coefficient (α) was determined in the surgical resident and surgical staff group, who had a Cronbach's alpha coefficient of 0.87 and 0.8, respectively. The reliability of patient and relatives questionnaire was not tested due to ethical problems.

Statistical analysis

In regards to descriptive statistics, qualitative data, including number and percentage was used. A comparison of scores of relevant questions among surgical residents, surgical staff, administrators and patients and relatives was conducted. The responses were compiled for each survey group, and the median response for each question calculated. Subsequently, the Kruskal-Wallis test was used to analyze the distribution of responses between the survey groups. In cases where a significant difference was identified via the Kruskal-Wallis test, a Dunn Bonferroni's post-hoc pairwise comparison test was used to compare variables among groups.

In certain cases, in which the questions compared two separate questionnaires, especially from the perception of surgical resident and surgical staff, the Mann-Whitney U test was used.

If some questions in the questionnaire regarding patients and relatives were not comparable to other groups, the researcher used descriptive statistics to present this information.

Analysis was done using SPSS Inc or PASW Statistics for Windows, Version 22.0, released in 2009. All data was analyzed and determined as statistically significant when the p-value was < 0.05.

RESULTS

A total of 401 responses were collected; 119 from surgical residents, 26 from surgical staff, 25 from administrators, and 231 from patients and relatives. The response rate for surgical residents, surgical staff, and administrators was 62.2%, 44.8%, and 43.1%, respectively. The number of years of experience working at Siriraj Hospital among administrators, surgical staff, and surgical residents is displayed in Table 1. Patients and their relative's ages are shown in Table 2.

A. Perception of resident participation on quality of patient care

1) How welcome resident participation was in the healthcare process is show in Table 3.

Patients and relatives were more welcoming of resident participation than administrators, surgical staff and surgical residents. The Kruskal-Wallis test showed that there was a statistically significant difference in perception of having surgical residents participate in patient care among the different populations ($p < 0.000$). Pairwise comparisons using the Dunn Bonferroni test revealed that patients and relatives had a significantly higher rating for patient involvement than administrators ($p = 0.031$), surgical staff ($p = 0.020$), and surgical residents ($p < 0.000$). This indicated that patients and relatives have more favorable views of resident participation in patient care than administrators and surgical staff. Surgical residents do not realize how welcoming patients and relatives are of their participation during treatment. The effect on overall quality of care when a surgical resident was involved in the patient care process in the hospital either in OPD, IPD, or in the operating room is shown in Table 4.

When patients and relatives were asked whether quality of care was better with residents involved, 80.9% agreed or strongly agreed and only 3.6% strongly disagreed or disagreed.

2) A comparison of perception amongst administrators, surgical staff, and surgical residents was done to understand the effect of resident participation on quality of care provided. It was divided into two aspects, OPD/IPD and operation room.

In both the OPD/IPD and operating room comparison, the Kruskal-Wallis test showed a statistically significant difference in quality-of-care perception among the different populations (administrators, surgical staff, and surgical residents) with a p-value of <0.000. In OPD/IPD, pairwise comparisons using the Dunn Bonferroni test revealed administrators and surgical staff reported significantly higher effects on overall quality of care provided in hospitals than surgical residents ($p = 0.006$ and $p < 0.000$ respectively). In the operating room, pairwise comparisons using the Dunn Bonferroni test revealed administrators and surgical staff rated a significantly higher effect in the hospital than surgical residents ($p < 0.000$ and $p < 0.000$), respectively.

This information indicates that surgical residents have a positive effect on the overall quality of care in patient care. This valuable effect on overall quality of care was realized and admired by all stakeholders, with the exception of surgical residents themselves.

3) Increasing resident autonomy has impacts on patient safety.

Whether increasing resident autonomy has an impact on patient safety is demonstrated in Table 5.

The Kruskal-Wallis test showed a statistically significant difference in perception in whether increasing resident autonomy would impact on patient safety among different populations (administrators, surgical staff, and surgical residents) with a p-value = 0.001. A Dunn Bonferroni's post-hoc pairwise comparison test was performed (administrator – surgical resident, pairwise comparison $p = 0.021$, surgical staff – surgical resident, pairwise comparison $p = 0.027$). This information indicated that surgical residents believed the more autonomy they

TABLE 1. Years of experience at Siriraj Hospital of administrators, surgical staffs, and surgical residents.

Experience in years at Siriraj Hospital							
	Administrator As an administrator who regulates training		Surgical staff As a surgical staff who trains and teaches surgical residents		Surgical resident As a surgical resident		
	N	%	N	%	Resident years	N	%
0-5	4	16.0	6	23.1	1	29	24.4
6-10	8	32.0	2	7.7	2	33	27.7
11-15	6	24.0	6	23.1	3	18	15.1
16-20	2	8.0	5	19.2	4	27	22.7
> 20	5	20.0	7	26.9	5	12	10.1
Total	25	100.0	26	100	Total	119	100

TABLE 2. Patient and relatives' ages.

Patient and relatives' ages		
Patient and relatives' ages	N	%
18-30	26	11.9
31-40	51	23.4
41-50	38	17.4
51-60	42	19.3
> 60	61	28
Total	218	100

TABLE 3. How welcoming patients and relatives were of resident participation in their healthcare.

Receptiveness of patients and relatives towards resident participation in their healthcare			
	Median*	Mean Rank**	n
Administrator	4 (3,5)	174.28	25
Surgical staff	4 (2,5)	172.44	26
Surgical resident	4 (2,5)	144.22	119
Patient and relative	4 (1,5)	235.64	230
	Median (range)	p < 0.000**	

*A five-point Likert scale

**The Kruskal-Wallis test

Dunn Bonferroni's pairwise comparison (administrator – patient, pairwise comparison p = 0.031), (surgical staff – patient, pairwise comparison p = 0.020), (surgical resident – patient, pairwise comparison p < 0.000)

TABLE 4. Effect of surgical residents' involvement in patient care on the overall quality of care provided in the hospital.

Patient's perception	n	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
My overall hospital and surgical care is of better quality with residents involved	225	2 (0.9%)	6 (2.7%)	35 (15.6%)	120 (53.3%)	62 (27.6%)

Perception of administrators, surgical staffs, and surgical residents were queried regarding the effect of resident participation in patient care and the effect on quality of care provided.

	in OPD and IPD		
	Median*	Mean Rank**	n
Administrator	5 (4,5)	106.86	25
Surgical staff	5 (4,5)	116.42	26
Surgical resident	4 (3,5)	74.26	119
	Median (range)	p < 0.000**	

*A five-point Likert scale
**The Kruskal-Wallis test
Dunn Bonferroni's pairwise comparison
(administrator – resident, pairwise comparison p = 0.006)
(surgical staff – resident, pairwise comparison p < 0.000)

	in the operating room		
	Median*	Mean Rank**	n
Administrator	5 (3,5)	112.42	24
Surgical staff	5 (4,5)	116.35	26
Surgical resident	4 (1,5)	72.62	119
	Median (range)	p < 0.000**	

*A five-point Likert scale
**The Kruskal-Wallis test
Dunn Bonferroni's pairwise comparison
(administrator – resident, pairwise comparison p < 0.000)
(Surgical staff – resident, pairwise comparison p < 0.000)

TABLE 5. Increasing resident autonomy has an impact on patient safety.

	Median*	Mean Rank**	n
Administrator	3 (2,5)	64.50	25
Surgical staff	3 (1,5)	65.85	26
Surgical resident	4 (2,5)	94.21	119
	Median (range)	p = 0.001**	

*A five-point Likert scale
**The Kruskal-Wallis test
Dunn Bonferroni's post-hoc pairwise comparison
(administrator – resident, pairwise comparison p = 0.021),
(surgical staff – resident, pairwise comparison p = 0.027)

had, the more safe care patients received. This perception was different from administrators and surgical staff who believed that too much autonomy for a surgical resident would decrease patient safety.

B. Administrator, surgical staff, and patient and relatives views of autonomy of residents in training

The question whether a surgical resident should perform procedures independently prior to completing residency and entering independent practice is shown in Table 6.

The Kruskal-Wallis test showed a statistically significant difference in perception of whether a surgical resident should perform procedures independently prior to graduating residency and entering independent practice among certain groups, with a p-value = 0.016. A Dunn Bonferroni's post-hoc pairwise comparison found a difference in perception only between surgical residents and patients and relatives (pairwise comparison

TABLE 6. A surgical resident should perform procedures independently prior to graduating residency and entering independent practice.

	Median*	Mean Rank**	n
Administrator	4 (2,5)	215.38	25
Surgical staff	4 (1,5)	173.92	26
Surgical resident	4 (2,5)	224.75	119
Patient and relative	4 (1,5)	189.34	230
	Medan (range)	p = 0.016**	

*A five-point Likert scale
**The Kruskal-Wallis test
Kruskal-Wallis test followed by Dunn Bonferroni's post-hoc pairwise comparison
(surgical resident – patient, pairwise comparison p = 0.024)

p = 0.024). It means surgical staff and administrators agree with patients and relatives regarding the goal of surgical resident training. The perception of whether a surgical resident should perform procedures independently prior to graduating residency and entering independent practice was in agreement among administrators, surgical staff, and patients and relatives.

The questions about if "resident independence" or autonomy is required to ensure a resident becomes a competent surgeon" are shown in Table 7. It reveals that administrators, surgical staff, surgical residents, all agreed that residents require autonomy during training to become a competent surgeon.

TABLE 7. Resident independence or autonomy is required for their development into a competent surgeon.

	Median*	Mean Rank**	n
Administrator	4 (2,5)	103.28	25
Surgical staff	4 (2,5)	78.42	26
Surgical resident	4 (2,5)	83.31	119
	Median (range)	p = 0.081**	

*A five-point Likert scale
**The Kruskal-Wallis test

Calibrating the autonomy of surgical residents was attempted by asking stakeholders to rate on a scale of 0 to 10 (0 being no independence and 10 the most) after various years of training. The researcher chose second-year residents because this was the first step of a resident having to adhere to his or her surgical sub-specialty and final year residents about to graduate and receive the Diplomate of the Thai Board of Surgery the following year. Administrators, surgical staff and each surgical resident were asked to rate on a scale of 0 to 10 (0 being no independence and 10 the most) the appropriate degree of autonomy for a second-year resident and a final year resident. The results are shown in Table 8.

The median scale of autonomy of final year residents rated by surgical residents was 7, whereas administrators and surgical staff rated the median scale of autonomy of a final year resident to be 8. The Kruskal-Wallis test was performed to evaluate the distribution of responses and there was no significant difference among these three groups. The surgical residents rated the median scale of autonomy of second-year residents as 5, whereas administrators and surgical staff rated the median scale of autonomy of the second-year residents as 5 and 3, respectively. The autonomy of surgical residents increased when surgical residents passed each year of training. Finally, all groups rated an appropriate level of autonomy for final year residents, without any statistic significant differences.

C. Administrators' role. Perception of the effect of regulation on reimbursement and liability concerns for the hospital and surgical resident autonomy

Administrators, surgical staff, and surgical residents were asked if they felt regulations regarding reimbursement were responsible for decreased autonomy. Perceptions of the effect of regulations on reimbursement on resident autonomy were investigated. The question if an increase in resident autonomy led to an increase in patient cost care is shown in Table 9. This information indicated that all three groups agreed that when resident autonomy increased, it was likely to increase cost of patient care.

Responses to the more serious question "if patients should get a discount on medical expenses when a surgical resident does an operation?" are shown in Table 10. Surgical residents, and patients and relatives had the same opinion that patients should get a discount on medical expenses, which was the complete opposite response provided by administrators and surgical staff. The Kruskal-Wallis test was performed to evaluate this opinion and the distribution of responses were significantly different ($p < 0.000$). Pairwise comparisons using the Dunn Bonferroni

TABLE 8. Appropriate level of independence for final year surgical residents and a second-year surgical resident on a scale of 0-10.

(No independence -> Fully independent (0-10))

	Final Year Resident Median*	Final Year Resident Mean Rank**	n	Second Year Resident Median*	Second Year Resident Mean Rank**	n
Administrator	8 (3,10)	99.58	25	5 (1,7)	75.50	25
Surgical staff	8 (4,9)	86.38	26	3 (0,8)	57.17	26
Surgical resident	7(3,10)	82.35	119	5 (1,9)	93.79	119
	Median (range)	p = 0.254**		Median (range)	p = 0.001***	

*A scale from 0 to 10
**The Kruskal-Wallis test
***Kruskal-Wallis test followed by Dunn Bonferroni's post-hoc pairwise comparison was used to compare continuous variables among second-year residents. (surgical staff - surgical resident, pairwise comparison p = 0.003)

TABLE 9. The question whether increasing resident autonomy would lead to increased cost of patient care.

	Median*	Mean Rank**	n
Administrator	3 (1,5)	80.86	25
Surgical staff	3 (1,4)	70.38	26
Surgical resident	3 (1,5)	89.78	119
	Median (range)	p = 0.138**	

*A five-point Likert scale
**The Kruskal-Wallis test

TABLE 10. Perception of question: "Should patients receive a discount on medical expenses when a surgical resident performs an operation?"

	Median*	Mean Rank**	n
Administrator	2 (1,4)	107.70	25
Surgical staff	2 (1,5)	104.83	26
Surgical resident	3 (1,5)	197.24	119
Patient and relative	4 (1,5)	216.84	222
	Median (range)	p < 0.000**	

*A five-point Likert scale
**The Kruskal-Wallis test
Kruskal-Wallis test followed by Dunn Bonferroni's post-hoc pairwise comparison
(administrator – patient, pairwise comparison p < 0.000)
(administrator – resident, pairwise comparison p = 0.001)
(surgical staff – patient, pairwise comparison p < 0.000)
(surgical staff – surgical resident, pairwise comparison p = 0.001)
(surgical resident – patient, pairwise comparison p = 0.684)

test revealed that surgical staff and administrators have the same opinion that medical expenses should not be reduced while surgical residents and patients (surgical resident - patient, pairwise comparison p = 0.684) both agreed that the cost should be reduced.

The question of whether increased resident autonomy caused liability concerns for the hospital is shown in Table 11. This table indicated that all three groups agreed that legal liability was not a concern when resident autonomy increased (p = 0.396)

D. Patient and relatives' views on resident participation in the health care process

This section will demonstrate patients' perception regarding all aspects of surgical residency, quality of care provided by surgical residents, informed consent of patients' ability to choose a "true" surgeon, and if surgical staff should be present in the operating room even in basic and uncomplicated operations as well

TABLE 11. Whether an increase in resident autonomy caused liability concerns for the hospital.

	Median*	Mean Rank**	n
Administrator	3 (1,5)	73.29	24
Surgical staff	3 (1,5)	84.69	26
Surgical resident	3 (1,5)	87.43	119
	Medan (range)	p = 0.396**	

*A five-point Likert scale
**The Kruskal-Wallis test

as acceptance of complications incurred by surgical residents. Patients and relatives opinions are shown in Table 12.

As long as surgical outcomes are the same or better with resident participation, 86.7% of patients agreed or strongly agreed to have residents involved in surgical care.

However, who the operating surgeon was and his/her experiences were still the main consideration. If patients have an operation performed by a surgical resident, 80.9 % of patients agreed or strongly agreed to wanting to know how many times the resident had done this operation. Another 72.5% of patients agreed or strongly agreed to wanting to choose the surgeon who operated on them.

Patients and relatives felt comfortable to allow surgical residents to perform only basic and uncomplicated operations without the need for surgical staff in the operating room, however, in complicated operations, 74.4% of patients strongly disagreed or disagreed with allowing a surgical resident to perform the operation by himself or herself without the presence of surgical staff in the operating room. In operations of increasing complexity, there was less willingness to allow resident involvement in surgical procedures.

The presence of surgical staff in operations had a major impact on resident autonomy. Even in complicated operations, 79.5% of patients agreed or strongly agreed to allow residents to carry out a complicated operation with surgical staff controlling the operating room.

Attention in the operation by surgical staff had a big influence on acceptance by patients of incurred operative complications. The majority of patients or 81.8%, (44.1% strongly disagreed or 37.7% disagreed) were unwilling to accept serious complications following an operation by a surgical resident without the presence of surgical staff in the operating room. Even for mild complications, 25.5% of patients could accept operations carried out by surgical residents alone.

E. Comparison of perception in improving surgical residents' autonomy in surgical staff and surgical residents

Table 13 summarizes the comparison of perception in increasing surgical residents's autonomy between surgical staff and surgical residents.

Most of the surgical staff and surgical residents opinions were compared in two questionnaires, one looking at the perception of surgical staff and the other at the perception of surgical residents. Differences in surgical staff and surgical residents were statistically significant for 5 of the 9 parallel items. Surgical staff provided too much direction in either patient care activities ($p = 0.004$) or operative procedures ($p < 0.000$).

Surgical staff tend to follow the plan of residents even if they prefer an alternative plan ($p = 0.432$), but when it comes to making important medical decisions, surgical staff did not take the input of surgical residents as seriously as expected ($p = 0.022$). When the surgical staff changed a surgical resident's method of treatment, a reason was seldom given to the resident to explain the change in treatment regimen ($p = 0.001$). Surgical staff gave too little feedback to surgical residents that could help motivate them to improve their performance ($p = 0.001$).

DISCUSSION

Individuals who receive support when given autonomy (e.g., sensitivity to their perspectives, acknowledgment of their feelings, provision of choices, minimization of controls) from important authority figures, are more motivated to pursue their goals, more satisfied with their work and life, and ultimately become high achievers compared to individuals who are forced or persuaded to pursue the goals of others.[15,16] Self-determination theory states that humans have a natural tendency for autonomous behavior, and those who are able to act autonomously learn and perform better.[6,15,17]

The preparedness of surgeons graduate for independent practice was concerned, owing to a lack of autonomy in their training.[18] This may be due to the fact that surgical residents had fewer opportunities to care for patients since their time as a medical student.[1] More importantly, during the time they practiced in provincial hospitals after finishing their MD, these surgical residents did not have the opportunity to practice surgical skills.[2] In addition, the number of residents seeking fellowship training or subspecialist training after residency has increased, which contributes to the major problem of lack of autonomy in residency.[3]

The issue of resident autonomy or independence becomes even more complicated when there is a discussion of how much autonomy should be afforded to a resident. The aim of this study was to evaluate perceptions regarding resident autonomy from hospital administrators, surgical staff, surgical residents, and patients and relatives.

The first part looked at the effect of resident participation on quality of patient care, and contained three questions:

1.1. How welcome resident participation was by patients and relatives.

In this study, patients and relatives welcomed resident participation in their health care process.

TABLE 12. Patient and relatives' opinions.

Patient and relative's opinions	n	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
As long as surgical outcomes are the same or better with resident participation, I agree with having residents involved in my surgical care.	227	2 (0.9%)	7 (3.1%)	21 (9.3%)	144 (63.4%)	53 (23.3%)
If I have an operation done by a surgical resident, I want to know how many similar previous cases he/she has handled before.	225	4 (1.8%)	9 (4.0%)	30 (13.3%)	105 (46.7%)	77 (34.2%)
You want to choose the surgeon who operates on you.	225	2 (0.9%)	14 (6.2%)	46 (20.4%)	83 (36.9%)	80 (35.6%)
If I had to have a basic and uncomplicated surgery, I would consent to a resident performing the operation,						
A. without the need for surgical staff to be in the operating room while the surgical resident is operating.	229	22(9.6%)	95 (41.5%)	36 (15.7%)	61 (26.6%)	15 (6.6%)
B. with surgical staff controlling the operating room.	229	3 (1.3%)	18 (7.9%)	23 (10.0%)	123 (53.7%)	62 (27.1%)
If I had to have a complicated surgery, I would consent to a resident performing the operation						
A. without the need for surgical staff to be in the operating room while surgical resident is operating.	230	89 (38.7%)	82 (35.7%)	21 (9.1%)	26 (11.3%)	12 (5.2%)
B. with surgical staff controlling the operating room.	205	8 (3.9%)	25 (12.2%)	9 (4.4%)	91 (44.4%)	72 (35.1%)
I could accept mild complications following an operation by a surgical resident without any surgical staff in the operating room.	204	30 (14.7%)	89 (43.6%)	33 (16.2%)	43 (21.1%)	9 (4.4%)
I could accept serious complications following an operation by a surgical resident without any surgical staff in the operating room.	204	90 (44.1%)	77 (37.7%)	6 (2.9%)	24 (11.8%)	7 (3.4%)

TABLE 13. Comparison of perceptions in improving surgical residents' autonomy between surgical staffs and surgical residents.

	Surgical staff			Surgical resident			
	Surgical staff (Median (Range))*	Surgical staff (Mean Rank)**	n	Surgical resident (Median (Range))*	Surgical resident (Mean Rank)**	n	Mann-Whitney test
Residents' self-confidence increases when attending staff allows more autonomy.	4 (3,5)	65.92	26	4 (2,5)	74.55	119	p = 0.283
Decreasing resident autonomy leads to a decreased sense of patient ownership.	4 (1,5)	68.02	26	4 (1,5)	74.09	119	p = 0.471
Surgical staff gives too much direction to a surgical resident regarding patient- care activities.	2 (1,5)	52.27	26	3 (1,5)	77.53	119	p = 0.004
Surgical staff gives too much direction to a surgical resident regarding operative procedures.	2 (1,5)	47.92	26	3 (1,5)	78.48	119	p < 0.000
Surgical staff encourages surgical residents to develop an independent thought process.	4 (2,4)	77.90	26	4 (2,5)	71.93	119	p = 0.478
Surgical resident's input is taken seriously by the surgical staff when making important medical decisions.	3 (1,4)	56.98	26	4 (1,5)	76.50	119	p = 0.022
Surgical staff allows the surgical resident to plan even if surgical staff prefers an alternative plan.	3 (1,5)	67.52	26	3 (1,5)	74.20	119	p = 0.432
When the surgical staff change the surgical resident's method of treatment, staff always explains the reasons before changing the treatment regimen.	4 (1,5)	94.88	26	4 (1,5)	68.22	119	p = 0.001
Surgical staff give feedback to surgical residents that helps them feel motivated to improve their performance.	4 (3,5)	96.65	26	4 (1,5)	67.83	119	p = 0.001

*A five-point Likert scale
** Mann-Whitney test

Surprisingly, it was higher than the welcome afforded by administrators, surgical staff, and surgical residents. Pairwise comparisons using the Dunn Bonferroni test revealed that patients and relatives had a significantly higher acceptance rate of resident involvement than administrators ($p = 0.031$), surgical staff ($p = 0.020$) and surgical residents ($p < 0.000$). This indicated that patients and relatives had more favorable views of resident participation in patient care than administrators and surgical staff. Surgical residents do not realize how much patients and their relatives welcome their participation in treatment.

1.2. The effect on overall quality of care when a surgical resident was involved in patient care, either in OPD, IPD, or in the operating room. In our study, 80.9% of patients and relatives agreed or strongly agreed that the quality of care was better when surgical residents were involved. Our study also compared the perceptions of the administrator, surgical staff, and surgical residents regarding the effect of resident participation on quality of care provided in OPD/IPD and the operating room.

In the OPD/IPD and operating room, pairwise comparisons using the Dunn Bonferroni test revealed that administrators and surgical staff reported statistically significant higher scores on overall quality of care than surgical residents did. This information indicates that surgical residents have a positive effect on overall quality of care provided in either OPD/IPD or the operating room. In the study by *Kempenich JW*[5], the team found that only 3% of the general public had responses indicating that residents had a negative effect on quality of care. This effect on overall quality of care was realized and admired by all groups of stakeholders except the surgical resident.

1.3. Increased resident autonomy has an impact on patient safety.

The prevalence of complications due to resident involvement was completely different to the perception of relevant stakeholders, including administrators, surgical staff, surgical residents, and patients.

Regarding the prevalence of complications due to resident involvement, *Castleberry et al*[19] found that although there was increased morbidity with residents involved, the 30-day mortality rate decreased and there was a lower "failure-to-rescue". Other studies found no increase in complications or morbidity with residents involved.[11,20]

In our study, Dunn Bonferroni's post-hoc pairwise comparison tests were performed and they confirmed that surgical residents believed that the higher the level of autonomy, the better the safety profile was for patients. This perception was different from administrators and surgical staff continue to believe that too much autonomy for a surgical resident will decease patient safety. The ones who have the duty to control autonomy of surgical residents believe that too much freedom will lead to less safety for patients.

The second part looked at the views of administrators, surgical staff, surgical residents, and the views of patients and relatives in training. This step focused on perceptions such as essential autonomy required to become a competent surgeon, proper level of surgical resident autonomy, level of autonomy currently present, and development of autonomy. This second step composed of three questions:

2.1. The reasons why a surgical resident should perform procedures independently prior to graduating residency and starting independent practice are shown in Table 6. The Kruskal-Wallis test found that administrators and surgical staff also agreed with patients and relatives about the goals of surgical training. The perceptions of if a surgical resident should perform procedures independently prior to graduating residency and entering independent practice were in agreement between administrators, surgical staff and patients and relatives.

2.2. The question "resident independence or autonomy is required for development of a resident into a competent surgeon" was explored in Table 7. It revealed that administrators, surgical staff, and surgical residents, all agreed that residents require autonomy during training to develop into a surgeon.

Surgical residents who were surveyed felt most strongly about the importance of performing procedures independently before graduation as important.[9] *Kempenich JW*[5] reported that although most of the teaching faculty, administrators, and the general public felt that residents should perform procedures independently before graduation, their responses were less enthusiastic than those of residents. Among the general public, 80.6% agreed or strongly agreed that residents need to perform procedures independently before graduation. When asked "should the surgical resident be able to perform the surgery on his/her own confidently before graduating resident training", 96% of surgical residents' were in total agreement or agreement, however, from the point of view of surgical staff, that was not as important. When patients were asked if they would consent to a final year resident performing a routine procedure independently, less of them agreed (73.1%; $p = 0.05$).[5]

2.3. Appropriate level of independence for final year surgical residents and second-year surgical resident on a scale of 0-10. In our study, the autonomy of surgical residents could be calibrated by asking them to rate on a scale of 0 to 10 (0 being no independence and 10 the

most) different years of resident training. The median scale of autonomy of final year residents was rated by an administrator, surgical staff, and surgical resident. The Kruskal-Wallis test was done to evaluate the distribution of responses and there was no significant difference among the three groups. The surgical residents rated median scale of autonomy of second-year residents as 5, whereas administrators and surgical staff rated the autonomy of second-year residents as 5 and 3, respectively. The autonomy of surgical residents increased when surgical residents completed each year of training.[4] Finally, all groups rated the appropriate level of autonomy for final year residents without any statistically significant differences.

The results of how much autonomy should be given to a resident, in our study, were different from results observed by *Sterkenburg et al.*[4] In our study, the final year resident rated autonomy as same as surgical staff whereas *Sterkenburg et at*[4] revealed that surgical staff consistently rated the appropriate level of independence as lower than residents. They went on to suggest that: "… it may be necessary for residents to overestimate their ability to stimulate learning."

The third part looked at the role of the administrator. Perceptions of the effect of regulations on reimbursement and liability concerns for the hospital on surgical resident's autonomy were questioned.

When surgical residents perform any operation, the operating time is definitely prolonged. Improper or inadequate treatment decisions, whether in the ward or in the operating room, will increase the cost for patients. If any complications following the operation occur, it will increase medical expenses. In our study, administrators, surgical staff, and surgical residents were asked if they felt regulations regarding reimbursement were responsible for decreased resident autonomy. Our results indicate that all three groups agreed when resident autonomy increased, it was likely to lead to increased cost of patient care.

The question "should patients receive a discount on medical expenses when a surgical resident performs an operation for a patient?" was explored in Table 10. Surgical residents, patients and relatives have the same opinion that patients should receive a discount on medical expenses, which is opposite to the opinion of administrators and surgical staff. Pairwise comparisons using the Dunn Bonferroni test revealed that surgical staff and administrators have the same opinion that medical expenses should not be reduced while surgical residents and patients both agreed that the cost should be reduced. This disagreement is statistically significant. An ethical consideration would be raised if some patients were operated on exclusively by surgical staff, (with the exclusion of any residents, if these patients paid more). Furthermore, how to compensate for higher medical expenses of a patient and relatives should be considered. Most likely, the patient and relatives should not pay extra because of unnecessary medical expenses incurred by surgical residents.

Although *Arriaga AF et al*[11] reported there were liability concerns over complications incurred by surgical residents, it was not clear if it is obstacle for the autonomy of a surgical resident. Our study indicated that all three groups (administrators, surgical staff, surgical residents) agreed that legal liability was not much of a concern as resident autonomy increased ($p = 0.396$).

The fourth part investigated the views of patients and relatives on resident participation.

Patient and relatives' opinions are shown in Table 12.

As long as surgical outcomes are the same or better with resident participation, 86.7% of patients agreed or strongly agreed to have residents involved in surgical care. If guarantees could be provided to patients that outcomes would be the same or better with resident participation, then the general public would be more receptive to their involvement in procedures.

Who the operating surgeon is and how much experience he/she has is still the main consideration. If patients had an operation done by a surgical resident, 80.9% of patients agreed or strongly agreed to wanting to know how many cases the resident had handled before. A total of 72.5% of patients agreed or strongly agreed to wanting to choose the surgeon who operated on them. In a survey of patients who underwent hysterectomy[14] (where 100% of patients had a resident involved in their care), 80% wanted to know what a resident would do during their operation and how they would be supervised.

Reichgott and *Schwartz*[21] found that most negative responses to resident participation were due to patients not being aware of what part of the process the resident would be involved in their care. Patients often had anxiety and did not want surgical residents to have too much autonomy in an operation. This might be due to the fact that patients did not know the true role of surgical residents at the time of admission. The surgical staff should ensure patients know the role of surgical residents. Attention to this issue could allow the teaching faculty to assuage patient fears.[5,21]

Patients and relatives felt comfortable to let surgical residents perform only basic and uncomplicated operations without the presence of surgical staff in the operating room, but in complicated operations, 74.4% of patients

strongly disagreed or disagreed to let the surgical resident perform the operation alone without the presence of surgical staff in the operating room. Regarding resident involvement in surgical procedures with increasing complexity, the general public was less willing. *Kempenich JW*[5] asked the general public if they would welcome resident participation if they did a routine vs complex surgical procedure, and there was statistically significant negative stance for complex surgical procedures.

In our study, the presence of surgical staff in the operation room had a major impact on resident autonomy. Even in complicated operations, 79.5% of patients agreed or strongly agreed to have a resident perform complicated operations with surgical staff controlling the operating room.

The attention of the surgical staff in the operation room had a big influence on patient acceptance of operative complications. The majority of patients, or 81.8% (44.1% strongly disagreed or 37.7% disagreed) could not accept serious complications following an operation by a surgical resident without surgical staff supervision in the operating room. Even in cases of mild complications, only 25.5% of patients could accept the absence of surgical staff.

The fifth part investigated differences in perception between surgical staff and residents in improving the autonomy of a surgical resident.

Surgical staff and surgical residents both endorsed autonomy as important for the development of residents. Surgical staff provided too much direction in either activities related to patient care ($p = 0.004$) or operative procedures ($p < 0.000$). Over direction by surgical staff may make residents more passive, and this may stimulate faculty to exert even more control. Any regulations have made it necessary to increase the presence of attending surgeons in the OR, and this may cause decreased resident autonomy as the OR staff and residents naturally defer to the attending surgeon. *Chalabian* and *Bremne*[9] performed a survey of surgical residents in the late 1990s which revealed that 71% of residents felt that "…mandatory presence of surgeons in the OR was bad."

When making important medical decisions, surgical staff did not take the input of surgical residents as seriously as expected ($p = 0.022$). When surgical staff changed the surgical resident's method of treatment, the staff seldom explained the reasons to the resident before changing the treatment regimen ($p = 0.001$). Hence, surgical residents may express their frustration about staff restriction on opportunities to make decisions and about unexplained changes in patient care plans. Lack of mutual trust is a serious threat to learning and in-patient care teams. Moreover, problems around communication still persist. Surgical staff do not provide enough feedback to surgical residents to help motivate them to improve performance. ($p = 0.001$)

The consistent difference between surgical staff and surgical residents' perceptions of autonomy suggests that a common underlying factor or set of related factors may drive observed differences. Hence, the researcher explored three possible underlying causes.[6]

1. Challenges to self-determination

Scholars have studied trust in leaders and stated having trust in their team leaders was a key driver for employee commitment and effectiveness.[22-24] Being able to build trust between junior and senior team members fosters organizational effectiveness, positive work attitude, goal acceptance, and better performance.[25,26]

Trust is developed between a resident and surgical staff when a competent resident executes a task appropriately.[8] Not challenging residents with appropriate responsibility and independence can lead to stunted learning and development.[3,4,9] Without trust, it is difficult for surgical staff to extend sufficient autonomy to residents.[8] Lack of mutual trust is a serious threat. In in-patient care teams, a major concern is that when surgical staff restrict the independence of "passive" residents whose competence is questioned, they get fewer opportunities for active learning.

Scaffolding strategy[6] is likely to encourage development of competence and enhance their relationships with team members and supervisors. Residents who are reluctant to act autonomously may benefit from more scaffolding in their education, so that they can gradually gain the confidence they need to assume a more independent role in patient care. *Biondi EA, et al*[6] encourages surgical staff to scaffold the learning of residents they view as passive including novices who are reluctant to take on independent roles by giving them decision-making responsibilities in increasingly complex situations after they prove themselves in less challenging settings.

To enhance the scaffolding scheme, the researcher recommends a series of "mentoring program" to address the communication and trust issues between surgical staff (mentors) and surgical residents (mentees). In medicine, mentoring refers to a dynamic and context-dependent process between experience clinicians and junior clinicians to advance the development of the mentees as well as create relationships with mutual benefit between the two parties.[27] In mentoring, goal setting and elaborating expectation at the beginning of the relationship are essential. Research studies regarding mentoring program

in medical school[28-30] suggested that the key success factors are to provide a choice of mentors who share similar clinical and specialties interest (regardless of differences in generation), uses of questionnaire, profile matching, mentor's professional orientation, work-life priorities and share values and interest between mentor and mentee. Mentoring program not only addresses the trust issue but also handles enhance communication through reflective and honest feedback.

In this study, the researcher recommends that surgical staff (mentors) should be trained to be able to assess surgical residents (mentees)' needs, competencies, experiences, personalities, values, beliefs and goals to create the mentoring environment that personalizes and increases their autonomy.

2. Generational differences

Descriptions of generation Y (individuals born after 1982) showed some of the attitudes expressed by surgical residents.[31,32] For example, surgical residents indicated they would like surgical staff to provide more specific expectations, more support, better explanations when residents' treatment plans were changed, and frequent feedback. Most responding residents belong to Generation Y, but the surgical staffs were a composite of generations.

3. Inaccurate self-assessments on the part of both surgical staff and surgical residents[33,34]

The resident and surgical staff groups sometimes rated themselves higher than other groups and there was a lack of insight into how their own behaviors could drive behaviors of the other group. For example, over direction by surgical staff might make surgical residents more passive, and this may stimulate surgical staff to exert more control.

Our study did have some limitations.

First, the response rate of surgical residents, surgical staff, and administrators was 62.2%, 44.8%, and 43.1%, respectively, however, this does not represent real results due to non-responder bias. The sample size calculation was not available for this study.

Second, generalizability may be restricted. The study was conducted within one residency program where autonomy issues may have elicited differences between residents and faculty that would not be evident elsewhere.

Third, the reliability of questionnaires, using the Cronbach's alpha coefficient (α), of the patients and relatives' questionnaire was not tested due to ethical concerns.

Despite these limitations, the strength of our study are that: 1. It characterizes views of resident autonomy from the perspective of four major groups who have a stake in surgical resident training. 2. This study is a multidisciplinary integration of knowledge between social science and medical education discipline.

ACKNOWLEDGEMENTS

The researcher would like to thank Dr. Sasima Tongsai from the Division of Clinical Epidemiology, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University for her continuous help with data processing and statistical analysis, Mrs. Sirirat Khawsaard and Mrs. Angkana Seetasang, from the Department of Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University for their continuous help with data collection.

Finally, the researcher would like to thank the Faculty of Medicine Siriraj Hospital, Mahidol Univeristy for granting me the full scholarship for enrolling M.S. Program in Human Resource and Organization Development, National Institute of Development Administration, which delivers professional knowledge in social science to conduct this study.

Conflicts of interest: The authors have no conflicts of interest to declare.

REFERENCES

1. Iramaneerat C, Rangsisombatsiri C. Satisfaction Survey of Graduates from the Doctor of Medicine Program of the Faculty of Medicine Siriraj Hospital. Siriraj Med Bull. 2010;3:2.
2. Lertrattananon D, Limsawart W, Dellow A, Pugsley. Does medical training in Thailand prepare doctors for work in community hospitals? An analysis of critical incidents. Human Resources for Health. 2019;17(62):1-8.
3. Lewis FR, Klingensmith ME. Issues in general surgery residency training. Ann Surg. 2012;256(4):553-9.
4. Sterkenburg A, Barach P, Kalkman C, Gielen M, Cate OT. When do supervising physicians decide to entrust residents with unsupervised tasks? Acad Med. 2010;85(9):1408-17.
5. Kempenich JW, Willis RE, Rakosi R, Wiersch J, Schenarts P. How do perceptions of autonomy differ in general surgery training between faculty, senior residents, hospital administrators, and the general public? A multi-institutional study. J Surg Educ. 2015:72(6):193-201.
6. Biondi EA, Varade WS, Garfunkel LC, Lynn JF, Craig MS, Cellini MM, et al. Discordance between resident and faculty perceptions of resident autonomy: Can Self Determination Theory help interpret differences and guided strategies for bridging the divide? Acad Med. 2015;90:462-71.
7. Teman NR, Gauger PG, Mullan PB, Tarpley JL, Minter RM. Entrustment of general surgery residents in the operating room: factors contributing to provision-of resident autonomy. J Am Coll Surg. 2014;219(4):778-87.

8. Cate OT. Trust, competence, and the supervisor's role in postgraduate training. BMJ. 2006;333(7571):748-51.
9. Chalabian J, Bremner R. The effects of programmatic change on resident motivation. Surgery. 1998;123(5):511-7.
10. Santry HP, Chokshi N, Datrice N, Guitron J, Moller MG. General surgery training and the demise of the general surgeon. Bull Am Coll Surg. 2008;93(7):32-8.
11. Arriaga AF, Elbardissi AW, Regenbogen SE, Green- berg CC, Berry WR, Lipsitz S. A policy-based intervention for the reduction of communication breakdowns in inpatient surgical care: results from a Harvard surgical safety collaborative. Ann Surg. 2011;253(5):849-54.
12. Babbott S. Commentary: watching closely at a distance-: key tensions in supervising resident physicians. Acad Med. 2010; 85(9):1399-400.
13. Phitayakorn R, Williams RG, Yudkowsky R, Harris IB, Hauge LS, Widmann WD, et al. Patient-care-related telephone communication between general surgery residents and attending surgeons. J Am Coll Surg. 2008;206(4):742-50.
14. Kim HN, Gates E, Lo B. What hysterectomy patients want to know about the roles of residents and medical students in their care. Acad Med. 1998;73(3):339-41.
15. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol. 2000;55:68-78.
16. Williams GC, Deci EL. The importance of supporting autonomy in medical education. Ann Intern Med. 1998;129:303-8.
17. Kusurkar R, Cate OT. AM last page: education is not filling a bucket, but lighting a fire: self-determination theory and motivation in medical students. Acad Med. 2013;88(6):904.
18. Greene FL. Transition to practice: a bold move. General Surgery News. 2013;40(N3):1.
19. Castleberry AW, Clary BM, Migaly J, Worni M, Ferranti JM, Pappas TN, et al. Resident education in the era of patient safety: a nationwide analysis of outcomes and complications in resident-assisted oncologic surgery. Ann Surg Oncol. 2013;20(12):3715-24.
20. Hernandez-Irizarry R, Zendejas B, Ali SM, Lohse CM, Farley DR. Impact of resident participation on laparoscopic inguinal hernia repairs: are residents slowing us down? J Surg Educ. 2012;69(6):746-52.
21. Reichgott MJ, Schwartz JS. Acceptance by private patients of resident involvement in their outpatient care. J Med Educ. 1983;58(9):703-9.
22. Watson Wyatt Worldwide. Work USA 2000: Employee commitment and the bottom line. Bethesda, MD: Watson Wyatt Worldwide. 2000.
23. Javidan M, Dorfman PW, Sully de Luque M, House RJ. In the eye of the beholder: Cross cultural lessons in leadership from project GLOBE. Academic of Management Perspectives. 2006;2:67-90.
24. Niljaeng N, Pruetipibultham O. An examination of the relationships of global leadership competency, trust in leader, and team process effectiveness in Thailand. Human Resource and Organization Development Journal. 2018;10(1):2-33.
25. Dirks KT. Trust in leadership and team performance: evidence from NCAA basketball. J Appl Psychol. 2000;85(6):1004-12.
26. Fulmer CA, Gelfand MJ. At what level (and in whom) we trust: Trust across multiple organizational level. Journal of Management. 2012;38:1167-230.
27. Wu JT, Wahab MT, Ikbal MF, Loo TWW, Kanesvaran R, Radha Krishna LK. Toward an interprofessional mentoring program in palliative care - a review of undergraduate and postgraduate mentoring in medicine, nursing, surgery and social work. J Palliat Care Med. 2016;6:292.
28. Ramanan RA, Taylor WC, Davis RB, Phillips RS. Mentoring matters. Mentoring and career preparation in internal medicine residency training. J Gen Intern Med. 2006;21:340-5.
29. Dimitriadis K, Borch PVD, Störmann S, Meinel FG, Moder S, Reincke M, et al. Characteristics of mentoring relationships formed by medical students and faculty. Med Educ Online, 2012:1(17):242.
30. Yehia BR, Cronholm PF, Wilson N, Palmer SC, Sisson SD, Guilliames CE, et al. Mentorship and pursuit of academic medicine careers: a mixed methods study of residents from diverse backgrounds. BMC Med Educ. 2014;14:26.
31. Eckleberry-Hunt J, Tucciarone J. The challenges and opportunities of teaching "Generation Y". J Grad Med Educ. 2011;3:458-61.
32. Shangraw RE, Whitten CW. Managing intergenerational differences in academic anesthesiology. Curr Opin Anaesthesiol. 2007;20:558-63.
33. Eva KW, Regehr G. Self-assessment in the health professions: A reformulation and research agenda. Acad Med. 2005;80(10 Suppl):S46-S54.
34. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. JAMA. 2006;296:1094-102.

Text Size Affects Eye Movement during Reading among Young Adults and Adults with Presbyopia

Noor Halilah Buari, M.H.Sc (Optom)*, Siti Nurnazihah Hamka, B.Optom*, Anis Nur Fazlyana Md-Isa, B.Optom*, Shauqiah Jufri, M.HSc (Optom)**

**Centre of Optometry Studies, Faculty of Health Sciences, Universiti Teknologi MARA Selangor, Kampus Puncak Alam, 42300 Bandar Puncak Alam, Selangor, Malaysia, **Topvision Eye Specialist Center, 55, Jalan Peninsula Utama 1, Peninsula Park, Pasir Penambang, 45000 Kuala Selangor, Selangor, Malaysia.*

ABSTRACT

Objective: Reading is an activity that indirectly informs a person's visual capacity to distinguish letters and words. Reading begins with eye movements, then substantial cognitive processing and synthesis, before becoming voice reading. Therefore, text is a factor that could impact reading quality through its control of eye movements. This study examined the eye movements of young adults and adults with presbyopia using texts of different sizes.

Materials and Methods: Twenty-five young adults and twenty-two adults with presbyopia and good vision were included in this study. Six text sizes of a passage were chosen as the reading stimuli. The eye movement of participants in saccades and fixation were captured, tracked, and analyzed using the Dikablis eye tracker glasses.

Results: Eye movement of young adults differed significantly ($p<0.05$) when reading texts of different sizes. The eyes moved more and had a wider saccadic angle as the font size increased. An increase in fixations or stopping of the eyes were observed with larger texts. Adults with presbyopia had significantly different eye movement patterns than young adults ($p<0.05$), whereby these participants stopped more frequently at longer periods and had a narrower saccadic angle.

Conclusion: Eye movements changed when reading texts of varied sizes and the movements differed between younger and older adults. These translate to altered visual searching and attention strategies with varied text readability, indicating that the oculomotor system adapts to the pattern, shape, and size of the presented reading material. This behavior could imply that cognitive processes have been altered to facilitate comprehension.

Keywords: Reading; fixations; saccadic; presbyopia; readers (Siriraj Med J 2022; 74: 650-657)

INTRODUCTION

Reading is more than just looking at a bunch of words, rather a process of transferring and processing knowledge from a presented reading material. Reading is a process of deciphering the meaning of written symbols and letters. To read, a person must be able to (i) recognize the words they see (word recognition), (ii) comprehend what the words mean (comprehension), and (iii) connect word and meaning so that reading becomes automatic and accurate (fluency).[1,2]

Reading materials come in a variety of sizes and formats. Books are the most common type of reading material. Newspapers, magazines, food labels, brochures, emails, reports, and a variety of other available reading materials that are used by children and adults in their daily reading activities. Given the wide range of reading

Corresponding author: Noor Halilah Buari
E-mail: noorhalilah@uitm.edu.my
Received 7 April 2022 Revised 16 August 2022 Accepted 16 August 2022
ORCID ID:http://orcid.org/0000-0002-7058-1176
http://dx.doi.org/10.33192/Smj.2022.76

materials available, they are presented in different ways. Reading materials come in varying font styles, text sizes, color, and layout, which may influence the outcome of reading. Font size, in particular, has been significantly shown to impact reading performance.[3-5] Readers with normal vision increased their reading speed as the font size increased.[5] while individuals with a blurred vision also have a similar affect in reading speed.[3]

The size of text also possibly influences eye movement during reading.[2,6-8] Eye movement is a microprocess of reading that causes changes in eye movement parameters such as fixations and saccades. Increased text size has resulted in altered saccadic eye movements.[7,8] Schoolchildren who read 25pt and 30pt sized texts, had increased amplitude of saccade movement.[8] Additionally, as the text size increased, the number of fixations increased, and the duration of fixations decreased.

When text was displayed on an LCD screen, the visual angle measured by saccade length, changed as the text size increased.[7] Here, saccadic eye movement remained consistent at smaller print sizes of up to 32pt text size. When eight different text sizes were compared, readers had the most prolonged fixation duration with a text size of 20pt.[7] Beymer *et al.*[6] studied online reading and discovered that participants read faster with larger fonts. When reading an online text, smaller text (10pt) retained the reader's attention for a more extended time than larger text (12pt).[6] Fixation also occurs more frequently with larger print sizes (32 pixels) than with smaller print sizes (24 pixels).[9]

Normal aging in late adulthood was likely associated with ocular and neural changes resulting in subtle deterioration of visual functions.[10,11] Changing optical features that occurs in older adults significantly reduces sensitivity towards contrast. Due to pupillary miosis, illuminance of the retina in aged eyes is diminished. The aging eye also has increased intraocular light dispersion and optical aberrations, impairing the contrasting appearance of an image.[12] These physiological changes indirectly affect vision performance, such as reading. Characters that are either too small or too huge impede reading speed in the aged population.[13] As a result, adults with reduced contrast sensitivity in low and high spatial frequency have difficulty reading and thus, additional lens with increased dioptric power are required to view standard size prints.[14] Healthy older adults with decreased visual processing speed on the other hand, have deficits in visual attention, associative learning, and executive function.[10,15,16] Thus, it was established that a generalized slowdown of information processing is likely caused by many aging-related cognitive impairments.

There is also a substantial change of the reading process in different age groups.[13,17-19] In preschoolers and elementary school children, reading speed improved with age.[19] Reading speed was fairly constant between teenagers and young adults, indicating that the reading skills had fully developed.[17,18] As people aged, their reading speed slowed. Adults read at approximately 9% slower than teenagers.[17,18] Slower reading could be due to some vision deficits in the healthy but aging eyes. After the age of 45, significant vision losses occur, and most notably in the middle and high spatial frequencies.[17]

Studies had established that changes in the size of texts caused eye movements to change as well. To investigate the cognitive processes of reading, the movement of the eyes were extensively investigated. However, studies in the past were conducted with a limited range of font size or reading materials were from electronic devices, such as online reading or reading from a digital screen (tablets, smartphones and laptops). Moreover, discrepancies exist regarding eye movements and there is limited research which compared reading using various font size on printed materials or among varied age groups. Data on these parameters are key as reading also occurs with printed or hard copy reading materials such as books, newspapers, food labels, brochures and medication labels on a daily basis, apart from reading through electronic displays. Given the age-related variation in reading ability, it is worthwhile to investigate eye movement patterns among young adults and adults with presbyopia. In this current study, the effect of text size and age on reading eye movement was examined between these two groups of readers.

MATERIALS AND METHODS

Participants

The eye movement behavior in reading different text sizes was conducted using a cross-sectional experiment of young adults (mean age: 22.28±1.46 years) and adults with presbyopia (mean age: age 49±6.65 years). The sample size was estimated using the formula $n=(Z\sigma/\Delta)^2$ computed using the G-power sample size software. The precision (σ) was set at 1. The standard deviation (Δ) was 2.3, and the Z for the 95 percent confidence interval was 1.96. As a result, each group needed a total of 25 participants.

This study included twenty-five young adults and twenty-two adults with presbyopia. Corruption of data led to, three (3) participants with presbyopia to be eliminated. The inclusion criteria were short-sightedness and long-sightedness with a low to moderate habitual refractive error (sphere correction between +2.00DS and -3.00DS;

astigmatism up to -2.00DC; and addition for near up to +3.00DS). The best-corrected distance and near visual acuity were 6/9 and N6 or better. Participants with any form of binocular vision impairments and ocular illnesses and eye diseases were excluded.

Reading stimuli

Reading passages in the Malay language from the Shauqiah-Ai Hong-Halilah Reading Passage Compendium (SAHRPC) were used as a reading stimuli.[20] The SAHRPC consisted of 42 passages divided into three sets of 13 each. Each set had 13 different Arial font text sizes, ranging from 1.2logMAR to 0.0logMAR. The passages were organized on a 0.1-step logarithmic scale in decreasing order. Each set of eight-page flip designs contained 13 printed passages in landscape format on a matte type A3 sized paper. A paragraph of four to five lines consisting of 50 words were arranged in each passage. Passages were taken from the Malaysian Ministry of Education's Malay language textbook, which is used in primary schools. This study used six Malay excerpts from the SAHRPC. The text sizes were 5pt, 8pt, 14pt, 20pt, 30pt, and 52pt, which were equivalent to 0.2logMAR, 0.4logMAR, 0.6logMAR, 0.8logMAR, 1.0logMAR, and 1.2logMAR text legibility, respectively. The six text sizes were chosen to resemble actual reading materials, with the text size of 52pt representing conventional newspaper headlines. Text of 30pt is similar to newspaper sub-headlines and children's books, 20pt is the same size as books for children aged seven to eight years old, 14pt is the equivalent to grade one to three children's textbooks, 8pt is the same size as tiny column newsprints, and 5pt is comparable to the small print found in The Bible or footnotes.

Eye Tracker

The Dikablis Eye Tracking Glasses Professional (Ergoneer, Germany) is a wireless, head-based eye-tracking device that fits as a glasses-like frame. It includes two small cameras in front of the eyes that record and detect eye movement. The camera captures a 384 × 288-pixel black-and-white video of the eye. The scene camera, positioned on the nose portion of the spectacles, is another small camera. The scene camera records the participant's viewing scene at a resolution of 1,920 × 1,080 pixels during data recording. The scene camera can be tilted up to 12 degrees upward and 81 degrees downward. The eye cameras may be adjusted to meet the wearer's interpupillary distance. The head unit can also be easily adjusted to fit the size of the head using the cord stopper on the drawstring. The head unit can be worn over polarized glasses or spectacles. The D-Lab software version 3.0, installed on a computer, was used to send and analyze the collected data. The Dikablis Eye Tracking Glasses Professional has various data visualization tools, including an eye tracker, heat maps, and a glance flow diagram.

Experimental procedures

The experiment was conducted in a room with ambient lighting and installed with a light booth and the reading stimuli. All participants provided a written informed consent before the experiment was commenced. The study was approved by the university's research ethics committee, in accordance with the Helsinki Declaration (Approval No: REC/426/17). The participants sat facing the reading stimuli (The Malay passage) placed on a reading stand at a working distance of 40 cm after passing the screening procedures. The passages were covered to prevent the participant from reading ahead of time. Along with their spectacle (if any), the participant wore the Dikablis eye tracker. The eye tracker was calibrated and the participant was asked to read aloud at a regular reading speed. The Dikablis eye tracker recorded the eye movement simultaneously during the reading process. The experiment was carried out individually for each text size.

Statistical analysis

Eye movement data consisting of frequency of saccades, saccade angles, number of fixations, and fixation durations were recorded and sent to the D-Lab program. The SPSS version 21.0 software was used for statistical analysis. The median, quartile, mean, and standard deviation were derived as descriptive data from the eye movements. The Kolmogorov-Smirnov test showed that young adults' eye movement data were normally distributed ($p>0.05$). Hence, eye movements with different text sizes among the participants were compared using the One-way ANOVA test. Scheffe post-hoc test was selected for further analysis. A comparison of eye movement behavior between young adults and adults with presbyopia was conducted using an independent t-test. The statistical significance was set at $p<0.05$.

RESULTS

This study compares the eye movement among young adults for texts in different font sizes. The number of saccades was found to be significantly increased ($F(5, 144) = 96.6$, $P<0.001$) when reading different text sizes. More saccadic movements were made when the text size increased, starting from 0.8logMAR to 1.2logMAR. Fig 1

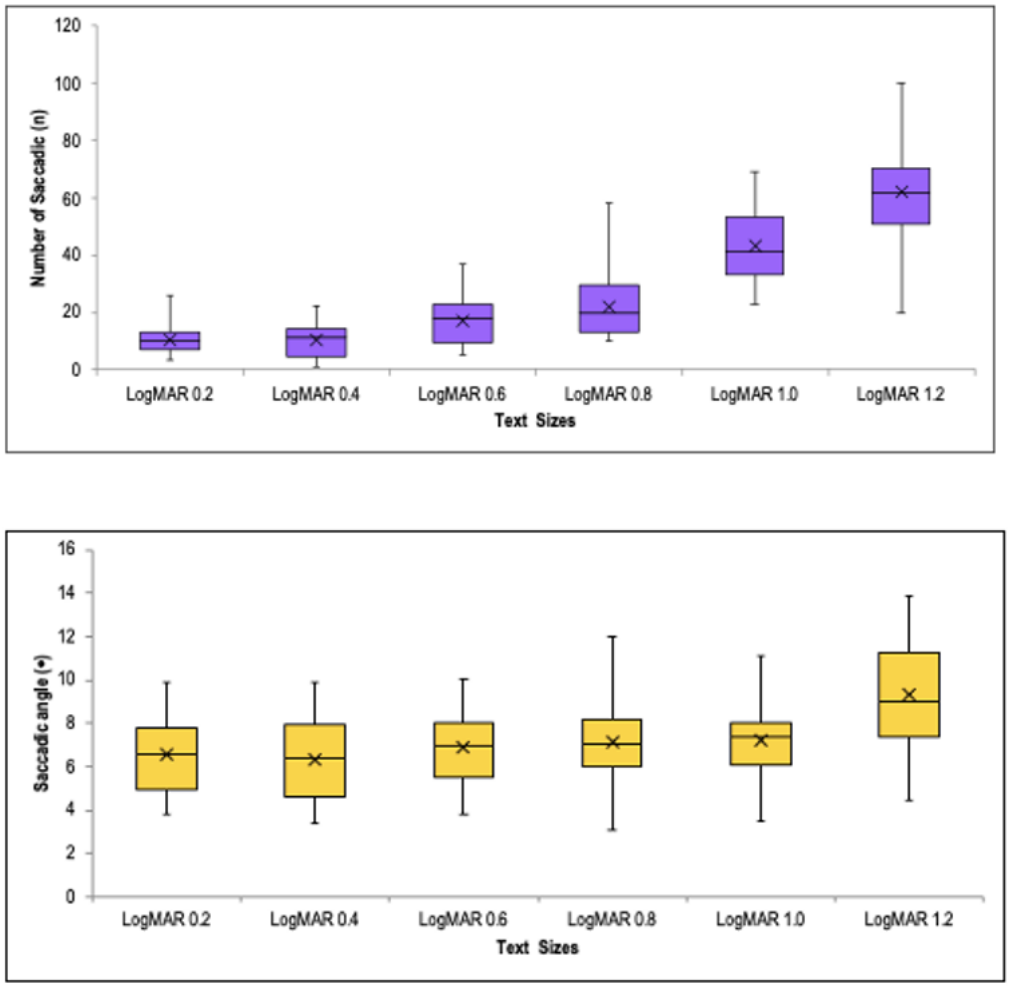


Fig 1. The saccadic eye movement behavior for different text sizes is shown in box plots. (a) Number of saccadic and (b) saccadic angles are shown above. The median is the horizontal line at the center of the box, and 'X' represents the mean value.

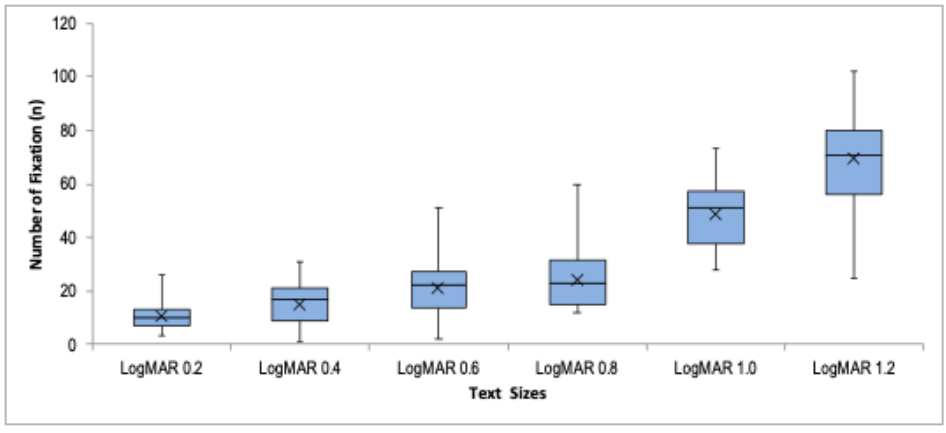
depicts the trend of the number saccadic movement. The saccadic angle became wider and more substantial as the text size increased, F (5, 144) = 7.03, p<0.001. When young adults read 1.2logMAR text size, which was comparable to 52pt, there were substantial changes in the saccadic angle data.

Fig 2 shows the frequency of fixations and fixation time. When font size increased, young adults made more stops during reading, F (5,144) =95.6, p<0.001. The number of fixations gradually rose with increasing text size, while fixation changed at text sizes 1.0logMAR and 1.2logMAR in young adults. The time the eye takes to stop during reading is called fixation duration. Fig 2 shows the fixation duration for different text sizes. Fixation duration was significantly longer for smaller text sizes, F (5,144)=7.04, p<0.001. At 0.2logMAR, 0.4logMAR, and 0.8logMAR, the eyes made more prolonged duration of stopping during reading.

Fig 3 shows the comparison of the number of saccadic movements and saccadic angles between adults with presbyopia and young adults. It was found that the number of saccadic movements were significantly different for different text sizes at 0.2logMAR (t(45)=-2.43, p=0.019), 0.6logMAR (t(45)=-2.77, p=0.008) and 0.8logMAR (t(45)=-2.32, p=0.025). Presbyopia readers made more saccadic movements compared to young adults.

The saccadic angle was significantly different at 0.2logMAR (t(45)=2.5, p=0.016) and 0.6logMAR (t(45)=2.5, p=0.015) between both groups of adults. Adults with presbyopia produced smaller and narrow saccadic angles when they moved from one fixation to another, especially with smaller text sizes.

Fig 4 shows the comparison of the number of fixation and fixation duration between young adults and adults with presbyopia. Adults with presbyopia made significantly (p<0.05) more pauses during a reading at all text sizes than young adults except at 1.2logMAR. Both adults with presbyopia and young adults had more fixation at larger text size than smaller text size. The



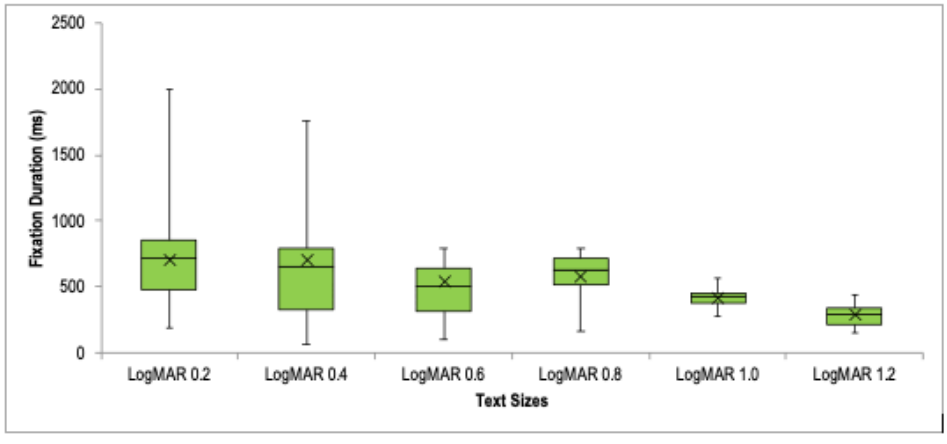
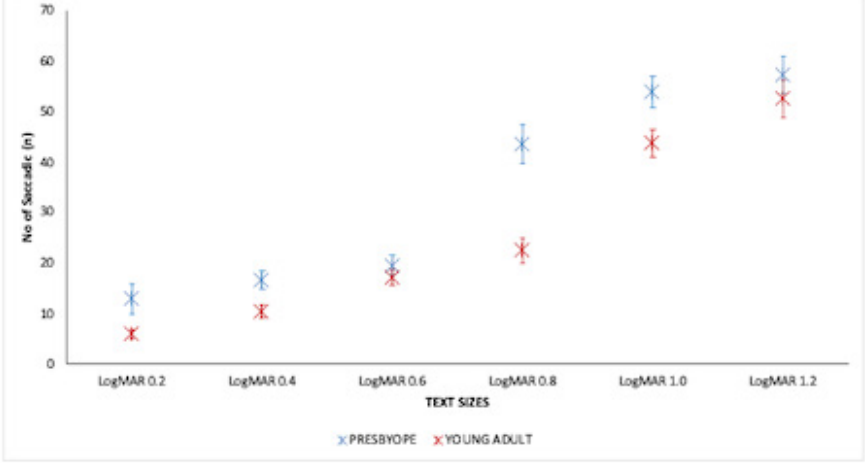


Fig 2. The fixation eye movement behavior for different text sizes is charted in box plots. The median value represents a middle horizontal line in the box, and the mean value is represented by the 'X' symbol within the box.



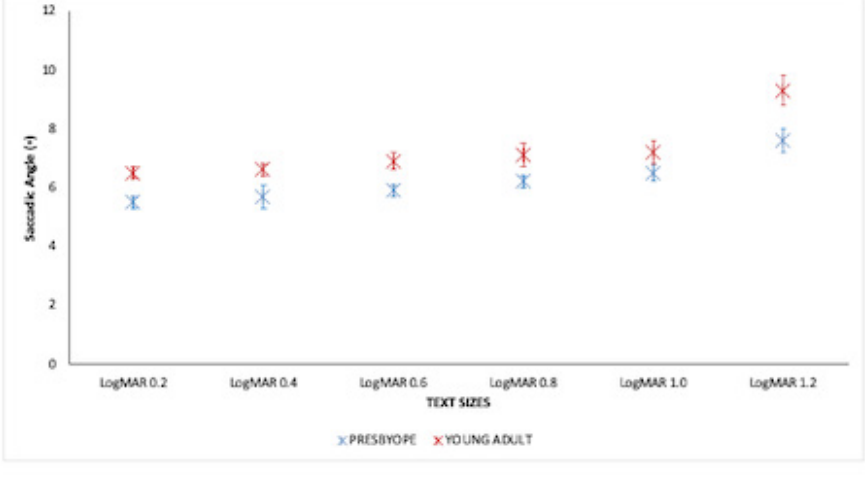


Fig 3. Comparison of saccadic eye movement for different text sizes between adults with presbyopia and young adults. (a) Number of saccadic and (b) saccadic angles is shown above.

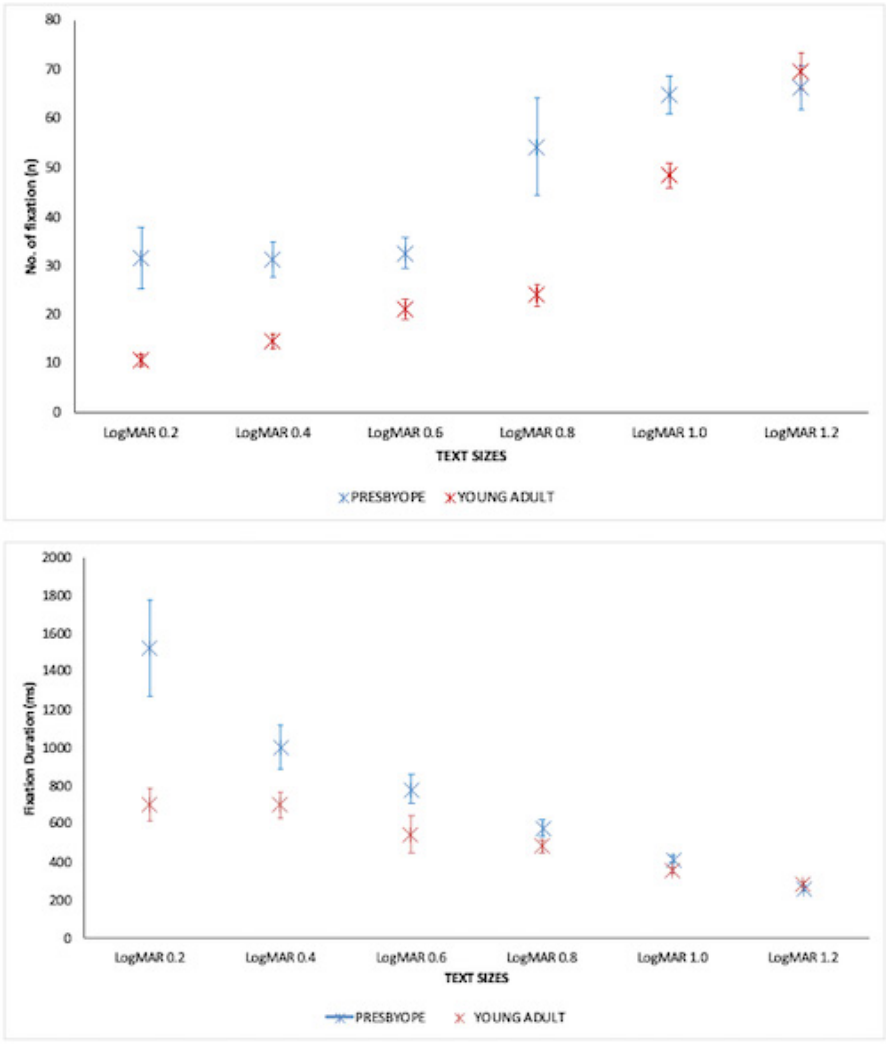


Fig 4. Comparison of fixation eye movement for different text sizes between adults with presbyopia and young adults. (a) Number of fixation and (b) fixation duration is shown above.

current data also demonstrates that fixation duration in adults with presbyopia was significantly slower at 0.2logMAR (mean:1523±254ms, t(45)=-3.83, p<0.001) and at 0.4logMAR (mean:1003±115ms, t(45)=-2.23, p=0.025) compared to the young adults readers. Fixation duration was longer at smaller text sizes than larger text sizes.

DISCUSSION

Two elements of eye movement behaviours investigated in this study were saccadic movement and fixation. Reading at a size of 20pt displayed a significant change in eye movement behaviour. The findings revealed that saccadic angle became broader as the text size increased, which was also demonstrated in a previous study.[7,8] The amplitude of the saccade movement was also increased in 7- to 12-year-old children as they read 25pt and 30pt texts.[8]

Investigation of the shape and size of text on an LCD screen discovered that saccade lengths in degrees of visual angle widened with large texts.[7] The findings of Franken *et al.*[7] mirrored the current study that also showed saccade movement changed at a larger text. Other studies however, found no significant differences in saccadic movement, both length and the number of saccades between different print sizes.[6,9] Considering the major difference, Baymer *et al.*[6] used reading stimuli with texts with the font sizes of 10pt, 12pt, and 14pt, while this study used a wider range of size. In the current study, saccades and saccadic lengths were similar at smaller text size, but, increased with increasing text size due to more jumping eye movements between the lines of the whole passage, which were laid out to fit the page layout. The participant's shifting visual searching strategies could account for the different outcome gained in this study compared to the literature from the past.

As text size increased, the number of fixations increased and fixation duration reduced. For example, a 30-point text size had a significant increase in fixation, whereby, smaller text size led to less frequent stops but more time processing visual information. Studies have revealed a similar pattern whereby the eye made fewer

and longer fixations with smaller text sizes.[6-9] Decreasing text size (10-pt or 24 pixels) resulted in significantly longer fixation durations.[6-9] At a text size of 2-pt, the fixation duration changed substantially.[7]

Atypical fixation behaviour was also revealed with reading materials of various text sizes. For smaller texts, the eye may spend more time fixating due to the crowding effect,[8,21] as longer stops were necessary for the eye to focus and extract visual information. Another possibility for longer stops is that letters or words printed closer to each other may interfere with the visual extraction process. Reading texts presented with larger spaces between the words did not seem to promote enhanced concentration.[22] However, readers found it difficult to read the smaller texts, which could explain how fixation behaviour diminishes with smaller texts. Changes in fixation behaviour also suggested that the eye employs visual attentional strategies to adapt to changes in different reading stimuli.

Compared to young adults, adults with presbyopia made more saccadic movements across different text sizes when eye movement data was analyzed. In terms of the size of the saccadic movement, adults with presbyopia made smaller and narrower angles between one fixation to another, especially with smaller text size. Furthermore, as the age increased, more stopping or pauses were made during reading for all text sizes. Fixations were also longer with smaller texts among adults with presbyopia compared to young adults.

Apart from a few other plausible explanations, this finding relates to reading performance, which has been shown to gradually decrease with age, even among individuals with good vision.[13,23] Contrast sensitivity have been demonstrated to decrease with age.[12,24-26] Physiological changes in the crystalline lens causes increment in light dispersion and aberration of sensitivity that reduces contrast.[12] Reduced contrast sensitivity also influenced visual-based cognitive abilities and visual speed processing.[12,26] Longer latencies in low contrast text affect information processing speed.[24] Previous research found that older adults have slower cognitive visual processing than younger adults.[10,27]

Decreased speed in reading could also be caused by deterioration in the transient system in an older adult.[10,27] The reduced spatial frequency may directly affect reading speed and change eye movements. Cognitive deficits in the elderly, such as visual attention and associative learning have been linked to changes in oculomotor function.[16,26,27] Age-related decline in motor function and ocular disease also affect eye movement.[27-30] Furthermore, older adults make more pronounced eye movement transitions compared to younger adults.[29] Young adults also had the quickest saccadic reaction compared to the aged participants. These findings demonstrate extremely strong age-related effects, corresponding to the various stages of normal nervous system development and degeneration.[31]

CONCLUSION

In conclusion, eye movement changes when young adults and presbyopia readers read passages with varying text sizes. Changes in saccade eye movements proposed that visual searching strategy had shifted and changes in fixation behavior suggested a shift in visual attentional strategy. Age also affects eye movement behavior during reading, as adults with presbyopia demonstrated significantly different saccadic eye movements and fixation compared to young adults. These changes, however, reveal that the eye movement behavior is adaptable to changes in the pattern, shape, and size of the reading materials. Trend of these eye movement parameters may provide some insight that cognitive and brain processing during reading are altered to facilitate understanding a variety of reading materials.

ACKNOWLEDGMENTS

The authors would like to thank all participants who participated in this study.

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

REFERENCES

1. Hung GK, Ciuffreda K. Models of the Visual System. New York: Kluwer Academic/Plenum Publishers; 2002.781p.
2. Rayner K, Slattery TJ, Bélanger NN. Eye movements, the perceptual span, and reading speed. Psychon Bull Rev. 2010; 17(6):834-9.
3. Alotaibi AZ. The effect of font size and type on reading performance with Arabic words in normally sighted and simulated cataract subjects. Clin Exp Optom. 2007;90(3):203-6.
4. Arditi A, Cho J. Letter case and text legibility in normal and low vision. Vision Res. 2007;47(19):2499-505.
5. Legge GE, Pelli DG, Rubin GS, Schleske MM. Psychophysics of reading—I. Normal vision. Vision Res. 1985;25(2):239-52.
6. Beymer D, Russell D, Orton P. An Eye Tracking Study of How Font Size and Type Influence Online Reading. In: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction [Internet]. 2008 [cited 2021 Jun 14]. p. 15–8. Available from: https://scienceopen.com/document?vid=6acc1c35-8d24-4d60-85b3-31e6385bc280
7. Franken G, Podlesek A, Možina K. Eye-tracking Study of Reading Speed from LCD Displays: Influence of Type Style and Type Size. JEMR. 2015;8(1):1-8.

8. Masulli F, Galluccio M, Gerard CL, Peyre H, Rovetta S, Bucci MP. Effect of different font sizes and of spaces between words on eye movement performance: An eye tracker study in dyslexic and non-dyslexic children. Vision Res. 2018;153:24-29.
9. Yen NS, Tsai JL, Chen PL, Lin HY, Chen ALP. Effects of typographic variables on eye-movement measures in reading Chinese from a screen. Behaviour & Information Technology. 2011;30(6):797-808.
10. Owsley C. Aging and vision. Vision Res. 2011;51(13):1610-22.
11. Owsley C. Vision and Aging. Ann Rev Vis Sci. 2016;2(1):255-71.
12. Artal P, Guirao A, Berrio E, Piers P, Norrby S. Optical Aberrations and the Aging Eye. Int Ophthalmol Clin. 2003;43(2):63-77.
13. Akutsu H, Legge GE, Ross JA, Schuebel KJ. Psychophysics of Reading--X. Effects of Age-related Changes in Vision. J Gerontol. 1991;46(6):P325-31.
14. Gupta N, Wolffsohn JS, Naroo SA. Comparison of near visual acuity and reading metrics in presbyopia correction. J Cataract Refract Surg. 2009;35(8):1401-9.
15. Rubin GS, Ng ESW, Bandeen-Roche K, Keyl PM, Freeman EE, West SK, et al. A Prospective, Population-Based Study of the Role of Visual Impairment in Motor Vehicle Crashes among Older Drivers: The SEE Study. Invest Ophthalmol Vis Sci. 2007;48(4):1483-91.
16. Salthouse TA. Relations between Cognitive Abilities and Measures of Executive Functioning. Neuropsychology. 2005;19(4):532-45.
17. Calabrèse A, Owsley C, McGwin G, Legge GE. Development of a Reading Accessibility Index Using the MNREAD Acuity Chart. JAMA Ophthalmol. 2016;134(4):398-405.
18. Chen AH, Khalid NM, Buari NH. Age factor affects reading acuity and reading speed in attaining text information. Int J Ophthalmol. 2019;12(7):1170-6.
19. Virgili G, Cordaro C, Bigoni A, Crovato S, Cecchini P, Menchini U. Reading Acuity in Children: Evaluation and Reliability Using MNREAD Charts. Invest Ophthalmol Vis Sci. 2004;45(9):3349-54.
20. Chen AH, Buari NH, Shauqiah J. The Development of SAH Reading Passage Compendium: A Tool for the Assessment of Reading Performance Related to Visual Function. IES. 2017;10(12):30-37.
21. Bricolo E, Salvi C, Martelli M, Arduino LS, Daini R. The effects of crowding on eye movement patterns in reading. Acta Psychologica. 2015;160:23-34.
22. Sharmin S, Špakov O, Räihä KJ. The Effect of Different Text Presentation Formats on Eye Movement Metrics in Reading. JEMR. 2012;5(3):1-9.
23. Beier S, Oderkerk CAT. The effect of age and font on reading ability. Visible Language. 2019;53(3):50-58.
24. Anstey KJ, Butterworth P, Borzycki M, Andrews S. Between- and Within-Individual Effects of Visual Contrast Sensitivity on Perceptual Matching, Processing Speed, and Associative Memory in Older Adult. Gerentology. 2006;52:124-30.
25. Hohberger B, Laemmer R, Adler W, Juenemann AGM, Horn FK. Measuring contrast sensitivity in normal subjects with OPTEC® 6500: influence of age and glare. Graefes Arch Clin Exp Ophthalmol. 2007;245(12):1805-14.
26. Skeel RL, Schutte C, Voorst W van, Nagra A. The Relationship Between Visual Contrast Sensitivity and Neuropsychological Performance in a Healthy Elderly Sample. J Clin Exp Neuropsychol. 2006;28(5):696-705.
27. Chan CYH, Chan AB, Lee TMC, Hsiao JH. Eye-movement patterns in face recognition are associated with cognitive decline in older adults. Psychon Bull Rev. 2018;25(6):2200-7.
28. Darchia N, Campbell IG, Feinberg I. Rapid Eye Movement Density is Reduced in the Normal Elderly. Sleep. 2003;26(8):973-7.
29. Firestone A, Turk-Browne NB, Ryan JD. Age-related deficits in face recognition are related to underlying changes in scanning behavior. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2007;14(6):594-607.
30. Chen AH, Jufri S, Congdon N. The Impact of Glaucomatous Visual Field Defects on Speed and Eye Movements during Reading: Siriraj Med J. 2021;73(1):17-25.
31. Munoz DP, Broughton JR, Goldring JE, Armstrong IT. Age-related performance of human subjects on saccadic eye movement tasks. Exp Brain Res. 1998;121(4):391-400.

Predicting the Need for Continuation of N-acetylcysteine Treatment among Acute Paracetamol Overdose Patients with Psi Parameter

Pattaraporn Mekavuthikul, M.D.*, Sunsern Cheamanunkul, Ph.D., Pinsumon Chomchai,***, Jariya Phuditshinnapatra, M.D.***

**Department of Preventive and Social Medicine, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, **Division of Science, Mahidol University International College, Mahidol University, Nakhon Pathom, Thailand, ***Shrewsbury International School Bangkok Riverside, Thailand.*

ABSTRACT

Objective: AcetaCalc was used to evaluate Psi's accuracy in predicting cases that required prolonged N-acetylcysteine (NAC) therapy, as well as Psi's optimal cut-off.

Materials and Methods: This is a retrospective study of patients with acute paracetamol overdose who were treated with NAC at Siriraj Hospital between 2007 and 2016. The Psi parameter was calculated using the Acetacalc after entering paracetamol concentrations, blood sampling times, and NAC onset times. Indications for NAC continuation is in accordance with the guidelines, which recommended that NAC treatment be continued if the follow-up aminotransferase reached 50 U/L or higher.

Results: We enrolled 403 patients, the proportion of NAC prolongation was 50.4 %. Psi was shown to be a significant predictor of NAC prolongation ($p < 0.001$) with area under the receiver operational characteristics curve 0.766 (95% confidence interval (CI) 0.719-0.813). The Psi cutoff with highest Youden index was 1.757 mM-hour. The sensitivities and specificities of the cutoff were 0.517 (95% CI 0.449-0.585) and 0.940 (95% CI 0.898-0.965), respectively.

Conclusion: Psi parameter calculated through AcetaCalc is a useful tool for the prediction of cases where extension of NAC therapy beyond the standard regimen is indicated.

Keywords: Paracetamol; acute liver injury; N-acetylcysteine; prognosis; psi (Siriraj Med J 2022; 74: 658-665)

INTRODUCTION

Hepatotoxicity from paracetamol overdose remains a significant healthcare burden in Thailand and worldwide.[1] The most definitive treatment for the overdose is the timely and sufficient use of N-acetylcysteine (NAC), administered as a 300 mg/kg intravenous infusion over 21 hours.[1] When NAC therapy is started early, usually within 8 hours after the overdose, hepatotoxicity can effectively be minimized. Factors that can contribute to the development of hepatotoxicity secondary to paracetamol overdose are a high initial serum paracetamol concentration and a delay in the initiation of NAC treatment. Studies have shown that the risks of hepatotoxicity (defined by aminotransferase concentration ≥ 1,000 U/L) when NAC is initiated before 8 hours after ingestion can range from 3.5-7.7%. That risk increases progressively to 10.3-22.2 % when NAC is given at 10-16 hours post ingestion. Subsequently, it can be as high as 12.9-45.1% when NAC is given later than 16 hours after ingestion.[1,2]

Corresponding author: Jariya Phuditshinnapatra
E-mail: jariya.ohu@mahidol.ac.th
Received 11 July 2022 Revised 14 August 2022 Accepted 20 August 2022
ORCID ID:http://orcid.org/0000-0002-6474-5613
http://dx.doi.org/10.33192/Smj.2022.77

In addition, even in instances where paracetamol-induced hepatitis has already occurred, NAC therapy beyond the initial 21-hour period has been shown to reduce the severity of hepatotoxicity, as well as the rate of complications and mortality.[3,4] The indication for prolonging NAC therapy is an aminotransferase concentration of 50 U/L or higher.[5] It is recommended that the infusion be maintained until aspartate aminotransferase concentration has decreased to half of the peak level or lower.[6] On the other hand, the assessment of hepatotoxicity risks has many interacting clinical parameters such as the time of ingestion, the dose of ingestion, the time to treatment initiation with NAC, and the initial paracetamol level.[1,2] These parameters make up a risk profile which, with the right tool namely the Psi parameter, can be individualized and extended to predict the need for prolonged NAC treatment in specific patients.[1,7] The Psi (Greek letter Ψ) parameter is a composite calculation that accounts for both timed serum paracetamol concentration and the time after ingestion until N-acetylcysteine therapy is initiated. It was developed based on a toxicokinetic model as a quasi-trapezoidal area-under–the-curve of paracetamol concentration and the duration of hepatic glutathione deficiency.[8,9] (Fig 1) Its purpose is to individualize each patient's hepatotoxicity risk which can help clinicians determine the disease prognosis with reasonable accuracy. Overall, the Psi parameter reflects exposure to N-acetyl-p-benzoquinone imine (NAPQI), paracetamol's hepatotoxic metabolite, prior to starting NAC. Higher paracetamol concentration and longer delay in NAC treatment results in a higher calculated Psi parameter.[10,11] The details about mathematical derivation of Psi can be found in previously published works.[8,9] Consequently, the utility and accuracy of the Psi parameter in predicting individual risk of hepatotoxicity have been substantiated in various publications.[9-12] In the Thai population, high Psi (≥ 5.0 mM-hour) predicts hepatotoxicity with sensitivity of 96.9% (95% confidence interval (CI) 84.3-99.4) and specificity of 91.5% (95% CI 87.1-94.5).[10] However, despite such excellent clinical data, its complex calculations severely limit its usefulness in busy clinical settings.[9] In 2021, AcetaCalc, a web-based application developed jointly by the Faculty of Medicine Siriraj Hospital, Mahidol University and Mahidol University International College, has made this task much simpler. Users can simply input paracetamol concentration, time after ingestion when paracetamol level was obtained, and the time to NAC initiation and the application can calculate Psi parameter, as well as other predictors of hepatotoxicity. This application can be accessed at https://sunsern.github.io/aceta-calc/#/tabs/info. In the present study, we evaluate the use of Psi parameter, which is derived with AcetaCalc, as a predictor of the need for prolonging NAC treatment among patients with acute paracetamol overdose.

MATERIALS AND METHODS

This was a retrospective review of patients who presented at Siriraj Hospital, Bangkok, Thailand from January 1, 2007 to December 31, 2016 with paracetamol overdose. Inclusion criteria included age 12 years or older and treatment with N-acetylcysteine. Patients were excluded if they fit one of the following criteria: mixed ingestion, staggered ingestion (overdose process longer than 1 hour) and abnormal initial aminotransferase concentrations. A standard case record form was used to extract the information from the medical records, including age, gender, type and dose of the overdose, initial paracetamol concentration, blood chemistry results, treatment, follow-up blood chemistry results and clinical outcomes. The study protocol was approved by the Siriraj Human Research Protection Unit (MU-MOU CoA 472/2021).

The Psi parameter was calculated using the AcetaCalc Application. Input data for Psi calculation included time of blood sampling (hours after overdose), measured paracetamol

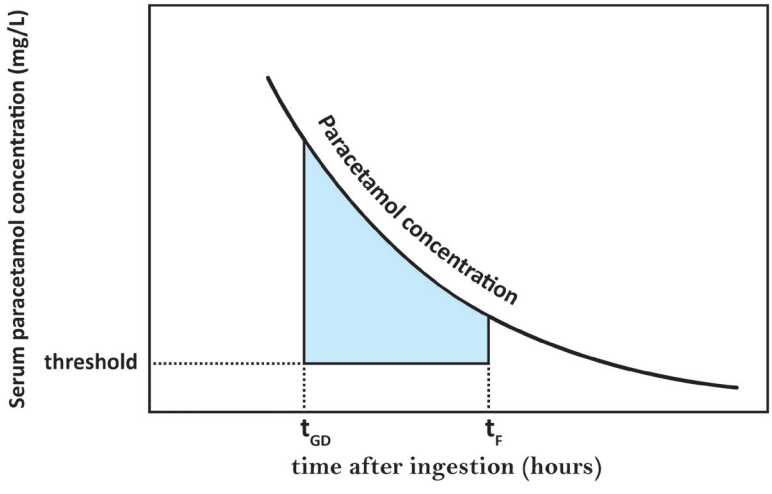


Fig 1. Calculation of Psi parameter
Shaded area represents Psi parameter, tGD = time of glutathione depletion (6 hours is used as a default), tF = time of N-acetylcysteine initiation, threshold = threshold paracetamol concentration (45 mg/L is used)

concentration (mg/L) and lag time from overdose to NAC initiation (hours). Patients fulfilled the primary outcome of the study if the follow-up aminotransferase concentration was 50 U/L or higher, indicating the need for prolonged NAC therapy. Hepatotoxicity was defined as an aminotransferase concentration of 1,000 U/L or above. For the purpose of comparison, extrapolated paracetamol concentration at 4 hours post ingestion ($[APAP]_{4hour}$) was calculated using the formula $[APAP]_{4hour} = C_t/2e^{>-(0.693/4)t}$ where C_t represents measured paracetamol concentrations and t indicated the time interval (hours) from ingestion to blood sampling.

During the study period, intravenous NAC was the mainstay treatment for acute paracetamol overdose. The standard regimen for NAC administration was 150 mg/kg in one hour, 50 mg/kg in four hours and 100 mg/kg in 16 hours, consecutively. Oral administration was used only when the patients had a contraindication to IV NAC or could not tolerate the intravenous regimen. NAC prolongation was carried out by the administration of 150 mg/kg/24 hours of NAC intravenously. In actual clinical settings, the decision to start and then discontinue the prolonged NAC therapy was based on each clinician's perception of whether a significant elevation and subsequent decline of aminotransferase had occurred.

Statistical analysis

Descriptive data were displayed as frequencies with percentages and means with 95% confidence intervals (CI). However, medians with interquartile ranges (IQR) were used for variables with non-normal distributions. Differences were tested with Student's t-test or Mann-Whitney U test. Proportions were tested with a chi-squared test or Fisher's exact test. Factors or co-variates with $p \geq 0.05$ were further tested with multiple logistic regression in order to predict cases that required NAC prolongation. Multiple logistic models were assessed using backward stepwise multiple logistic regression. Receiver operating characteristics (ROC) curve and area under the curve (AUC) with 95% CI were used to assess the accuracy of Psi in predicting the outcome. The optimal cutoff was selected using the highest Youden index.[13] Validities of the predictions were evaluated by sensitivity, specificity, positive likelihood ratio (LR+) and negative likelihood ratio (LR-) with 95% CIs. Statistical analyses were performed using PASW 18 (Release version18.0.0) statistical program.

The sample size estimation for this study was performed based on the "rule of 10 for logistic regression".[14] Since we expected no more than four factors or co-variates in the predicting model, the minimal number of events in this study was estimated to be 40. Therefore, we required at least 40 cases that fulfill the indication for NAC prolongation to achieve a desirable statistical power.

RESULTS

During the study period, 1,286 patients presented to Siriraj Hospital due to paracetamol overdose. Among these, 883 cases were excluded. Therefore, we enrolled 403 into the analyses. (Fig 2) The subjects consisted of 332 females (82.4%) with a median age of 23 years (IQR 2028, range 13-62). Compared with the group receiving standard NAC duration, the NAC prolongation group had significantly higher extrapolated four-hour paracetamol concentrations,

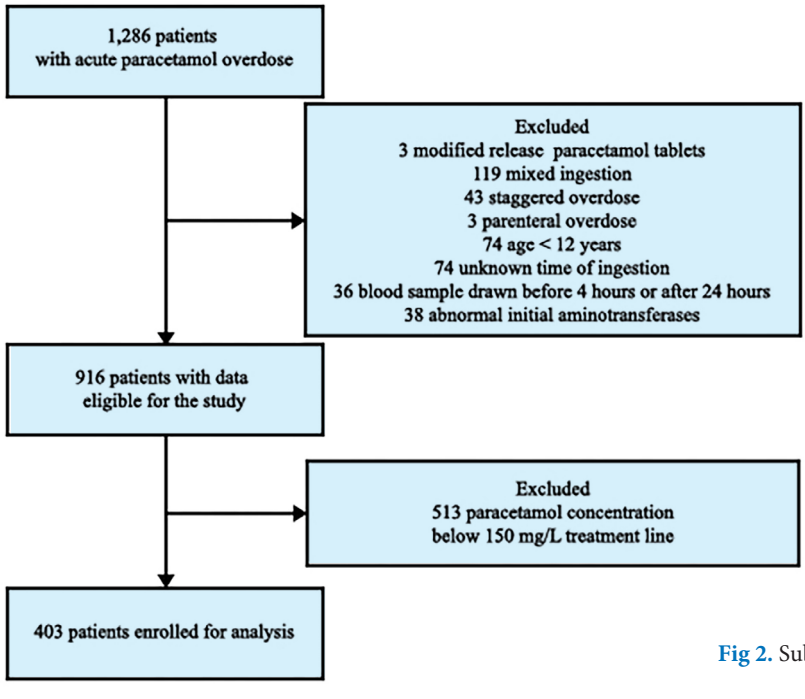


Fig 2. Subject enrollment flow

Psi values, peak aminotransferase concentrations and proportion of hepatotoxicity and longer time-to-NAC-therapy, while having a significantly lower proportion of patients who received decontamination with activated charcoal. (Table 1) None of the patients experienced liver failure and no mortality occurred in this study.

Both the Psi parameter and the decontamination with activated charcoal were entered into the multiple logistic regression analysis. However, activated charcoal yielded no statistical significance (*p*-value 0.401) in the multiple logistic regression model (Model 1). Activated charcoal was removed since the resultant Nagelkerke R^2 of Model 2 with Psi parameter alone was higher than Model 1 (a higher score meant better fit of the model). (Table 2) When using the equation Logit P = -0.918 + (0.824 * Psi), Psi was a significant predictor of the need to prolong NAC treatment (*p*-value <0.001). Fig 3 demonstrates the scatter plots of $[APAP]_{4hour}$ and onset of NAC therapy, as classified by cases with and without the need for NAC prolongation. There is a clear pattern of the need for NAC prolongation in cases with high $[APAP]_{4hour}$ or delayed NAC onset.

The ROC curve for Psi in predicting NAC prolongation has an AUC of 0.766 (95%CI 0.721-0.806) and is shown in Fig 4. The highest Youden index was found at the Psi concentration of 1.757 mM-hour. The 1.757 mM-hour cutoff yielded a sensitivity 51.7% and a specificity 94.0%. When the cutoff was increased to 2.948 mM-hour, the specificity achieved a maximal value of 99.5%. (Table 3)

TABLE 1. Demographic and clinical characteristics of the overall subjects, groups with and without N-acetylcysteine (NAC) prolongation.

Characteristics	Overall (403 case)	NAC prolongation (203 cases)	No NAC prolongation (200 cases)	P-value
Female (frequency (%)	332 (82.4)	167 (82.3)	165 (82.5)	0.951
Age (median (IQR) (years)	23 (20-28)	24 (21-27)	23 (19-31)	0.387
Paracetamol dose (mg/kg)	281.7 (200.0-400.0)	294.1 (202.0-408.2)	256.4 (200.0-378.8)	0.250
$[APAP]_{4hour}$ (mg/L)	237.3 (181.4-355.5)	356.0 (289.3-474.1)	195.5 (170.6-222.8)	<0.001
NAC onset (hours)	7 (5-9)	8 (6-10)	5 (4-7)	<0.001
Psi (mM-hour)	0.476 (0.001-2.250)	1.873 (0.001-4.995)	0.001 (0.001-0.639)	<0.001
Initial AST (U/L)	13 (10-18)	16 (11-31)	14 (10-30)	0.742
Initial ALT (U/L)	11 (8-17)	15 (8-31)	14 (9-30)	0.861
Peak AST (U/L)	37 (22-79)	76 (54-457)	24 (19-30)	<0.001
Peak ALT (U/L)	41 (19-86)	86 (60-572)	26 (15-37)	<0.001
Peak INR	1.1 (1.0-1.2)	1.1 (1.0-1.2)	1.1 (1.0-1.2)	0.048
Activated Charcoal	87 (21.6)	22 (10.8)	64 (32.5)	<0.001
Hepatotoxicity	45 (11.2)	45 (22.2)	0 (0)	<0.001

Abbreviations: NAC, N-acetylcysteine; IQR, interquartile range; $[APAP]_{4hour}$, Extrapolated 4-hour paracetamol concentration; AST, aspartate aminotransferase; ALT, alanine aminotransferase

TABLE 2. Logistic regression models of Psi and activated charcoal administration as predictors of N-acetylcysteine prolongation.

Models	Factors or co-variates	Regression coefficients	P-value	-2 log likelihood	Nagelkerke R^2
1	Psi	0.786	<0.001	423.678	0.285
	Activated charcoal	-0.253	0.401		
2	Psi	0.824	<0.001	424.394	0.378

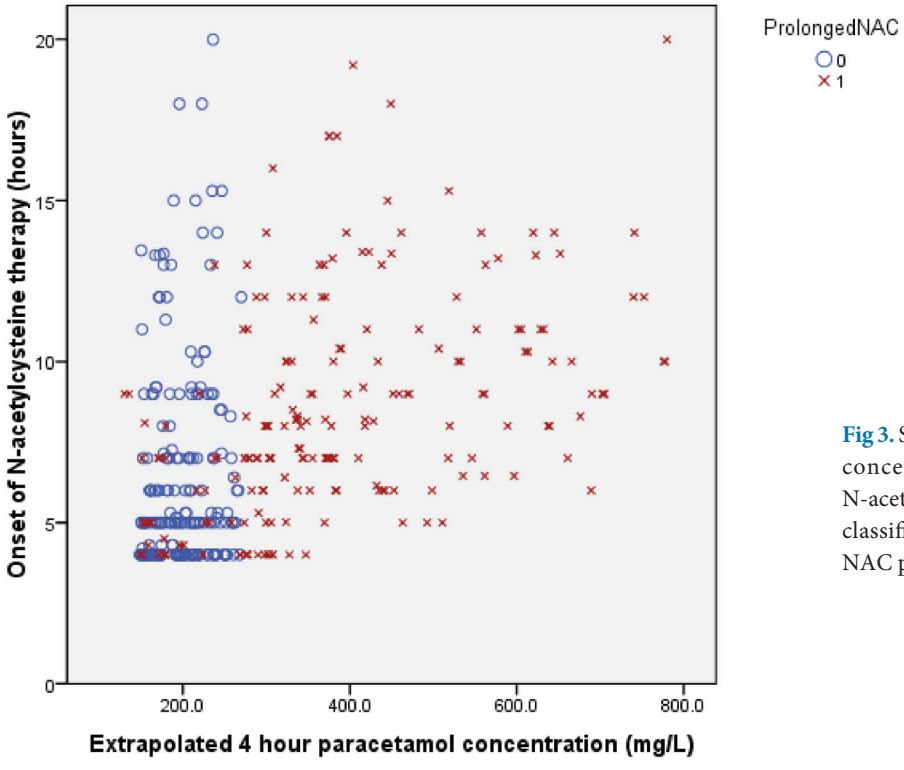


Fig 3. Scatter plots of extrapolated paracetamol concentration at 4 hours and onset of N-acetylcysteine (NAC) treatment (cases are classified as cases with (1) and without (0) NAC prolongation)

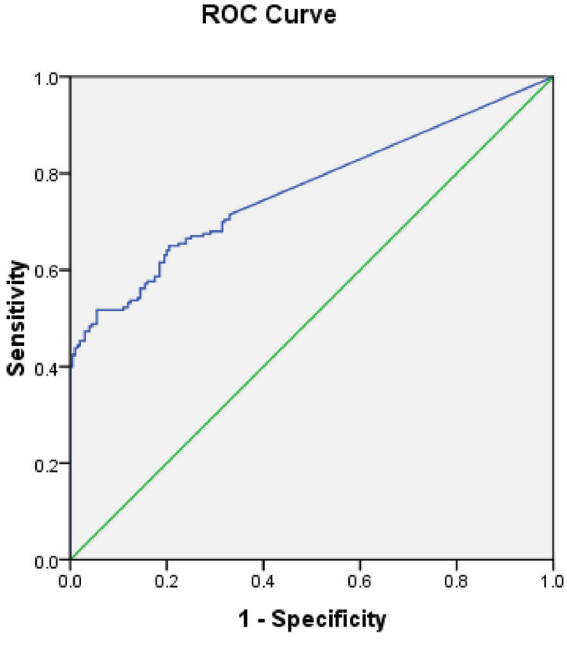


Fig 4. Receiver operating characteristics curve of Psi for predicting prolongation of N-acetylcysteine therapy

TABLE 3. Diagnostic validities of Psi for predicting N-acetylcysteine prolongation at various cutoff concentrations (* remarks cutoff psi concentration with highest Youden index).

Psi cutoff (mM-hour)	Sensitivity (95% CI)	Specificity (95% CI)	LR+ (95% CI)	LR- (95% CI)
2.948	39.9 (33.4-46.8)	99.5 (97.2-99.9)	79.80 (11.22-567.87)	0.60 (0.54-0.68)
1.757*	51.7 (44.9-58.5)	94.0 (89.8-96.5)	8.62 (4.90-15.16)	0.51 (0.44-0.60)
0.257	70.9 (64.3-76.7)	67.0 (60.2-73.1)	2.15 (1.73-2.67)	0.43 (0.34-0.55)
0.001	100.0 (98.1-100.0)	0.0 (0.0-0.019)	100.0 (100.0-100.0)	-

Abbreviations: CI, confidence interval; LR+, positive likelihood ratio; LR-, negative likelihood ratio

DISCUSSION

NAC therapy for paracetamol overdose is one of the most studied antidotal treatments, as apparent by the numerous guidelines for its administration. When given early, its efficacy in preventing hepatotoxicity is well-established. Currently, it is most often given in either a two-bag (200 mg in four hours and 100 mg in 16 hours) or a three-bag (150 mg/kg in one hour, 50 mg in 4 hours and 100 mg/kg in 16 hours) regimen over a period of 21 hours.[1,2] Serum paracetamol concentration and time-to-NAC therapy are two regularly assessed clinical parameters for clinicians when prognosticating the outcome of patients with paracetamol overdose.[2,8,10] The availability of these information makes the derivation of Psi parameter possible. Because the treatment regimens are so well adopted by clinicians, the associated demands for healthcare resources such as the frequency of laboratory monitoring efforts, the amount of antidote needed, as well as the patient's length of stay can often be reasonably predicted. However, in a subset of patients, paracetamol-induced hepatotoxicity can occur despite the completion of a standard NAC administration. In these cases, the continuation of NAC has been shown to significantly reduce mortality and complications. Postulated mechanisms of the action include NAC acting as an inflammatory modulator, increasing oxygen delivery and utilization and improving blood flow in the microvasculature.[1,3,4] According to the current guideline on the treatment of acute paracetamol poisoning, NAC continuation is recommended when the aminotransferase is elevated, as determined by AST or ALT of 50 U/L or higher.[1,5] Despite the seemingly low value, we believe that this cutoff offers maximal safety for patients with paracetamol induced hepatotoxicity. Therefore, we used this value to select cases for the primary outcome of this research. NAC continuation means intravenous infusion of NAC at the rate 150 mg/kg/24hours after completing the standard 21-hours NAC regimen.

In this study, the Psi parameter is shown to be an accurate predictor of NAC prolongation. The ROC's AUC of 0.766 implies that Psi has an acceptable accuracy in discriminating cases with and without the need for NAC prolongation. Although decontamination with activated charcoal is also significantly associated with the need for NAC prolongation, we postulate the mechanism to be the reduction of serum paracetamol due to activated charcoal's effects which, in turn, affects the calculation of the Psi value. Subsequently, when activated charcoal is eliminated from the logistic regression model, the Psi parameter remains as a sole and adequate predictor of the need for NAC prolongation. Our study illustrates the tendency for increasing need of prolonging NAC therapy as a function of higher paracetamol concentration and greater delay in NAC administration (Fig 3). Similarly, Cairney et al, reported a phenomenon whereby incidences of acute liver injury, as defined by aminotransferase > 150 U/L, gradually increased as a function of paracetamol nomogram groups. The incidence was 6% in the 0-100 mg/L nomogram group and progressed to as high as 27% in the > 500 mg/L nomogram group. The rates of acute

liver injury were lower in patients who were treated with NAC within 8 hours.[15] Of note, the term 'nomogram group' in this study is similar to the $[APAP]_{4hour}$ concentration groups in our study.

We proposed that the Psi cutoff of 1.757 mM-hour be used as the criterion to predict the need for NAC prolongation, since it yielded good specificity, although the sensitivity was mediocre. (Table 3) On the other hand, the 2.948 mM-hour cutoff had very high specificity, but the sensitivity became unacceptably low. The cutoff of 0.257 is shown in Table 3 because it is the second lowest concentration cutoff, next to 0.001 mM-hour. The cutoff value of 0.001 mM-hour is also significant because it has 100% sensitivity for the need for NAC prolongation, in concordance with previous findings that Psi of 0.001 mM-hour has a very high sensitivity for hepatotoxicity (aminotransferase > 1,000 U/L).[12]

The findings in our study have relevant clinical implications. Firstly, when Psi is calculated at the onset of treatment, physicians can expect to have to continue NAC therapy beyond the standard regimen if its value is 1.757 mM-hours or higher. This has significant ramification on the expected length of stay in the hospital and suggests that, in these cases, the reimbursement scheme may need to be adjusted. Secondly, when Psi is at its lowest possible value of 0.001 mM hour the probability of requiring NAC beyond the routine protocol should be very low since the value signifies early NAC treatment and low paracetamol concentration. Furthermore, the omission of a follow-up aminotransferase level after completion of the standard NAC therapy can also be justified based on such reasoning. Thirdly, our study shows that decontamination with activated charcoal can significantly reduce the need for NAC continuation beyond the routine protocol. This recapitulates the findings of previous studies and reiterates the importance of adequate gastric decontamination.[16-18]

The limitations of this study are in its retrospective nature. The data in medical records are intended for clinical services, and some errors in information obtained from the medical records may exist. The most important piece of information that can affect the result of this study is the time of paracetamol overdose, since it is a reference point from which the time of blood sampling and NAC initiation are calculated. In this study, a large number of subjects were excluded because they did not fulfill the requirement for Psi application. However, we do not expect them to result in any distortion of the results. For the future, we suggest that the study question is re-evaluated in a prospective observational study.

CONCLUSION

Psi parameter, a composite value of paracetamol concentration, time of blood sampling and onset of N-acetylcysteine treatment, is a useful tool to help clinicians predict the need for the continuation of NAC treatment beyond the standard regimen. The Psi parameter can be derived conveniently with the use of a computer application.

REFERENCES

1. Chiew AL, Buckley NA. Acetaminophen Poisoning. Crit Care Clin. 2021;37(3):543-61.
2. Rumack BH. Acetaminophen hepatotoxicity: the first 35 years. J Toxicol Clin Toxicol. 2002;40(1):3-20.
3. Harrison PM, Keays R, Bray GP, Alexander GJ, Williams R. Improved outcome of paracetamol-induced fulminant hepatic failure by late administration of acetylcysteine. Lancet. 1990; 335(8705):1572-3.
4. Keays R, Harrison PM, Wendon JA, Forbes A, Gove C, Alexander GJ, et al. Intravenous acetylcysteine in paracetamol induced fulminant hepatic failure: a prospective controlled trial. Bmj. 1991;303(6809):1026-9.
5. Chiew AL, Reith D, Pomerleau A, Wong A, Isoardi KZ, Soderstrom J, et al. Updated guidelines for the management of paracetamol poisoning in Australia and New Zealand. Med J Aust. 2020;212(4):175-83.
6. Curtis RM, Sivilotti ML. A descriptive analysis of aspartate and alanine aminotransferase rise and fall following acetaminophen overdose. Clin Toxicol (Phila). 2015;53(9):849-55.
7. Wong A, Graudins A. Risk prediction of hepatotoxicity in paracetamol poisoning. Clin Toxicol (Phila). 2017;55(8):879-92.
8. Chomchai S, Chomchai C, Anusornsuwan T. Acetaminophen psi parameter: a useful tool to quantify hepatotoxicity risk in acute acetaminophen overdose. Clin Toxicol (Phila). 2011;49(7): 664-7.
9. Sivilotti ML, Good AM, Yarema MC, Juurlink DN, Johnson DW. A new predictor of toxicity following acetaminophen overdose based on pretreatment exposure. Clin Toxicol (Phila). 2005; 43(4):229-34.
10. Chomchai S, Lawattanatrakul N, Chomchai C. Acetaminophen Psi Nomogram: a sensitive and specific clinical tool to predict hepatotoxicity secondary to acute acetaminophen overdose. J Med Assoc Thai. 2014;97(2):165-72.
11. Sivilotti ML, Yarema MC, Juurlink DN, Good AM, Johnson DW. A risk quantification instrument for acute acetaminophen overdose patients treated with N-acetylcysteine. Ann Emerg Med. 2005;46(3):263-71.
12. Chomchai S, Chomchai C. Predicting acute acetaminophen hepatotoxicity with acetaminophen-aminotransferase multiplication product and the Psi parameter. Clin Toxicol (Phila). 2014;52(5): 506-11.
13. Akobeng AK. Understanding diagnostic tests 3: Receiver operating characteristic curves. Acta Paediatr. 2007;96(5): 644-7.
14. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in

logistic regression analysis. J Clin Epidemiol. 1996;49(12): 1373-9.
15. Cairney DG, Beckwith HK, Al-Hourani K, Eddleston M, Bateman DN, Dear JW. Plasma paracetamol concentration at hospital presentation has a dose-dependent relationship with liver injury despite prompt treatment with intravenous acetylcysteine. Clin Toxicol (Phila). 2016;54(5):405-10.
16. Chiew AL, Gluud C, Brok J, Buckley NA. Interventions for paracetamol (acetaminophen) overdose. Cochrane Database Syst Rev. 2018;2(2):Cd003328.
17. Chomchai S, Mekavuthikul P, Phuditshinnapatra J, Chomchai C. Sensitivity of dose-estimations for acute acetaminophen overdose in predicting hepatotoxicity risk using the Rumack-Matthew Nomogram. Pharmacol Res Perspect. 2022;10(1):e00920.
18. Hoegberg LCG, Shepherd G, Wood DM, Johnson J, Hoffman RS, Caravati EM, et al. Systematic review on the use of activated charcoal for gastrointestinal decontamination following acute oral overdose. Clin Toxicol (Phila). 2021;59(12):1196-227.

Determinants of Modern Contraceptive Usage among Married Women: A Mixed-Methods Study in a Rural Community of India

Trina Sengupta, MBBS*, Bobby Paul, DCH, M.D.(PSM), Lina Bandyopadhyay, M.D. (PSM)***, Ranjan Das, M.D.(PSM)****, Ankush Banerjee, MBBS, M.D.******

**Department of Preventive & Social Medicine, All India Institute of Hygiene & Public Health, Kolkata, India, **Department of Occupational Health, All India Institute of Hygiene & Public Health, Kolkata, India, ***Department of Preventive & Social Medicine, All India Institute of Hygiene & Public Health Kolkata, India, ****Department of Preventive and Social Medicine, All India Institute of Hygiene & Public Health Kolkata, India, *****Department of Preventive & Social Medicine, All India Institute of Hygiene & Public Health Kolkata, India.*

ABSTRACT

Objective: A woman's multifaceted feelings, knowledge, and perceptions of their intimate relations greatly influence their contraceptive behavior. In addition, women empowerment has been increasingly recognized as a key factor in family planning and reproductive health outcomes. This study aimed to assess modern contraceptive usage and its determinants among currently married women of reproductive-age (WRA) in rural Bengal.

Materials and Methods: This mixed-method study was conducted in a rural area of Hooghly District, West Bengal from April to September 2021. The quantitative strand of the study was conducted by interviewing 210 currently married WRA. The qualitative strand was conducted via focussed group discussions among husbands/mothers-in-law and in-depth interviews with healthcare workers. SPSS software was utilized for quantitative data analysis and factors associated with contraceptive usage were analyzed using logistic regression models. Qualitative data were analyzed thematically.

Result: Currently 114 (54.8%) study participants were using modern contraceptive methods. Education (aOR=7.65, 95% CI=1.85-31.67), empowerment through freedom from family domination (aOR=5.56, 95% CI=1.30-23.66), attitude on contraception (aOR=4.67, 95% CI=1.26-17.19), and family planning counselling (aOR=4.41, 95% CI=1.12-17.33) were found to be significantly associated with modern contraceptive usage. Lack of couple counselling, family support, and knowledge gap was identified as the major barriers to contraceptive usage.

Conclusion: Since a woman's decision-making ability significantly affects their sexual and reproductive health outcomes, effective measures should be undertaken to empower them by creating awareness regarding their rights and freedom to make strategic life choices. Couple counselling should be prioritized to enhance male involvement and eliminate perceived barriers.

Keywords: Contraceptive; empowerment; family support; India; mixed-methods (Siriraj Med J 2022; 74: 666-674)

INTRODUCTION

Family Planning (FP) is a cost-effective investment, the timely intervention of which can help in reducing the impact of high population growth in any country. India became the pioneer country in the world to launch a National Family Planning Programme in 1952. The current slogan of this program: "*Jodi Zimmeder Jo Plan Kare Parivar*" (Responsible couples are those who plan

Corresponding author: Ankush Banerjee
E-mail: ankush.banerjee20@gmail.com
Received 3 August 2022 Revised 24 August 2022 Accepted 7 September 2022
ORCID ID:http://orcid.org/0000-0003-2762-123X
http://dx.doi.org/10.33192/Smj.2022.78

their family) reflects the broader holistic aspects of family planning rather than just mere achievement of replacement level of fertility.[1] Modern contraceptive services which are available under this program are broadly of two types- spacing methods (condoms, oral contraceptive pills, intrauterine devices), and permanent methods (sterilization techniques). The present aim of this program in India is to emphasize the need for a reduction in the number of unintended pregnancies, proper birth spacing, and attainment of the ideal age of first pregnancy.

This current initiative has helped the country to traverse a long way in achieving its goal of slowing the population growth as recent data as per the National Family Health Survey (NFHS-5) statistics have shown that India's TFR (Total Fertility Rate) has come down to the replacement level of 2. Despite this nationwide laudable achievement, some states in India (Madhya Pradesh, Rajasthan, Uttar Pradesh, Jharkhand) are still lagging behind their desired fertility rates. Moreover, India still houses approximately 9.4% of eligible couples with an unmet need for FP, reflecting a significant gap between their reproductive intentions and contraceptive behavior.[2] Family planning plays a central role in women's health by reducing the mortality rate of unsafe abortions and undesired pregnancies.[3] Previous studies have shown that approximately 15.6 million abortions have been performed in India in the year 2015, which was associated with a high rate of unintended pregnancy (70.1 per 1000 women aged 15-49 years).[4] NFHS-5 also showed that the state of West Bengal deserves special mention as the proportion of total unmet needs is particularly high in rural areas (7.8%) as compared to urban areas(5.2%).

Lower rates of contraceptive usage in India; especially in rural areas are largely driven by gender inequality and lack of female autonomy over family planning choices. A study by Shakya et al done in rural India had shown that women empowerment was higher in those couples who received increasing communication regarding contraception.[5] Women empowerment which has been defined as *'the expansion of people's ability to make strategic life choices in a context where this ability was previously denied to them'* has increasingly been recognized as a key factor affecting FP and reproductive health outcomes among women.[6,7] In developing nations like India where gender discrimination is very prominent (particularly in rural areas), a complete understanding of how gender-based power influences the ability to access and use contraceptives is the need of the hour. Moreover, the influence of family-level stakeholders on a woman's choice of contraceptives as well as deficiencies at the health sector level needs further exploration, thus mandating the necessity of mixed-method research. With this backdrop, the present study aimed to assess the contraceptive usage patterns and their major determinants (through quantitative strand) among the currently married women of the reproductive age group (WRA). Again the perspective about using modern contraceptives among family-level stakeholders (husbands and mothers-in-law in the case of this particular study) and also the felt barriers of healthcare workers with regards to providing family planning services to the rural community were explored through the qualitative strand of the study.

MATERIALS AND METHODS

This cross-sectional study with a mixed-methods approach (convergent parallel design) was conducted from April to September 2021in the rural service area of the Rural Health Unit and Training Centre (RHUTC), Singur, Hooghly District, West Bengal. Two primary health centres and 12 sub-centers are situated in the study area from where family planning services are provided to the community comprising of 64 villages. The quantitative strand of this study was conducted among the currently married WRA (15 to 49 years of age), residing in the study area for at least five years. Those who did not give written informed consent were critically ill or had undergone hysterectomy or oophorectomy were excluded. For the qualitative strand, family members comprising of husbands and mothers-in-law of study participants, as well as healthcare workers working in the study area for at least 1 year, were selected.

Sampling:

According to the National Family Health Survey 5 (NFHS 5), the prevalence of modern contraceptive usage among currently married women in West Bengal was found to be 60.6%.[2] Considering *P=0.606*, an absolute error of precision *(L)=10%*, design effect=2 (for cluster sampling in the first stage), and non-response rate=5% (for simple random sampling done in the second stage) the sample size estimated using standard Cochran's formula was *201*.[8] Since a two-stage cluster sampling technique was applied comprising 15 clusters, the final sample size came to be 210.

A list of all 64 villages (along with the population of each village) situated in the rural service area of RHUTC, Singur was taken. From that list, 15 villages were selected through a probability proportional to size (PPS) method. From each of those 15 villages, 16 currently married WRA, residing at those villages were selected by Simple Random Sampling. These selected participants were

approached at their residences with help of field-level healthcare workers. For the qualitative strand, participants were selected purposively and data was collected till the point of data saturation.

Data Collection, Study Tools, and Parameters used:

The quantitative part of the study was conducted through face-to-face interviews among the currently married WRA. A predesigned, pretested structured questionnaire was prepared [translated into the local language of Bengali] which was face and content validated by a team of public health experts. It consisted of the following domains-

a) *Socio-demographic characteristics* which included age, religion, educational status, occupation, socio-economic status, type of family, and number of children.

b) *Knowledge regarding contraceptives and attitude towards their usage:* The knowledge-based section consisted of 12 items where the participants had to respond as "True, False, or Don't Know" (Cronbach's alpha=0.73). The correct response was given a score of '1' while the wrong response or "Don't know" fetched a score of '0'. Total scores ranged from 0 to 12 while the cut-off for having satisfactory knowledge was taken to be the 75th percentile of the total attained score (=8). Attitude toward using contraceptives consisted of a 10-items questionnaire, distributed across three domains: Attitude towards perceived benefits of contraception, perceived barriers from family-level stakeholders, and perceived self-barriers to contraceptive use (Cronbach's alpha= 0.67). Each item had three options (Disagree, Neutral, and Agree) with scores ranging from *-1* to *1*. The total (ranging from -10 to 10) was calculated by adding scores of each domain. Cut-off for having a favorable attitude was taken to be the 75th percentile of the total attained score (=5).

c) *Women's Empowerment Scale adopted from Compendium of Gender Scale by C-change* (previously applied in a similar demographic setting in Bangladesh): It consisted of 18 items, distributed across three sub-scales: *i) women's mobility* (8 items) in which each respondent was given a score of '1' for each place she had visited and an additional score of '1' if she had ever gone there alone. Thus, the scores ranged from 0 to 8 [Cut-off for being empowered was taken as 75th percentile of the attained total score (=5)] *ii) Freedom from Family Domination* (4 items); the responses were scored as 1 for 'Yes' and 0 for 'No' response. A woman was classified as "empowered" if she said that none of the mentioned items ever happened to her or as "not empowered" if any of these items had happened to her, *iii) Economic Security* (4 items): A score of '1' was assigned for each of the following items: if a woman owned her house or land; owned any productive asset; had her cash savings or her savings were ever used for business or money-lending. A woman with a score of ≥2 was classified as being empowered. Levels of women empowerment were measured separately for each sub-scale.[9]

d) *The current usage of modern contraceptives by the study participants* was the outcome variable of this study. Participants were asked whether they have used any of the modern contraceptive methods (oral contraceptive pills, IUDs, condoms, injectables, tubectomy) in the past 6 months. Any participant who had utilized at least one of the mentioned methods was considered a modern contraceptive user.

For the qualitative strand of the study, two focus group discussions (FGDs) were conducted among family-level stakeholders of the study participants. One FGD was conducted among six husbands (median age= 30 years) and the other among six mothers-in-law (median age=67 years) to explore their diverse view-point about family planning. Each FGD took place for about 45 minutes and was conducted with the help of an FGD guide. In addition, in-depth interviews of health workers [one public health nurse (56 years), one multipurpose worker female (42 years), and three accredited social health activists (median age=35 years)] were carried out to explore their perceived barriers to providing contraceptive services to the community. They were interviewed with a pre-tested, semi-structured interviewer guide. All FGDs and IDIs were audio-recorded with the prior consent of the study participants.

Data analysis

Quantitative data were analyzed by Microsoft Excel (2016) & SPSS software (IBM Corp. Chicago. USA. version 16). Appropriate descriptive statistics were utilized for denoting the outcome variables as well as the independent variables. After excluding multicollinearity (variance inflation factor > 5), factors associated with the current usage of modern contraceptives among the study participants were analyzed by a test of significance (p-value <.05) via univariate logistic regression analysis separately. All the biologically plausible significant variables in the respective univariate analysis were included in the final multivariable models. The data obtained through

FGDs and IDIs were simultaneously processed using a manual thematic analysis approach. The records were listened to and transcribed verbatim in Microsoft Word Version 2016. The transcripts were read thoroughly, the important sentences were underlined and the main ideas derived from them were labeled as codes. Appropriate codes were then placed under appropriate themes.

RESULTS

Background characteristics of the study participants

Among 210 study participants, the median age was found to be 28 years (IQR=23-32 years). The majority (82.4%) were in the age group of 20-34 years. 58 (27.6%) participants had an education level of primary or below while only 6.2% of participants were involved in some other occupation and the rest were home-maker. 84 (40%) belonged to socio-economic class III and 79 (37.6%) belonged to class IV according to Modified B.G. Prasad's Scale 2020.[10] (Table 1)

Contraceptive usage patterns among the study participants

Currently, 114 (54.8%) study participants were using any of the modern methods of contraceptives. Among all participants, ligation was used by 21.4% while condom was utilized by 19% of the participants. (Fig 1)

TABLE 1. Descriptive characteristics of the study participants (N=210).

Parameters	Categories	Number (Percentage)
Age (in completed years)	15-19	2 (1.0)
	20-34	173 (82.4)
	35-49	35 (16.6)
Religion	Hindu	190 (90.4)
	Muslim	20 (9.6)
Respondent's education¥	Illiterate	19 (9.0)
	Primary or below	58 (27.6)
	Middle	54 (25.7)
	Secondary	43 (20.5)
	Higher Secondary	21 (10.0)
	Graduate or above	15 (7.2)
Husband's education	Illiterate	15 (7.2)
	Primary or below	66 (31.4)
	Middle	51 (24.3)
	Secondary	29 (13.8)
	Higher Secondary	27 (12.9)
	Graduate or above	22 (10.4)
Occupational status of the respondents	Home-maker	197 (93.8)
	Other professionals	13 (6.2)
Type of family	Joint	81 (38.6)
	Nuclear	129 (61.4)
Socio-economic status£	Class I (upper class)	3 (1.4)
	Class II (upper middle class)	39 (18.6)
	Class III (middle class)	84 (40.0)
	Class IV (lower middle class)	79 (37.6)
	Class V (lower class)	5 (2.4)
No of children	≤2	143 (68.1)
	>2	67 (31.9)

¥ below primary= below 5th standard, primary=passed 5th standard, middle= passed 8th standard, secondary= passed 10th standard, higher secondary= passed 12th standard

£ according to modified B.G Prasad's scale 2020

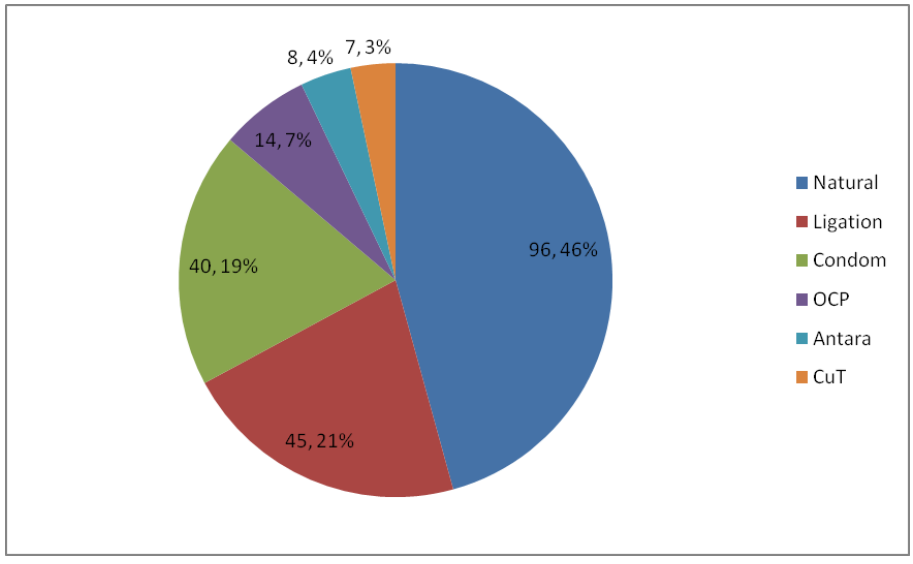


Fig 1. Pie-Diagram showing the current pattern of contraceptive usage among the study participants (n=210).

Knowledge regarding contraceptives and Attitudes toward their usage among the study participants

51.9% (n=109) of the participants had satisfactory knowledge about contraceptives (median score=7, IQR= 6-8) while only 29.5% (n=61) had a favorable attitude towards their usage (median score=3, IQR=1-5).

Women's Empowerment among the study participants

58.6% of women were found to be empowered through the 'Women's Mobility' scale whereas 56.2% were empowered on the 'Freedom from Family Domination' scale. Only 20% of the participants were found to be empowered through the 'Economic Security' sub-scale. It was detected that women who were using any of the modern contraceptive methods were more empowered in all the three sub-scales compared to those not using any. (Fig 2)

Factors associated with the current usage of modern contraceptives by the study participants

Significant factors of modern contraceptives use among study participants were educational status of middle school and above [aOR=7.65, 95% CI = 1.85-31.67], favourable attitude towards modern contraceptive usage [aOR=4.67, 95% CI = 1.26-17.19], empowerment through freedom from family domination [aOR = 5.56, 95% CI = 1.30-23.66], recipients of family planning counselling [aOR = 4.41, 95% CI = 1.12-17.33]. The

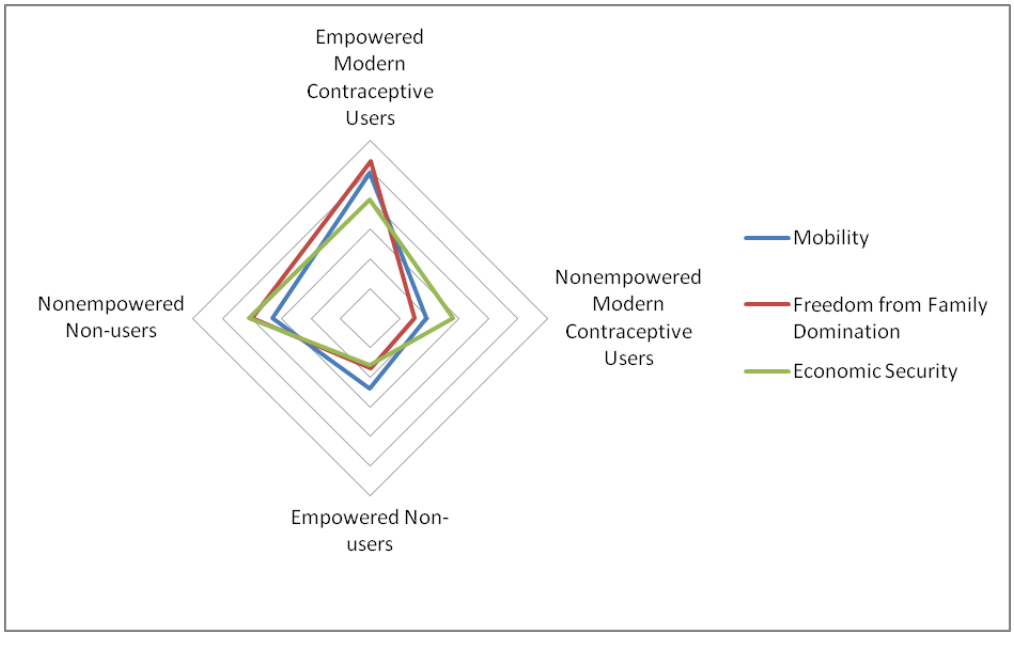


Fig 2. Radar diagram showing the association of usage and non-usage of modern contraceptive methods with woman empowerment.

multivariable regression model deduced was of good fit (non-significant Hosmer-Lemeshow test, P-value >.05) while 42% to 63% of the variance of the outcome variable could be explained by this model. (Table 2)

Qualitative exploratory findings

Qualitative exploration from the two FGDs among family-level stakeholders and 5 IDIs among healthcare workers revealed two major themes: a) Perspective about

TABLE 2. Factors associated with modern contraceptive usage among the study participants: Logistic Regression Analysis (N=210)

Parameters	Total No.	Usage of Modern Contraceptives n (%)	Unadjusted OR[l] (95% CI)	Adjusted OR (95% CI)
Education				
Below middle school	77	20 (25.9)	1 (Ref)	1 (Ref)
Middle school and above	133	94 (70.6)	6.89 (3.04-15.62)	7.65 (1.85-31.67)
Husband's Education				
Below middle school	81	22 (27.1)	1 (Ref)	1 (Ref)
Middle school and above	129	92 (71.3)	6.73 (3-15.07)	3.21 (0.87-11.7)
Type of Family				
Joint Family	81	28 (34.5)	1 (Ref)	1 (Ref)
Nuclear Family	129	89 (66.7)	3.57 (1.67-7.62)	3.31 (0.96-11.44)
Socio-economic[ł] Status				
Below middle class	84	25 (29.7)	1 (Ref)	1 (Ref)
Middle class and above	126	86 (70.6)	5.62 (2.56-12.29)	1.44 (0.37-5.61)
No. of Children				
≤2	143	65 (45.4)	3.11 (1.37-7.02)	3.24 (0.86-12.24)
>2	67	49 (73.1)	1 (Ref)	1 (Ref)
Knowledge regarding modern contraceptives				
Satisfactory	109	77 (70.6)	4.18 (1.97-8.84)	2.10 (0.58-7.55)
Unsatisfactory	101	37 (36.6)	1 (Ref)	1 (Ref)
Attitude towards usage of contraceptives				
Favorable	81	77 (70.6)	5.95 (2.58-13.70)	4.67 (1.26-17.19)
Unfavorable	129	37 (36.6)	1 (Ref)	1 (Ref)
Women's mobility				
Empowered	123	82 (66.7)	3.54 (1.68-7.48)	1.10 (0.26-4.68)
Non-empowered	87	32 (36.8)	1 (Ref)	1 (Ref)
Freedom from family Domination				
Empowered	118	89 (75.4)	8.31 (3.71-18.62)	5.56 (1.30-23.66)
Non-empowered	92	25 (27.1)	1 (Ref)	1 (Ref)
Economic Security				
Empowered	94	67 (71.2)	3.66 (1.72-7.77)	2.86 (0.69-11.81)
Non-empowered	116	47 (40.5)	1 (Ref)	1 (Ref)
Received FP Counselling				
Yes	136	101 (74.2)	3.66 (1.72-7.77)	4.41 (1.12-17.33)
No	74	13 (17.5)	1 (Ref)	1 (Ref)

l OR- Odds Ratio, CI- Confidence Interval

ł according to B.G Prasad Scale 2020

Hosmer-Lemeshow test statistic=0.61, Cox and Snell's R^2=0.42 & Nagelkerke's R^2=0.63

modern methods of contraceptives and b) Barriers to adopting modern family planning methods. Communication gap, woman's authority in choice of contraception, and lack of couple counselling were the major codes under the theme 'Barriers to adopting modern family planning methods'. Lack of knowledge and experience emerged as the most important code under the theme of 'Perspective about modern methods of contraception. (Table 3).

TABLE 3. Juxtaposed Findings of both Qualitative and Quantitative Inquiry on Modern Contraceptive Usage among study participants.

Associated Survey Themes	Quantitative Components	Qualitative Components with Quotable quotes	Qualitative Codes
Barriers to adopting modern family planning methods	"One can feel embarrassed in discussing contraceptives with spouse"-46.7% responded "YES"	Less interaction about sexual and reproductive life with spouses.	Communication Gap
	"Husband's objection to contraceptive methods can prevent a woman from using it"- 52.3% responded 'YES'	Husband's domination in decision-making.	Women's Authority in Contraceptive Choices
	"Does your partner know about newer contraceptive methods like Antara"-87.4% responded "NO"	Deficient knowledge of husbands about various contraceptive choices. "In our time there was no one to teach us about reproductive health. It is important to teach sex education in schools"- a 35 years old male quoted.	Knowledge Gap
	"Have you received a couple counseling"- 98% replied "NO"	Men do not feel enough empowered for choosing contraceptives due to a lack of couple counseling. *"ASHA didi told my wife to use contraceptives, but it would have been better if both of us were counselled in private. We could share more things then"- a 28-year-old husband remarked. "When we visit house to house we only get the women at home, their husbands are at work then. And especially we target the women who come for ANC or PNC clinics at health-centres, for FP counseling"*- a 38-year-old ASHA told.	Lack of Couple Counselling
Perspective about Modern methods of Family Planning	"Change in Mother-in-law's attitude may improve contraceptive use"- 46.2% replied "YES"	Preformed notions about harmful side-effects of modern-day methods of contraception and lack of experience. *"These modern-day girls don't discuss their lives with us, the elderly. They won't take our advice too. So I don't talk about this with my daughter-in-law and I personally never used any contraceptives in our times"*- a 68-year-old mother-in-law remarked.	Lack of knowledge and experience

DISCUSSION

The present study tried to address holistically all the aspects of FP and elicited some major determinants of practicing modern FP methods among currently married women of reproductive age group in a rural area of Bengal. A study by Ahirwar RK et al, done in central India showed that 88.5% of study participants had never used any contraceptive methods while the current study found a considerable proportion of reproductive-aged married women using any of the modern contraceptive methods (54.8%).[11] Another research work done in rural Maharashtra, India by Dixit A et al found this percentage to be only 38.3% (more than one-third of the study sample).[12] The current study detected that tubal ligation was the most commonly used method of contraception which was found concordant with a study conducted at the national level in India by Ewerling F et al, where it was found that the majority of the reproductive-aged married women were using modern methods of contraceptives mostly in the form of sterilization. Condom and oral contraceptives were the second and third most commonly used contraceptive methods.[13] Another study done by Talungchit et al in Thailand showed that the most commonly used method of contraception among teenage multigravida and primigravida was oral contraceptive pills, while only 5% among primis and 25% among multigravidas were using contraceptives.[14] A retrospective cohort study conducted at a medical college clinic in Thailand showed that approximately 15% of women had never used any contraceptive methods.[3] A research work conducted by Chopra S et al demonstrated that acceptance of a permanent method of contraception among a tribal population in northern India was only 5%.[15]

Although the status of women in India has improved over time, across different dimensions, gender discrimination and patriarchal social norms still remain a burning issue in this nation, especially in rural areas. Only 20% of participants in the present study were found to be empowered by means of economic security. Moreover, this study found that women who are free from any kind of family domination were more likely to use modern contraceptives. Another study of rural Maharashtra, India by Reed E et al elicited that there was a significant association between woman's access to money and the usage of condoms or other methods of contraceptives. Other significant determinants detected in that research work were women's control over reproductive health decision-making and freedom of movement to seek health care.[16] A study conducted in Egypt by Samari G et al found that determinants of women empowerment like household decision-making, non-acceptance towards intimate partner violence, and joint decision-making power are significantly associated with modern contraceptive usage.[17]

This mixed-methods study is a strength in itself as the qualitative exploration led to an in-depth understanding of the perspectives of modern contraceptive usage and its important barriers among the family-level stakeholders and healthcare providers. Previous studies in India and abroad had explored some of the major barriers to modern contraceptive usage such as a woman's fear of side effects or other health concerns and the absent cafeteria approach.[18-20] A scoping review on determinants of unmet need of family planning in low and middle income countries by Wulfian et al showed that the reasons behind the non-usage of contraceptives among women were mostly opposition from husbands, their fear of infidelity and fear of side effects.[21] From the health workers' point of view, the barriers that had been explored by prior studies are low prioritization of contraceptive training, disputes over funding, and an overburdened health system.[22,23] In addition to these above findings, the current study found some new emerging barriers like lack of couple counselling, misinformation from peer groups, generalized fear, and misconception about modern contraceptive methods among the elder generation. Hence appropriate and suitable interventions are necessary so that they can gradually adjust to and overcome these pre-existing as well as emerging barriers to modern family planning use.

Limitations

As the study participants came from a unique population, the generalizability of the present study had been compromised. Since this study was of cross-sectional nature, a causal relationship between the variables and contraceptive usage could not be established. Moreover, as some information obtained was recall-based, bias might be possible.

CONCLUSION

Investing in family planning is the most intelligent step that developing nations like India can undertake to improve their socio-economic and maternal-child health scenario. Thus, in order to improve the overall family planning practice, certain measures like awareness generation, education, extending help for self-empowerment, and most importantly economic independence among women need to be prioritized. Support from healthcare facilities like couple counseling is needed to overcome barriers like the lack of involvement of family-level stakeholders,

especially the husbands in process of making a decision on contraception. Ensuring the availability of all kinds of modern contraceptives in remote and rural health centres should also be done sustainably.

ACKNOWLEDGMENTS

The authors were grateful to the officer-in-charge and staff of RHUTC, Singur, for their constant support and co-operation during this difficult pandemic phase throughout our study. We extend our heartfelt gratitude to all the participants who participated in this study.

Declaration of Conflicting Interests: The authors declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Ethical Issues: This study was conducted after ethical clearance from the Institutional Ethics Committee of All India Institute of Hygiene and Public Health, Kolkata. Appropriate written informed consent was taken from the participants before conducting the study. They were assured of the confidentiality of the data provided by them. All other ethical principles as per the Declaration of Helsinki were strictly adhered to.

REFERENCES

1. Centre for Health Informatics of the National Health Portal. National Family Planning Programme [Internet]. New Delhi, India: National Health Portal; 2017 [cited 2022 Mar 23]. Available from: https://humdo.nhp.gov.in/about/national-fp-programme/
2. International Institute for Population Sciences. National Family Health Survey-5 [Internet]. Mumbai, India: International Institute for Population Sciences;2019-21 [cited 2022 May 20]. Available from: http://rchiips.org/nfhs/
3. Parkpinyo N, Panichyawat N, Sirimai K. Early Removal of the Etonogestrel Contraceptive Implant and Associated Factors Among Users at the Urban Family Planning Clinic in Siriraj Hospital, Bangkok, Thailand. Siriraj Med J. 2021;73(6):399-405.
4. Singh S, Shekhar C, Acharya R, Moore AM, Stillman M, Pradhan MR, et al. The incidence of abortion and unintended pregnancy in India, 2015. Lancet Glob Health. 2018;6(1):e111-20.
5. Shakya HB, Dasgupta A, Ghule M, Battala M, Saggurti N, Donta B, et al. Spousal discordance on reports of contraceptive communication, contraceptive use, and ideal family size in rural India: a cross-sectional study. BMC Womens Health. 2018;18(1):147.
6. Kabeer N. Resources, agency, achievements: Reflections on the measurement of women's empowerment. Dev Change. 1999;30(3):435-64.
7. Blanc AK. The effect of power in sexual relationships on sexual and reproductive health: an examination of the evidence. Stud Fam Plann. 2001;32(3):189-213.
8. Saha I, Paul B. Essential of Biostatistics & Research Methodology. 3rd Edition. Kolkata: Academic Publishers; 2020.
9. Nanda G. Compendium of Gender Scales [Internet]. Washington DC, USA: FHI 360/C-Change; 2011 [cited 2022 May 20]. Available from: http://gender.careinternationalwikis.org/_media/c-change_gender_scales_compendium.pdf.
10. Dalvi TM, Khairnar MR, Kalghatgi SR. An update of BG Prasad and Kuppuswamy socio-economic status classification scale for Indian population. Indian J Pediatr. 2020;87(7):567-8.
11. Ahirwar RK, Kumar S, Gupta S, Niranjan A, Prajapati K, Rawal S. A study to assess the socio-demographic factor and knowledge, attitude, and practice of family planning methods in Lohpeeta migrants tribe located in Shivpuri central India. J Family Med Prim Care. 2021;10(3):1308-12.
12. Dixit A, Johns NE, Ghule M, Battala M, Begum S, Yore J, et al. Male–female concordance in reported involvement of women in contraceptive decision-making and its association with modern contraceptive use among couples in rural Maharashtra, India. Reprod Health. 2021;18(1):139.
13. Ewerling F, McDougal L, Raj A, Ferreira LZ, Blumenberg C, Parmar D, et al. Modern contraceptive use among women in need of family planning in India: an analysis of the inequalities related to the mix of methods used. Reprod Health. 2021;18(1):173.
14. Talungchit P, Lertbunnaphong T, Russameecharoen K. Prevalence of repeat pregnancy including pregnancy outcome of teenage women. Siriraj Med J. 2017;69(6):363-9.
15. Chopra S, Dhaliwal L. Knowledge, attitude and practices of contraception in urban population of North India. Arch Gynecol Obstet. 2010;281(2):273-7.
16. Reed E, Donta B, Dasgupta A, Ghule M, Battala M, Nair S, et al. Access to money and relation to women's use of family planning methods among young married women in rural India. Matern Child Health J. 2016;20(6):1203-10.
17. Samari G. Women's empowerment and short-and long-acting contraceptive method use in Egypt. Cult Health Sex. 2018;20(4):458-73.
18. Jain M, Caplan Y, Ramesh BM, Isac S, Anand P, Engl E, et al. Understanding drivers of family planning in rural northern India: An integrated mixed-methods approach. PLoS One. 2021;16(1):e0243854.
19. McClendon KA, McDougal L, Ayyaluru S, Belayneh Y, Sinha A, Silverman JG, et al. Intersections of girl child marriage and family planning beliefs and use: qualitative findings from Ethiopia and India. Cult Health Sex. 2018;20(7):799-814.
20. Shafi S, Mohan U. Perception of family planning and reasons for low acceptance of NSV among married males of urban slums of Lucknow city-A community based study. J Family Med Prim Care. 2020;9(1):303-9.
21. Wulifan JK, Brenner S, Jahn A, De Allegri M. A scoping review on determinants of unmet need for family planning among women of reproductive age in low and middle income countries. BMC Womens Health. 2016;16:2.
22. Kc H, Shrestha M, Pokharel N, Niraula SR, Pyakurel P, Parajuli SB. Women's empowerment for abortion and family planning decision making among marginalized women in Nepal: a mixed method study. Reprod Health. 2021;18(1):28.
23. Walker SH, Hooks C, Blake D. The views of postnatal women and midwives on midwives providing contraceptive advice and methods: a mixed method concurrent study. BMC Pregnancy Childbirth. 2021;21(1):411.

Simulated Surgical Model Design for Myringotomy and Tympanostomy Tube Insertion in Children using Medical Image Processing and 3D-Printing Technologies

Kawin Pratumaneechai, M.A.*, Archwin Tanphaichitr, M.D., Thanongchai Siriapisith, M.D.***, Vannipa Vathanophas, M.D.**, Kitirat Ungkanont, M.D.**, Puttinun Patpituck, Ph.D.***

**Vejnitatphattana School, Medical Educational Technology Program, **Department of Otorhinolaryngology, ***Department of Radiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Objective: Researchers aimed to design surgical simulation models using medical image processing and 3D-printing technologies to train otolaryngologie residents with correct surgical techniques and study their skills improvement.

Materials and Methods: The models were produced for three age ranges (group A: 8-12 years old, group B: 3-7 years old, and group C: 10 months - 2 years old). Eleven residents were practiced from older to younger child models. Overall surgical time and results were evaluated to determine improvement. Both residents and specialists assessed satisfaction surveys after training.

Results: The median operational time was significantly reduced by 64.57% in model A and 50.24% in model B ($p < 0.05$). Operating time and surgical skills improved in order from models A, B, and C. Model C showed the most improvement with correct operational techniques in myringotomy incision (66.7%, $p = 0.003$) and tympanostomy tube insertion (48.5%, $p = 0.011$). Residents' and specialists' satisfaction assessments exhibited prominent satisfaction results with surgical simulation model training.

Conclusion: Surgical simulation models training enhanced residencies' confidence and improved correct surgical techniques. Residencies can gradually practice skills from fundamental to more complicated techniques in younger child model where symptom occurs.

Keywords: Myringotomy; tympanostomy tube insertion; medical image processing; 3D-print; surgical simulation (Siriraj Med J 2022; 74: 675-683)

INTRODUCTION

Otitis media with effusion (OME) is accumulation of fluid in middle ear that causes inflammation and fluid build-up behind eardrum. 90% occurred in children between six months and four years.[1,2] Build-up of fluid in middle ear affects tympanic membrane and middle ear functions, leading to conductive hearing loss and occasional pain from pressure changes while also affecting speech, cognition, behavioral problems, and language development. Ear tube insertion is a procedure whereby doctor inserts a tympanostomy tube into eardrum to ameliorate ear infections and allow drainage of excess fluid

Corresponding author: Puttinun Patpituck
E-mail: puttinun.pat@mahidol.ac.th
Received 25 April 2022 Revised 11 September 2022 Accepted 19 September 2022
ORCID ID:http://orcid.org/0000-0002-2669-7999
http://dx.doi.org/10.33192/Smj.2022.79

from middle ear. The myringotomy and tympanostomy tube insertion require precise surgical skills under a microscope.[3-5]

Medical practitioners should improve their knowledge of operational procedures and practice performing with correct surgical techniques to reduce surgical risks. The major problem is the procedures are difficult to learn and practice. Anatomical structures of children are significantly different from adults concerning anatomy and physiological differences in bone growth and development. The research study by Ungkanont et al., also found that 70% of the children with cleft palate had their first myringotomy before they were 3 years old, while 62.5% of normal children had their first ventilation tubes within their first 5 years. The mean age at first myringotomy in children with cleft palate was 1.3 years old, which required precision surgical techniques for operation.[6] The complication from myringotomy and tympanostomy tube insertion can also develop to granular myringitis in the patient with a history of other ear diseases. Having the ability to better operations can reduce the risk factors for the development of complications after surgery.[7] Many simulation models are designed for practicing myringotomy and tympanostomy tube insertion. But use easily found materials, such as boxes and syringes to simulate the ear holes,[8-10] which model appearance is unrealistic.[11-13] Also the medical models are designed as adult, and have not been designed as a child, where the symptom most occurs.[14] Hence researchers aimed to design and build more realistic simulation models of the children in various age ranges that allow unlimited practice with the operating microscope and surgical instruments used during the actual operation, which can be a beneficial teaching tool for medical practitioners. The designed simulation models were continued to be beneficially used for training the otolaryngology residents at the Department of Otorhinolaryngology, Siriraj Hospital, Mahidol University, until the present. In addition, the researchers suggested that the study subjects should be conducted with a larger population in future studies

Medical imaging data conversion to surgical simulation models

Image data were collected from CT scans of healthy children's heads and related organs. Patients' age ranges were categorized into three groups; group A: 8 to 12 years old, group B: 3 to 7 years old, and group C: 10 months to 2 years old. Two-dimensional image data from CT scans were transferred to Materialize Mimics software by uploading DICOM files and selecting the particular tissues.[15,16] Anatomical positions of tympanic membrane, ossicular ligaments, and oval window were captured to ensure that ear canal, tympanic cavity, and middle ear were accurate. 2D image data processing in Mimics software makes it possible to generate three-dimensional models in stereolithography (.stl) file format that can be designed on a computer-aided design software (Pixologic ZBrush).[15-17] Connective parts between ears and head were designed, and model head circumference was adjusted to match average data from World Health Organization (WHO). Standard ratios of model perimeters as follows; group A: 52 – 54 cm, group B: 49 – 51 cm and group C: 45 – 48 cm. (Fig 1)

The designed 3D models were imported to Ultimaker Cura 4.2.1 software to convert to G-code (gcode) that commands path and instructs 3D printing machine (Anet A8 Plus).[18,19] The selected thermoplastic material was Polylactic-acid plastic (PLA). Injector temperature was set at 220 °C, with build plate temperature at 70 °C and layer detail as 'Normal' (0.15 mm). Solid 3D printed models appeared horizontal lines of plastic filaments. Hence surfaces of the models were polished to remove lines from silicone casting molds. Yellow polyester putty was used as a primer before workpieces were polished using No. 800 and No. 1000 sandpapers. Clean water was applied to surface before spraying with a gray primer (Leyland Polypropylene Primer) to detect any rough areas on 3D printed models. Primer provided a smooth surface that was easy to separate from mold. Various types of silicone rubber are used for casting medical models, depending on underlying objectives and level of realism required by application. 1300 silicone rubber was chosen

Fig 1. Surgical simulation models designs for three age ranges
(Model A, Model B and Model C)

to create ear models, with RTV-357 silicone rubber to create head models. A food-grade plastic bag with 15-18 microns thickness was selected by specialists in pediatric otolaryngology for artificial eardrum. When viewed from a microscope, the material is slightly opaque with a glossy surface, and surface tension is similar to eardrum. The artificial eardrum was attached by applying a thin layer of latex glue over cylindrical end of ear model, using an O-ring (No. 113) to tighten. After straightening plastic sheet, a skin-color marker was used to draw the line to refer to malleus bone location.

MATERIALS AND METHODS

Ethical consideration

The research was approved by the Institutional Review Board of the Faculty of Medicine Siriraj Hospital, Mahidol University (SIRB) (Si 251/2020). The study has requested permission to use CT scan image data of healthy children's heads and related organs from the Department of Radiology, Siriraj Hospital, Mahidol University. There are three groups of patients categorized by the age range: the age of 10 months to 2 years old, 3 to 7 years old, and 8 to 12 years old, with three patients in each group, for a total of 9 patients. The names are concealed, and the data only indicates their ages and genders. The subjects of the study are first-year otolaryngology residents who had not previously experienced myringotomy and tympanostomy tube insertion. The researchers announced participants who volunteered by posting an invitation poster at the Department of Otorhinolaryngology, Siriraj Hospital, Mahidol University inviting residencies to participate in the practice session. If any participants did not want to join this research project, their study and grades would not be affected. The clarification participant information sheet was distributed to the participants who voluntarily registered before joining the research project, which informed the study's objectives, methods, data collection, and expected benefits. The participants' data were kept in confidence, and the research findings will be reported in the overall results.

Recruitment and instruction

Residencies who volunteered to participate in the study were advised the myringotomy and tympanostomy tube insertion procedures before training. Instruction included lectures and a demonstration of operational techniques using surgical simulation models. The list of equipment used for training includes; 1) microscope, 2) ear speculum, 3) alligator forceps, 4) myringotomy knife, 5) straight pick, 6) tympanostomy tube or grommet; polyethylene tube no.90, 7) simulated surgical models, and 8) adjustable table

Data collection

Two videos recorded otolaryngologie residents' training using a microscope (OPMI Pico) to view tympanic membrane, DSLR camera (Canon 7D) captured hand movements and the use of equipment. (Picture 2) Participant information was encoded by a number instead of student's name and ID, with faces and voices concealed. Participants performed in order from model A, B, and C, the older to younger child model. The second training session was repeated in the same setting after one week. The videos were arranged side-by-side to show time duration of operation.

Fig 2. Myringotomy and tympanostomy tube insertion training.

Case record form and surveys

1. The residencies' operation times were recorded from the beginning to the last procedure. The three specialists in pediatric otolaryngology watched videos and assessed residencies' proficiencies in each procedure in the case record form. By the scoring criteria, 10 (completed), 5 (not completed), and 0 (not performed) in the corresponding score box.

2. The satisfaction survey was used to assess the residencies' satisfaction in surgical simulation model training. The residencies rated their level of agreement as strongly agreed, agreed, disagree, and strongly disagree regarding knowledge, understanding, and confidence in operation. The comments section is open for residencies to suggest improvement ideas for the surgical simulation model training.

3. The specialists in pediatric otolaryngology also

completed a satisfaction survey to assess the effectiveness of the training with simulated surgical models compared with their regular teaching experiences, rated by the level of satisfaction as very satisfied, satisfied, neutral, and not satisfied. The comments section is open for specialists in pediatric otolaryngology to suggest improvement ideas for the surgical simulation model training.

Statistical analysis

1) The residencies' operation time were compared, ranked, and calculated statistically by Wilcoxon's Signed Ranks Test. PASW Statistics (SPSS) version 18.0 (SPSS Inc., Chicago, IL., USA) was used for the statistical analysis. If the p-value less than 0.05, the result is concluded that the operation time was reduced with statistical significance. And would be concluded that the medical models could enhance faster operation time.

2) The analysis of the residencies' operation skill results; McNemar Bowker Test was used in each procedure to test the significance of the score comparison in Test 1 and Test 2, and calculated the testing results with the statistical SPSS Program. When the p-value less than 0.05, it is concluded that the residencies' skill was developed in the particular operative procedure with statistical significance.

3) Satisfaction survey analysis; the competencies of the medical models were rated by level of agreement as strongly agreed, agreed, disagree, and strongly disagree regarding knowledge, understanding, and confidence in operation. The researchers summarized the statistical scoring results in the table with frequency distribution and conversed them into percentages.

RESULTS

Improvements in operational time

The otolaryngologie residents spent a shorter time performing model A and B, with statistically significant results. The median operational time was 64.57% faster in model A, while model B was 50.24% faster. The Wilcoxon signed-rank test's statistical analysis showed a two-tailed significance ($p < 0.05$) in model A ($p = 0.007$) and model B ($p = 0.003$). Results in model C gave fluctuating operational time ($p = 0.147$) Overall operational time were reduced from model A, B, and C in order, indicating that residencies have gained familiarity by frequent repetition practices. (Table 1)

TABLE 1. Operational time using the simulated surgical models.

	n	Time (sec) Median	IQR	Minimum	Maximum
Model A Test 1	11	587	325	197	1195
Model A Test 2	11	208	263	125	515
Test 1 – Test 2	11	379	363		
Percentage		64.57%			
P-value[a]		0.007*			
	n	Time (sec) Median	IQR	Minimum	Maximum
Model B Test 1	11	613	205	156	759
Model B Test 2	11	305	280	157	483
Test 1 – Test 2	11	308	159		
Percentage		50.24%			
P-value[a]		0.003*			
	n	Time (sec) Median	IQR	Minimum	Maximum
Model C Test 1	11	382	263	104	1044
Model C Test 2	11	204	267	88	527
Test 1 – Test 2	11	178	270		
Percentage		46.6%			
P-value[a]		0.147			

[a]Wilcoxon signed rank test

Operational skill results

The operational skill results also indicated that residencies improved their correct surgical techniques in order from model A, B, and C. The number of proficiencies that showed statistical significance evaluated by the McNemar-Bowker test was increased from older to younger child model ($p < 0.05$). (Table 2) For instance, model A showed one proficiency improvement 54.4% in checking the tympanostomy tube position on completion ($p = 0.004$). Model B showed three proficiencies improvement as 42.4% in correction of holding the myringotomy knife in the right direction ($p = 0.039$). Correction of incision size showed 54.6% improvement ($p = 0.021$), and 47.7% improvement in checking the tympanostomy tube position on completion ($p = 0.057$). While model C showed four proficiencies improvement as 42.4% in microscope camera adjustment ($p = 0.046$). The correction of incision size showed 66.7% improvement ($p = 0.003$). The correction of tympanostomy tube insertion showed 48.5% improvement ($p = 0.011$), and 42.4% improvement in checking the tympanostomy tube position on completion ($p = 0.039$).

Additional results in model A also occurred that some residencies had difficulty making an incision in the correct position. The incision should be made on the anterosuperior quadrant or anteroinferior quadrant of tympanic membrane. Some residencies could not estimate the incision size during first performance. The incision is too wide and causes to dropped the tympanostomy tube into middle ear, which is unacceptable in actual operation. Results in model B also showed that some residencies used straight pick to push the tympanostomy tube through the myringotomy incision using the outer flange instead of inner flange. Model C results also showed that some residencies encountered difficulties pushing tympanostomy tube into the incision in narrower ear canal and often accidentally touched the malleus, which is also unacceptable in actual operation.

TABLE 2. Operational skills using the simulated surgical models.

Item	Proficiency	P-value[b] (Percentage of Improvement) Model A	Model B	Model C
1	Adjust the microscope camera to clearly see the eardrum	18.2% (0.613)	6% (0.816)	42.4% (0.046*)
2	Select the correct ear speculum	66.6% (0.388)	18.1% (0.312)	12.2% (0.5)
3	Hold the myringotomy knife in the right direction	24.2% (0.065)	42.4% (0.039*)	36.4% (0.109)
4	Correct incision position	30.4% (0.168)	48.4% (0.076)	36.4% (0.2144)
5	Correct incision size	30.2% (0.5287)	54.6% (0.021*)	66.7% (0.003*)
6	Use the knife without touching the ear canal or causing the eardrum to tear apart	12% (0.5313)	12.2% (0.774)	24.2% (0.359)
7	Handle the tympanostomy tube properly using alligator forceps	48.5% (0.078)	6.1% (1)	36.4% (0.302)
8	Insert the tympanostomy tube in the correct direction at the myringotomy incision	24.2% (0.348)	-12.1% (0.847)	48.5% (0.011*)
9	Use the straight pick to push the tympanostomy tube or alligator forceps to insert the tympanostomy tube through the myringotomy incision	12.1% (1)	60.6% (0.15)	-15.2% (0.668)
10	Check the tympanostomy tube positionon completion	54.4% (0.004*)	47.7% (0.057*)	42.4% (0.039*)

[b] McNemar-Bowker test

Recidencies satisfaction with the training

The satisfaction assessment showed strongly agreed results in all competencies. (Table 3) The major influencing competencies indicated that the simulated surgical models effectively increased residencies' operational skills ($\bar{x}$ = 3.91), and the models should be expanded to encompass different medical practices. ($\bar{x}$ = 3.91). The minor influencing competencies showed that the simulated surgical models increased residencies' confidence when operating ($\bar{x}$ = 3.82). The simulated surgical models were effective and easy to understand learning techniques ($\bar{x}$ = 3.82) and had a level of difficulty in practicing skills suitable for residencies year ($\bar{x}$ = 3.82). Additional comments from residencies also noted that using the same tympanostomy tube repeatedly during training caused the outer flange to deteriorate. The residencies suggested changing the tympanostomy tube after repeating the training, while the material used for the tympanic membrane was too elastic, making it difficult to attach to the tympanostomy tube.

Satisfaction of the specialists in pediatric otolaryngology with simulated surgical models training

The simulated surgical models can simulate complicated hands-on surgical procedures to supplement observation before operating on patients. The major influence on specialists' satisfaction indicated that simulated surgical models improved residencies' understanding in operational processes ($\bar{x}$ = 3.67). The residencies gained confidence through practicing the operational techniques ($\bar{x}$ = 3.67) and were consistent with the learning objectives according to the training content ($\bar{x}$ = 3.67). The simulated surgical models were easy to use, durable, and easy to maintain for repetitive training ($\bar{x}$ = 3.67) and should be further developed to simulate other operational techniques in the future ($\bar{x}$ = 3.67). (Table 4) Additional comments from the specialists in pediatric otolaryngology also suggested that the position of the malleus was uncertain. This issue caused some residencies to accidentally touch the malleus. In some cases, the tympanic membrane was not tight and less realistic, while the O-ring should not be too tight because this caused narrowing of the ear canal and presented difficulties for some residencies.

DISCUSSION

The results showed that the residencies have gained more confidence and enhanced surgery fundamentals through repetitive training with surgical simulation models. Residencies can understand operational procedures with

TABLE 3. Training satisfaction using the simulated surgical models.

Item	Competency	Mean ($\bar{x}$)	SD	Result
1	Simulated surgical models improved knowledge and understanding of the operational processes	3.73	0.47	strongly agreed
2	Simulated surgical models increased operational skills	3.91	0.30	strongly agreed
3	Simulated surgical models increased confidence when performing the operation	3.82	0.40	strongly agreed
4	Simulated surgical models were practical and easy-to-understand learning techniques	3.82	0.40	strongly agreed
5	Simulated surgical models were easy to use, durable and simple to maintain	3.55	0.52	strongly agreed
6	Simulated surgical models were effective tools that increased their roles as practitioners	3.73	0.47	strongly agreed
7	Simulated surgical models had a level of difficulty in practicing skills suitable forresidence's year	3.82	0.40	strongly agreed
8	Simulated surgical models should be developed to practice skills in other clinical areas	3.91	0.30	strongly agreed

* Four-points scale: 1 = strongly disagree; 4 = strongly agreed

TABLE 4. Satisfaction of the specialists in pediatric otolaryngology with simulated surgical models training.

Item	Competency	Mean ($\bar{x}$)	SD	Result
1	Simulated surgical models improved residencies' operational processes	3.67	0.58	very satisfied
2	Simulated surgical models increased residencies' operational skills	3.33	0.58	satisfied
3	Simulated surgical models increased residencies' operational confidence	3.67	0.58	very satisfied
4	Simulated surgical models were consistent with the learning objectives according to the training content	3.67	0.58	very satisfied
5	Simulated surgical models had a level of difficulty in practicing skills appropriate for the residencies' year	3.33	0.58	satisfied
6	Simulated surgical models increased interaction between the instructor and residencies	3.33	0.58	satisfied
7	Simulated surgical models were cost-effective and suitable for the number of trainees	3.33	0.58	satisfied
8	Simulated surgical models were easy to use, durable and easy to maintain	3.67	0.58	very satisfied
9	Simulated surgical models should be further developed for other operational simulations	3.67	0.58	very satisfied

* Four-points scale: 1 = not satisfied; 4 = very satisfied

an improvement of operation times in Model A (64.57%, p = 0.007) and Model B (50.24%, p = 0.003) and develop better skills in correct incision size in Model B (54.6%, p = 0.021) and Model C (66.7%, p = 0.003), and inserting the tympanostomy tube in the correct direction in Model C (48.5%, p = 0.011), which these skills are essential to reduce surgical complications. The major influence on residents' satisfaction showed that the surgical simulation models effectively increased residencies' operational skills ($\bar{x}$ = 3.91). Residencies can encounter problems or unforeseen circumstances and help improve skills for practitioners who lack the experience to reduce the risks of surgery with real patients. To date, residencies can only learn by observing medical lecturers or seniors. Hence, the second major influence on residents' satisfaction showed that the models should be expanded to encompass different medical practices. ($\bar{x}$ = 3.91) The specialists' satisfaction assessment also showed prominent satisfaction results with surgical simulation model training. The major influence on specialists' satisfaction showed that simulated surgical models improved residencies' understanding in operational processes ($\bar{x}$ = 3.67), gained confidence through practicing the operational techniques ($\bar{x}$ = 3.67), and were consistent with the learning objectives according to the training content ($\bar{x}$ = 3.67). The models were easy to use, durable, and easy to maintain for repetitive training ($\bar{x}$ = 3.67) and should be further developed to simulate other operational techniques in the future ($\bar{x}$ = 3.67). The specialists' satisfaction was evaluated with the normal teaching experiences and found that the simulation-trained residencies were notably outperformed. Simulators can provide a safe and standardized method for surgery training without risks. They allow trainees to practice their surgical skills, contribute detailed feedback by performance assessment, and enable better patient safety and standards of care. The evidence exhibits that surgical skills are acquired through simulation training, and specialists are positively considered to transfer simulation training to the clinical teaching setting and improve operative outcomes.

From the study, the surgical simulation models can be further developed to become more realistic. The following areas of improvement were discussed. 1) The ear canal in Model C is narrower than Model A and B. Some residencies encountered difficulties during operation in model C that were more complicated than Model A and B. 2) Scanned file of the ear canal is too narrow, which may cause problems for residencies in operating procedures. CT scans were captured with the head of the patient (temporal area) lying on a pillow, which caused uncertainties in ear canal size. Future simulation models need to adjust the ear canal width to be suitable for production and training purposes. 3) The material used to imitate the tympanic membrane had a muscular and bouncing surface tension, which caused difficulties when placing the tympanostomy tube in the incision. A more appropriate material should be used. Whereas, the appearance of the studied material when looking through a microscope camera was already similar to real tympanic membrane. 4) A skin color marker pen used to imitate the malleus was uncertain. Marking the position of the malleus should have a standard setting. 5) Some residencies pressed down on the outer-ear during training, which compressed the gap inside the middle ear part and made it difficult to insert the tympanostomy tube in the incision. The gap between the tympanic membrane and the middle ear should be increased. With more realistic simulation model design, medical practitioners can practice with the microscope and surgical instruments as the actual operation. Real-time myringotomy simulations using virtual reality (VR) are beneficial, with savings on manufacturing costs.[20-22] However, the tangible surgical simulation models lack response to feedback from the practitioners as other simulations. With current technology, surgical simulation models can incorporate capacitive sensing technology to track equipment placement and quantitatively measure operator proficiency in live surgical procedures.[23] Researchers suggest combining medical models with innovative features such as sensors and intelligent tracking systems that can enhance realistic experiences for medical practitioners. Researchers believe that more research studies will need to develop and transfer the benefits of surgical simulation training to clinical teaching and aim to create surgical simulation models for other symptoms in the future.

The limitation in this study was due to a small number of sample subjects, which affected the study's statistical results. Hence, the researchers recommended that future research should be conducted with more sample subjects in the study.

CONCLUSION

The myringotomy procedure and tympanostomy tube insertion require specialized training with the microscope. Using surgical simulation models as a learning tool increased confidence and improved the expertise of the residencies to reduce surgical risks and improve their knowledge of operational procedures. Training with various model age ranges also allows residencies to frequent repetition practices from fundamental to experiment with more complicated techniques in the younger child model where the symptom most occurs.

ACKNOWLEDGEMENTS

The authors extend their deepest gratitude to the Veinitatphattana School, Department of Otorhinolaryngology, and Department of Radiology, Faculty of Medicine Siriraj Hospital, Mahidol University for their support and contributions to the design and development of the simulator, and for providing the location for testing and clinical advice.

Conflict of interest: The authors declare they have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

1. Ito M, Takahashi H, Iino Y, Kojima H, Hashimoto S, Kamide Y, et al. Clinical practice guidelines for the diagnosis and management of otitis media with effusion (OME) in children in Japan, 2015. Auris Nasus Larynx. 2017;44(5):501-8.
2. Wallace IF, Berkman ND, Lohr KN, Harrison MF, Kimple AJ, Steiner MJ. Surgical treatments for otitis media with effusion: a systematic review. Pediatrics. 2014;133(2):296-311.
3. Escamilla Y, Aguilà AF, Saiz JM, Rosell R, Vivancos J, Cardesín A. Tympanostomy tube emplacement in children with secretory otitis media: analysis of effects and complications. Acta Otorrinolaringologica (English Edition). 2009;60(2):84-9.
4. Rosenfeld RM, Schwartz SR, Pynnonen MA, Tunkel DE, Hussey HM, Fichera JS, et al. Clinical practice guideline: Tympanostomy tubes in children. Otolaryngol Head Neck Surg. 2013;149 (1 Suppl):S1-35.
5. Rosenfeld RM, Shin JJ, Schwartz SR, Coggins R, Gagnon L, Hackell JM, et al. Clinical Practice Guideline: Otitis Media with Effusion (Update). Otolaryngol Head Neck Surg. 2016;154(1 Suppl):S1-S41.
6. Ungkanont K. Long-term Outcome of the Management of Otitis Media with Effusion in Children with and Without Cleft Palate Using the House-brand Polyethylene Ventilation Tube Insertion. Siriraj Med J. 2021;73(4):245-51.
7. Suvarnsit K, Chantarawiwat T, Prakairungthong S, Limviriyakul S, Atipas S, Pitathawatchai P. Granular Myringitis Treatment at

Siriraj Hospital. Siriraj Med J. 2020;72(6):502-7.
8. Hong P, Webb AN, Corsten G, Balderston J, Haworth R, Ritchie K, et al. An anatomically sound surgical simulation model for myringotomy and tympanostomy tube insertion. Int J Pediatr Otorhinolaryngol. 2014;78(3):522-9.
9. Molin N, Chiu J, Liba B, Isaacson G. Low cost, easy-to-replicate myringotomy tube insertion simulation model. Int J Pediatr Otorhinolaryngol. 2020;131(109847):1-4.
10. Wiet GJ, Deutsch ES, Malekzadeh S, Onwuka AJ, Callender NW, Seidman MD, et al. SimTube: A National Simulation Training and Research Project. Otolaryngol Head Neck Surg. 2020;163(3):522-30.
11. Barber SR, Kozin ED, Dedmon M, Lin BM, Lee K, Sinha S, et al. 3D-printed pediatric endoscopic ear surgery simulator for surgical training. Int J Pediatr Otorhinolaryngol. 2016;90:113-8.
12. Kuru I, Maier H, Muller M, Lenarz T, Lueth TC. A 3D-printed functioning anatomical human middle ear model. Hear Res. 2016;340:204-13.
13. Sparks D, Kavanagh KR, Vargas JA, Valdez TA. 3D printed myringotomy and tube simulation as an introduction to otolaryngology for medical students. Int J Pediatr Otorhinolaryngol. 2020;128(109730):1-4.
14. Volsky PG, Hughley BB, Peirce SM, Kesser BW. Construct validity of a simulator for myringotomy with ventilation tube insertion. Otolaryngol Head Neck Surg. 2009;141(5):603-8 e1.
15. Hwang JJ, Jung YH, Cho BH. The need for DICOM encapsulation of 3D scanning STL data. Imaging Sci Dent. 2018;48(4):301-2.
16. Kamio T, Suzuki M, Asaumi R, Kawai T. DICOM segmentation and STL creation for 3D printing: a process and software package comparison for osseous anatomy. 3D Print Med. 2020;6(1):1-12.
17. Edelmers E, Kazoka D, Pilmane M. Creation of Anatomically Correct and Optimized for 3D Printing Human Bones Models. Applied System Innovation. 2021;4(3):1-14.
18. Kumar Sharma R, Dhiman S, Negi S. Basics and applications of rapid prototyping medical models. Rapid Prototyping Journal. 2014;20(3):256-67.
19. Rengier F, Mehndiratta A, von Tengg-Kobligk H, Zechmann CM, Unterhinninghofen R, Kauczor HU, et al. 3D printing based on imaging data: review of medical applications. Int J Comput Assist Radiol Surg. 2010;5(4):335-41.
20. Ho AK, Alsaffar H, Doyle PC, Ladak HM, Agrawal SK. Virtual reality myringotomy simulation with real-time deformation: development and validity testing. Laryngoscope. 2012;122(8):1844-51.
21. Huang C, Cheng H, Bureau Y, Agrawal SK, Ladak HM. Face and content validity of a virtual-reality simulator for myringotomy with tube placement. J Otolaryngol Head Neck Surg. 2015;44(40):1-8.
22. Huang C, Cheng H, Bureau Y, Ladak HM, Agrawal SK. Automated Metrics in a Virtual-Reality Myringotomy Simulator: Development and Construct Validity. Otol Neurotol. 2018;39(7):e601-e8.
23. Vaitaitis VJ, Dunham ME, Kwon YC, Mayer WC, Evans AK, Baker AJ, et al. A Surgical Simulator for Tympanostomy Tube Insertion Incorporating Capacitive Sensing Technology to Track Instrument Placement. Otolaryngol Head Neck Surg. 2020;162(3):343-5.

Survival Analysis of and Prognostic Factors for Metastatic Epidural Spinal Cord Compression Compared between Preoperative Known and Unknown Primary Tumors

Nitiwut Saenmanot, M.D.*, Monchai Ruangchainikom, M.D.*, Thanase Ariyawatkul, M.D.*, Ekkapoj Korwutthikulrangsri, M.D.*, Soraya Saenmanot, Ph.D., Panya Luksanapruksa, M.D.*, Werasak Sutipornpalangkul, M.D.*, Sirichai Wilartratsami, M.D.*, Chatupon Chotigavanichaya, M.D.***

**Department of Orthopedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand, **Faculty of Public Health, Mahidol University Amnatcharoen Campus, Amnatcharoen, Thailand.*

ABSTRACT

Objective: To analyze the median survival time of and prognostic factors for metastatic epidural spinal cord compression (MESCC) secondary to preoperative unknown primary tumor (pre-op UPT) compared to MESCC secondary to preoperative known primary tumor (pre-op KPT).

Materials and Methods: This retrospective cohort study reviewed all consecutive MESCC patients who underwent surgical decompression with or without stabilization within 72 hours of admission during 2010 to 2016. Survival was compared between the pre-op UPT and pre-op KPT groups, and preoperative and postoperative prognostic factors for survival were analyzed.

Results: A total of 169 patients (pre-op UPT: 51, and pre-op KPT: 118) were enrolled. The survival rate at 3, 6, and 12 months was 84.3%, 58.8%, and 47.1% in the pre-op UPT group, and 72.0%, 48.3%, and 34.7% in the pre-op KPT group, respectively. The median survival time secondary to lung cancer was significantly longer in the pre-op UPT group (6.0±1.4 months) than in the pre-op KPT group (3.6±0.2 months) (p=0.031). Multivariate analysis revealed survival time to be influenced by preoperative known or unknown primary tumor status, revised Tokuhashi score, the adjuvant therapy, and postoperative complications, including myocardial infarction, gastrointestinal bleeding, and urinary tract infection.

Conclusion: MESCC secondary to preoperative unknown primary tumor patients who had the clinical presentation with acute progressive neurological deficits who need urgency spine surgery has comparable survival to MESCC secondary to preoperative known primary tumors.

Keywords: Metastasis; spinal cord compression; unknown primary tumor; survival time; prognosis; urgency decompression (Siriraj Med J 2022; 74: 684-692)

INTRODUCTION

Cancer is one of the leading causes of death.[1] Spinal metastasis was found in seventy percent of cancer death undergoing autopsy, and 10% of spinal metastasis patients developed neurological deficits.[2,3]

Metastatic epidural spinal cord compression (MESCC) patients who can realize potential benefit from surgical decompression should be urgently treated to improve patient functional status, mental status and to prevent persistent loss of motor and/or sensory function.[4,5] Surgical

Corresponding author: Monchai Ruangchainikom
E-mail: monchai.ortho@gmail.com
Received 30 April 2022 Revised 12 September 2022 Accepted 13 September 2022
ORCID ID:http://orcid.org/0000-0003-0525-6390
http://dx.doi.org/10.33192/Smj.2022.80

management is usually considered in patients who have a life expectancy of greater than 3 months.[6]

Survival time prediction has important clinical implications, including decision-making relative to the potential benefit of surgical treatment or palliative cancer therapies. There are many scales/scoring systems for predicting the prognosis of MESCC patients, especially Tomita[7] and revised Tokuhashi[8], which are the most widely accepted scoring systems. These assessment tools are useful when standard diagnostic strategy is complete, to include tumor marker, Tc-99m bone scan, positron emission tomography (PET) scan, chest-abdomen-pelvis computed tomography (CT) scan, magnetic resonance imaging (MRI) of the spine with contrast, and tissue biopsy. However, these investigations take time that may delay proper management of spinal metastasis patients who have acute neurological deficit as the first clinical presentation. These patients whose primary tumor was initially not definitively known were described as MESCC with preoperative unknown primary tumor (MESCC with pre-op UPT).

Urgent surgical decompression is quite common in MESCC patients with pre-op UPT. To the best of our knowledge, there is no study that has compared survival time between MESCC with pre-op UPT and MESCC with the preoperative known primary tumor (MESCC with pre-op KPT) in this urgent neurological compromise setting.

Thus, survival analysis of all consecutive MESCC with pre-op UPT or pre-op KPT who underwent surgical decompression was the main objective of this study. The secondly aim was to identify prognostic factors associated with median survival time in both groups.

MATERIALS AND METHODS

This retrospective cohort study reviewed all 193 consecutive MESCC patients who underwent surgical decompression with or without stabilization within 72 hours of admission to Siriraj Hospital during 2010 to 2016. This study was approved by Siriraj Institutional Review Board (SIRB) of the Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand (Si 013/2016).

To be eligible for inclusion, MESCC patients who underwent surgical decompression must have had all of the following components of preoperative management: 1) complete medical history and physical examination; 2) standard laboratory analysis, including tumor markers; 3) plain radiography of involved bone and chest X-ray; and, 4) MRI spine with contrast medium, with imaging of the whole spine the sagittal view. MESCC patients with incomplete data or who died from a non-cancer-related cause were excluded.

Patient survival time was investigated by telephone call and a review of medical records. Preoperative and postoperative assessment parameters were reviewed and recorded, as follows: general demographic data, smoking or nonsmoking, American Spinal Injury Association (ASIA) impairment scale score, site of pathologic spinal level, number of spinal metastases, pre-op KPT, pre-op UPT, final identified primary tumor site, revised Tokuhashi score after complete investigation, adjuvant therapy, operative time and operative complications.

Statistical analysis

Chi-square test was used to compare categorical variables (results shown as number and percentage), and Student's *t*-test was used for continuous variables (results shown as mean plus/minus standard deviation). Survival analyses were performed by Kaplan-Meier method, with subsequent group comparison by log-rank test. Prognostic factors associated with survival time were identified by Cox proportional hazards model. A p-value of < 0.05 was considered statistically significant. SPSS for Windows version 18.0 was used for all statistical analyses.

RESULTS

Participants

Of the 193 MESCC patients that were evaluated for eligibility, 24 were excluded. Of those, 19 were excluded for having incomplete patient data or because we could not determine their postoperative status. Another 5 cases were excluded because they died of a non-cancer-related cause (3 suicides, 1 trauma, and 1 murder). The remaining 169 MESCC patients were included in our final analysis (Fig 1).

Descriptive data

The baseline characteristics of the 169 enrolled MESCC patients (101 males, 68 females) were evaluated and compared between the pre-op UPT and pre-op KPT groups. The mean age of study patients was 54.9±12.8 years. Fifty-one cases (30.2%) with preoperative unknown primary tumor site were identified. All MESCC patients presented with incomplete spinal cord lesion, and most (59.2%) had ASIA impairment scale grade D. Almost three-quarters (69.2%) of MESCC patients who underwent urgent spinal decompression presented with more than one level of spinal metastasis, and the most commonly affected spinal region was thoracic spine (50.3%). Regarding the postoperative revised Tokuhashi score after complete

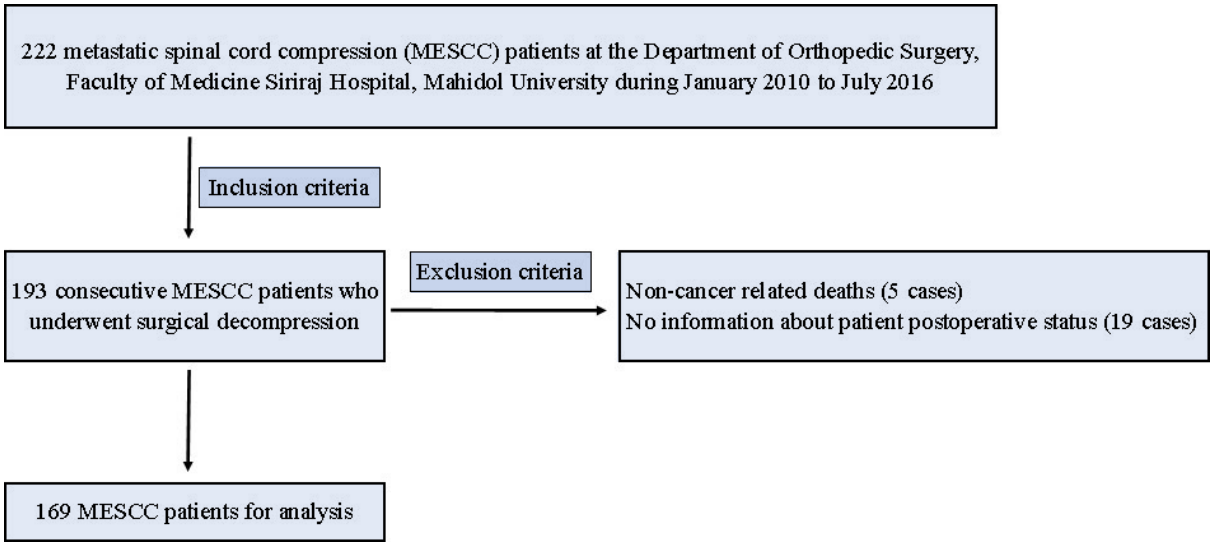


Fig 1. Flow chart describing the patient enrollment process.

investigation according to the standard diagnostic strategy, most patients (55.6%) had a score that fell into the 0-8 group.

The distribution of identified primary tumor site in the pre-op UPT and pre-op KPT groups is shown in Table 2. The most common identified primary tumor site was the lung in both groups. From histological study, eight of nine cases in the cancer of unknown primary site (CUP) patients were adenocarcinoma, and the other was undifferentiated carcinoma.

Survival after surgery

The median survival time between groups was not significantly different, but there was a trend towards longer survival time in the pre-op UPT group (8.4 months, 95% CI: 0.8-16.1) than in the pre-op KPT group (5.1 months, 95% CI: 3.5-6.8) (p=0.127). Kaplan-Meier survival method and log-rank test were used to estimate survival and compare the results between groups (Fig 2). The survival rate at 3, 6, and 12 months was 84.3%, 58.8%, and 47.1% in the pre-op UPT group, and 72.0%, 48.3%, and 34.7% in the pre-op KPT group, respectively.

Concerning the lung being the most common primary tumor site in both groups, the median survival time secondary to lung cancer was 3.6 months (95% CI: 3.2-4.0) in the pre-op KPT group, and 6.0 months (95% CI: 3.3-8.7) in the pre-op UPT group (p=0.031) (Fig 3).

Prognostic factors associated with survival time

Univariate analysis (Table 3) showed the following prognostic factors to be significantly associated with survival time: American Spinal Injury Association (ASIA) Impairment Scale at presentation, patient smoking status, number of levels of spinal involvement, revised Tokuhashi score after inclusion of all standard diagnostic data (especially identification of the type of metastatic tumor from histologic finding), adjuvant treatment after surgical intervention, and postoperative complications, including cerebral infarction (stroke), myocardial infarction (MI), gastrointestinal bleeding (GI bleeding), urinary tract infection (UTI), and pressure ulcer.

Multivariate analysis (Table 4) revealed preoperative primary tumor of known or unknown status, adjuvant therapy, revised Tokuhashi score after collection of all standard diagnostic data, and the postoperative complications MI, GI bleeding, and UTI to be independent prognostic factors associated with survival.

DISCUSSION

Interest in the predicted survival time of MESCC patients has increased over the last few years because it is one of the most important factors for guiding decision-making in MESCC patients relative to whether patients with neurological deficit that require urgent care should be given palliative care or operative management. In this study, the prevalence of MESCC patients who presented with acute progressive neurological deficit that indicated for spinal decompression was common with secondary to primary unknown tumor. Moreover, we observed a comparable between median survival time in the pre-op UPT group and in the pre-op KPT group. Regarding lung as the primary tumor site, which was the most common primary tumor site in the pre-op UPT group (37.3%), we found a significantly longer median survival time in the pre-op UPT group than in the pre-op KPT group. This is an interesting finding, especially since patients with

TABLE 1. Demographic and clinical characteristics compared between the preoperative unknown primary tumor site group (Pre-op UPT) and the known primary tumor site group (Pre-op KPT).

Characteristics	Pre-op UPT (n=51)	Pre-op KPT (n=118)	*P*-value
Gender, n (%)			
Male	39 (76.5%)	62 (52.5%)	0.003
Female	12 (23.5%)	56 (47.5%)	
Age			
Mean age (±SD)	54.9±11.5	54.9±13.3	0.986
ASIA impairment scale			
B	9 (17.7%)	19 (16.1%)	0.418
C	9 (17.7%)	32 (27.1%)	
D	33 (64.7%)	67 (56.8%)	
Pathologic spinal level			
Cervical spine	7 (13.7%)	14 (11.9%)	0.917
Thoracic spine	27 (52.9%)	58 (49.2%)	
T-L junction (T12-L1)	9 (17.7%)	24 (20.3%)	
L2-3 or cord level	8 (15.7%)	22 (18.6%)	
Revised Tokuhashi score			
Score 0-8	28 (54.9%)	66 (55.9%)	0.116
Score 9-11	20 (39.2%)	33 (28.0%)	
Score 12-15	3 (5.9%)	19 (16.1%)	
Number of levels of spinal metastasis			
1 level	19 (37.3%)	33 (28.0%)	0.296
2 levels	15 (29.4%)	31(26.3%)	
≥3 levels	17 (33.3%)	54 (45.8%)	
Adjuvant therapy			
None	7 (13.7%)	11 (9.3%)	0.092
Chemotherapy (CMT)	2 (3.9%)	3 (2.5%)	
Radiotherapy (RT)	26 (51.0%)	42 (35.6%)	
CMT and RT	16 (31.4%)	62 (52.5%)	

Abbreviation: ASIA indicated American Spinal Injury Association

the lung as the primary tumor site are given the lowest score of 0, the poorest prognosis primary tumor origin category, when using the Tokuhashi scoring system.[8]

MESCC with preoperative unknown primary tumor with acute progressive neurological deficits presents a major decision-making challenge for a spine surgeon, and it makes it difficult for the surgeon to offer the patient accurate information specific to prognosis, survival, and management. These unknowns can lead to surgeon reluctance to perform urgent spinal decompression and stabilization, but a failure to do so can lead to adverse outcomes for the patient compared to those who receive urgent intervention. The general recommendation is to perform surgery only in patients with a life expectancy of greater than 3 months, and the common survival prediction scoring systems are commonly based on primary tumor, performance status, number of levels of spine metastasis, neurological status, and presence of visceral metastases.[6-9] A recent guideline from the Netherlands Comprehensive Cancer Organization relative

TABLE 2. Distribution of identified primary tumor sites compared between the pre-op UPT and pre-op KPT groups.

Identified primary tumor site	Pre-op KPT (n=118)	Pre-op UPT (n=51)
CA Lung	28 (23.7%)	19 (37.3%)
CA Breast	28 (23.7%)	1 (2.0%)
CA Prostate	10 (8.5%)	5 (9.8%)
Hematologic malignancy	10 (8.5%)	9 (17.6%)
CA Nasopharynx	8 (6.8%)	1 (2.0%)
CA Liver	7 (5.9%)	4 (7.8%)
CA Cervix	7 (5.9%)	0 (0.0%)
CA Colon	6 (5.1%)	0 (0.0%)
CA Thyroid	4 (3.4%)	1 (2.0%)
CA Kidney	2 (1.7%)	0 (0.0%)
CA Bladder	1 (0.8%)	0 (0.0%)
Others	7 (5.9%)	2 (3.9%)
Cancer of unknown primary site (CUP)	-	9 (17.6%)

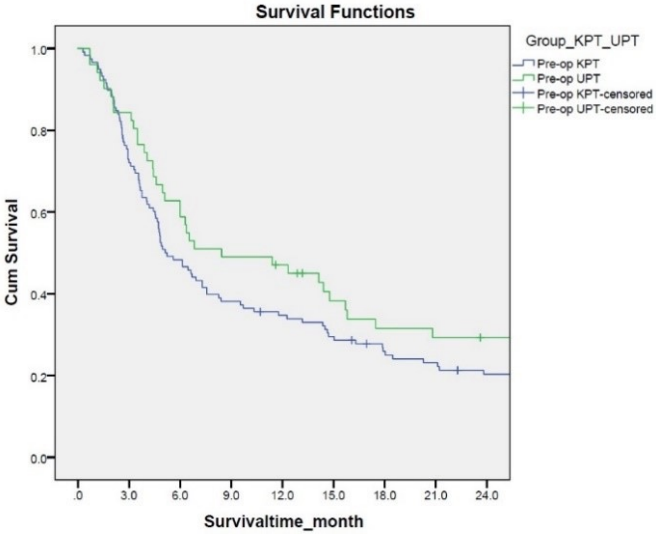


Fig 2. Kaplan-Meier survival graphs compared between pre-op unknown primary tumor site group (Pre-op UPT, green line) and the pre-op known primary tumor site group (Pre-op KPT, blue line).

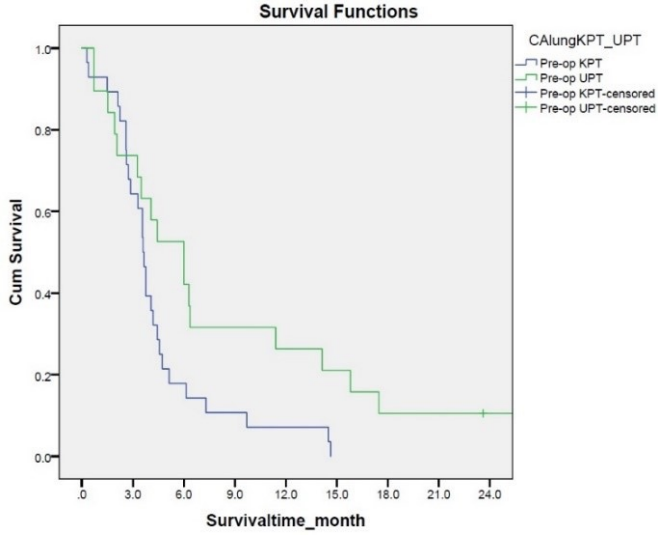


Fig 3. Kaplan-Meier survival graphs compared between those with pre-op unknown primary tumor site with post-op determination of primary lung cancer (lung CA) (Pre-op UPT, green line) and those with pre-op primary tumor site known to be the lung (Pre-op KPT, blue line).

TABLE 3. Univariate analysis for prognostic factors that predict survival time in MESCC patients who underwent surgical decompression within 72 hours of hospital admission.

Variables	n	Univariate Analysis		
		HR	95% CI	*P*-value
Preoperative primary tumor				
known	118	1.334	0.920 – 1.933	0.129
unknown	51	1		
Gender: males/females	101/68	1.088	0.777 – 1.523	0.625
Age				
17 – 40 years	26	1.641	0.871 – 3.091	0.125
41 – 50 years	34	1.105	0.595 – 2.054	0.751
51 – 60 years	50	1.206	0.678 – 2.146	0.523
60 – 70 years	39	0.864	0.468 – 1.594	0.640
> 70 years	20	1		
ASIA impairment scale				
B	28	1.998	1.276 – 3.129	*0.002**
C	41	1.108	0.737 – 1.664	0.622
D	100	1		
Adjuvant therapy				
None	18	2.344	1.345 – 4.086	*0.003**
Chemotherapy(CMT)	5	1.186	0.371 – 3.787	0.774
Radiotherapy(RT)	68	1.092	0.763 – 1.563	0.630
Combined CMT and RT	78	1		
Smoker (+/-)	43/126	1.471	1.010 – 2.142	*0.044**
Pathologic level				
Cervical spine	30	1	0.888 – 2.720	0.122
Thoracic spine	21	1.554	0.670 – 2.424	0.460
T-L junction (T12-L1)	85	1.274	0.718 – 2.532	0.353
L2-3 or cord level	33	1.3481		
Spinal related symptom period				
< 1 week	101	0.992	0.450 – 2.190	0.985
1 – 4 weeks	61	0.812	0.362 – 1.820	0.613
>4 weeks	7	1		
Morbidity stage				
ASA 1	7	1		
ASA 2	85	0.990	0.430 – 2.281	0.982
ASA 3	75	1.233	0.532 – 2.858	0.625
ASA 4	2	2.2461	0.449 – 11.233	0.325
Group Tokuhashi score				
Score 0-8	94	5.308	2.909 – 9.686	< 0.001*
Score 9-11	53	1.623	0.862 – 3.053	0.133
Score 12-15	22	1		

TABLE 3. Univariate analysis for prognostic factors that predict survival time in MESCC patients who underwent surgical decompression within 72 hours of hospital admission. (Continued)

Variables	n	Univariate Analysis		
		HR	95% CI	*P*-value
Number of spinal level involvement				
1 level	52	1		
2 levels	46	1.652	1.063 – 2.567	0.025*
≥ 3 levels	71	1.713	1.141 – 2.572	0.009*
Group operative time (hours)				
< 3 hours	12	1		
3 – 4 hours	60	0.758	0.391 – 1.467	0.410
4 – 5 hours	54	1.028	0.531 – 1.992	0.934
5 – 6 hours	26	1.042	0.501 – 2.170	0.912
> 6 hours	17	1.074	0.491 – 2.345	0.859
Post-op complications				
Cerebral infarction (+/-)	3/166	3.754	1.172 – 12.028	0.026*
Myocardial infarction (+/-)	11/158	2.307	1.209 – 4.402	0.011*
Pneumonia (+/-)	42/127	2.803	1.925 – 4.081	<0.001*
Gastrointestinal bleeding (+/-)	11/158	3.973	2.108 – 7.489	<0.001*
Urinary tract infection (+/-)	68/101	1.650	1.177 – 2.313	0.004*
Thromboembolism (+/-)	18/151	1.590	0.942 – 2.686	0.083
Pressure ulcer (+/-)	54/115	1.521	1.072 – 2.157	0.019*

Abbreviations: ASIA; indicated American Spinal Injury Association, ASA; American Society of Anesthesiologists grade of physical status

to MESCC secondary to preoperative unknown primary tumor recommends that, if it is possible, MRI of the whole spine and PET-CT of the thorax/abdomen should be performed, and that tissue biopsy should be obtained within 1 day.[10] An attempt to complete all recommended investigations and tissue biopsy would delay the critical time needed for recovery of injured neural tissue. The results of our study revealed a survival time of greater than 3 months in 72.0%-84.3% of MESCC patients. Yalamanchili, *et al*. found that rapid progression is common in patients who present with neurological deficit. They found that 30% of patients with weakness could progress to paraplegia within 1 week, and that the likelihood of regaining neurological function was very poor when paraplegia was present for more than 24 hours.[11] Preserving the remaining functional neural tissue, increasing the chance of neural tissue injury recovery, and improving ambulatory status all play an important role in patient survival and quality of life and mental status.[5] A meta-analysis of spinal metastasis by Luksanapruksa, *et al*. found neurological deficits and ambulatory status to be commonly reported prognostic factors for survival.[12-15]

Cancer of unknown primary site (CUP) was not uncommon (17.6%) in this study, and the most common histopathologic finding was adenocarcinoma. This is similar to previous studies that reported a prevalence of CUP in patients with MESCC of 13.4%-14.5%, and the tissue pathology was usually adenocarcinoma.[16,17] CUP is usually associated with more aggressive behavior and shorter survival time, which are derived from both biologic condition, such as prior immunoediting and/or featuring a high degree of immunosuppression, and lack of a specific guideline for clinical management.[17,18]

Our univariate analysis revealed several potential prognostic factors significantly associated with survival time. Subsequent multivariate analysis that included those factors revealed independent prognostic factors for survival time. Among those, we found preoperative primary tumor of known or unknown status to play an important role

TABLE 4. Multivariate analysis for prognostic factors that predict survival time in MESCC patients who underwent surgical decompression within 72 hours of hospital admission.

Variables	n	Multivariate analysis		
		HR	95% CI	*P*-value
Pre-op primary tumor				
known	118	1.657	1.060 – 2.590	*0.027**
unknown	51	1		
ASIA impairment scale				
B	28	0.992	0.583 – 1.689	0.978
C	41	0.752	0.472 – 1.197	0.229
D	100	1		
Adjuvant therapy				
None	18	3.359	1.760 – 6.414	*<0.001**
Chemotherapy(CMT)	5	2.401	0.691 – 8.346	0.168
Radiotherapy(RT)	68	0.944	0.616 – 1.447	0.792
Combined CMT and RT	78	1		
Smoker (+/-)	43/126	1.435	0.930 – 2.214	0.103
Pathologic level				
Cervical spine	30	1		
Thoracic spine	21	1.855	0.978 – 3.518	0.058
T-L junction (T12-L1)	85	0.892	0.443 – 1.798	0.750
L2-3 or cord level	33	1.987	0.983 – 4.017	0.056
Group Tokuhashi score				
Score 0-8	94	6.854	3.351 – 14.016	<0.001*
Score 9-11	53	1.892	0.973 – 3.679	0.060
Score 12-15	22	1		
Number of spinal level involvement				
1 level	52	1		
2 levels	46	1.259	0.731 – 2.169	0.407
≥ 3 levels	71	0.937	0.542 – 1.622	0.817
Post-op complications				
Cerebral infarction (+/-)	3/166	1.427	0.354 – 5.752	0.618
Myocardial infarction (+/-)	11/158	3.104	1.450 – 6.648	0.004*
Pneumonia (+/-)	42/127	1.402	0.870 – 2.260	0.165
GI bleeding (+/-)	11/158	3.565	1.702 – 7.465	0.001*
Urinary tract infection (+/-)	68/101	1.519	1.018 – 2.265	0.041*
Thromboembolism (+/-)	18/151	0.960	0.498 – 1.850	0.904
Pressure ulcer (+/-)	54/115	1.262	0.824 – 1.935	0.285

Abbreviation: ASIA indicated American Spinal Injury Association

in survival time. We also found independent association between the revised Tokuhashi score and survival. Similar to other studies[16,17], adjuvant therapy was identified as an important prognostic factor for survival; however, differences in survival were reported among different adjuvant treatments and different types of tumors.[19-21] Lastly, we also found postoperative complications, including myocardial infarction, gastrointestinal bleeding, and urinary tract infection, to be independent prognostic factors to decreased life expectancy. A systematic review by Bakar, *et al*. found a high prevalence of various types of postoperative complications in MESCC patients that ranged in prevalence from 5% to 42.6%.[22]

Limitations

This study has some mentionable limitations. First, our study's retrospective design suggests the potential for missing or incomplete data. However, we endeavored to exclude all cases with incomplete data. Second, the small number of each identified primary tumor type except for lung cancer means that Kaplan-Meier survival analysis could only be performed for MESCC secondary to lung cancer. Third and last, the findings of this study could not analyze the decision making for adjuvant treatments in various types of primary tumor and ununiformed optimal chemotherapy in the period of this study with rapid development of chemotherapy.

CONCLUSION

Survival time of MESCC patients who had the clinical presentation with acute progressive neurological deficits depends on multiple prognostic factors, however; the preoperative unknown primary tumor origin is not negative factor for survival in palliative spine surgery.

ACKNOWLEDGEMENTS

We thank the personnel from the Siriraj Hospital record and research section for their assistance, the Department of Orthopedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand for their support, and Mr. Pongsathorn Samphaotong for assisting in journal submission process.

Conflict of interest: The authors declare that there is no conflict of interest.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69(1):7-34.
2. Rose PS, Buchowski JM. Metastatic disease in the thoracic and lumbar spine: evaluation and management. J Am Acad Orthop Surg 2011;19(1):37-48.
3. Jacobs WB, Perrin RG. Evaluation and treatment of spinal metastases: an overview. Neurosurg Focus 2001;11(6):e10.
4. Furstenberg CH, Wiedenhofer B, Gerner HJ, Putz C. The effect of early surgical treatment on recovery in patients with metastatic compression of the spinal cord. J Bone Joint Surg Br 2009;91(2):240-4.
5. Chavasiri C, Sukprasert N, Chavasiri S. Depression, Social Support, and Coping Stratergies in Individuals with Spinal Injury Depression with Spinal Injury Patients. Siriraj Med J 2021;73(8): 518-25.
6. Metastatic Spinal Cord Compression: Diagnosis and Management of Patients at Risk of or with Metastatic Spinal Cord Compression. Cardiff (UK); 2008.
7. Tomita K, Kawahara N, Kobayashi T, Yoshida A, Murakami H, Akamaru T. Surgical strategy for spinal metastases. Spine (Phila Pa 1976) 2001;26(3):298-306.
8. Tokuhashi Y, Matsuzaki H, Oda H, Oshima M, Ryu J. A revised scoring system for preoperative evaluation of metastatic spine tumor prognosis. Spine (Phila Pa 1976) 2005;30(19):2186-91.
9. Tokuhashi Y, Uei H, Oshima M, Ajiro Y. Scoring system for prediction of metastatic spine tumor prognosis. World J Orthop 2014;5(3):262-71.
10. Bollen L, Dijkstra SPD, Bartels R, Graeff A, Poelma DLH, Brouwer T, et al. Clinical management of spinal metastases-The Dutch national guideline. Eur J Cancer 2018;104:81-90.
11. Yalamanchili M, Lesser GJ. Malignant spinal cord compression. Curr Treat Options Oncol 2003;4(6):509-16.
12. Hosono N, Ueda T, Tamura D, Aoki Y, Yoshikawa H. Prognostic relevance of clinical symptoms in patients with spinal metastases. Clin Orthop Relat Res 2005;(436):196-201.
13. Pointillart V, Vital JM, Salmi R, Diallo A, Quan GM. Survival prognostic factors and clinical outcomes in patients with spinal metastases. J Cancer Res Clin Oncol 2011;137(5):849-56.
14. Tatsui CE, Suki D, Rao G, Kim SS, Salaskar A, Hatiboglu MA, et al. Factors affecting survival in 267 consecutive patients undergoing surgery for spinal metastasis from renal cell carcinoma. J Neurosurg Spine 2014;20(1):108-16.
15. Luksanapruksa P, Buchowski JM, Hotchkiss W, Tongsai S, Wilartratsami S, Chotivichit A. Prognostic factors in patients with spinal metastasis: a systematic review and meta-analysis. Spine J 2017;17(5):689-708.
16. Paholpak P, Sirichativapee W, Wisanuyotin T, Kosuwon W, Jeeravipoolvarn P. Prevalence of known and unknown primary tumor sites in spinal metastasis patients. Open Orthop J 2012; 6:440-4.
17. Wanman J, Grabowski P, Nystrom H, Gustafsson P, Bergh A, Widmark A, et al. Metastatic spinal cord compression as the first sign of malignancy. Acta Orthop 2017;88(4):457-62.
18. Aizenberg MR, Fox BD, Suki D, McCutcheon IE, Rao G, Rhines LD. Surgical management of unknown primary tumors metastatic to the spine. J Neurosurg Spine 2012;16(1):86-92.
19. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ, Wu Y-L, et al. Lung cancer: current therapies and new targeted treatments. Lancet 2017;389(10066):299-311.
20. Teo MY, Rathkopf DE, Kantoff P. Treatment of Advanced Prostate Cancer. Annu Rev Med 2019; 70: 479-99.
21. Sartor O, de Bono JS. Metastatic Prostate Cancer. N Engl J Med 2018;378(7):645-57.
22. Bakar D, Tanenbaum JE, Phan K, Alentado VJ, Steinmetz MP, Benzel EC, et al. Decompression surgery for spinal metastases: a systematic review. Neurosurg Focus 2016;41(2):E2.

Types and Levels of Colostomy in Children with Anorectal Malformation

Ravit Ruangtrakool, M.D., Cholapa Pintawekiat, M.D.

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

ABSTRACT

Objective: Divided colostomy for anorectal management is often recommended due to reports of higher complications associated with loop colostomy. This study was conducted to compare outcomes and complications in colostomies in children with anorectal malformations according to type and level of colostomy.

Materials and Methods: A retrospective study was performed in children with anorectal malformations who underwent a colostomy at Siriraj Hospital between December 2003 and June 2018.

Results: Out of 167 patients, 159 had a loop colostomy while 8 had a divided colostomy. Overall complication rates were 33.3% for loop colostomy and 62.5% for divided colostomy (p = 0.100). Urinary tract infection was the most frequently encountered complication in both loop and divided colostomies, at 23.7% and 50%, respectively (p = 0.094). The prolapse rate in the loop colostomy group was 8.8 % and 0% in the divided colostomy group (p = 0.376). Overall complication rates with respect to location of stoma also did not differ (p = 0.706). Prolapse rates were 15.8 % in transverse colostomy and 7.1 % in sigmoid colostomy (p = 0.231). Overall complications rates of colostomy closure in loop and divided colostomy was 7.5% and 12.5%, respectively (p = 0.672). Non-inferiority was demonstrated by the differences in overall complications of loop and divided colostomy (p = 0.008).

Conclusion: There was no difference in incidence of complications between type or location of colostomy performed in children with anorectal malformations. Loop colostomy was non-inferior to divided colostomy in respect to overall complications.

Keywords: Anorectal malformation; loop colostomy; divided colostomy; colostomy prolapse; urinary tract infection; complication (Siriraj Med J 2022; 74: 693-698)

INTRODUCTION

Among all congenital anomalies, the gastrointestinal anomaly was the second most common system involved (33.67%).[1] Anorectal malformations are a common gastrointestinal anomaly encountered by pediatric surgeons worldwide. There is a wide spectrum of malformations, ranging from simple cutaneous fistula to cloacal malformations. Colostomy, with subsequent definite repair is the standard treatment in those with non-low type anorectal malformation. Loop colostomy was the only preferred option in Division of Pediatric Surgery at Siriraj Hospital for more than five decades, until divided colostomy was firstly introduced by Peña A, who developed posterior sagittal anorectoplasty, the most popular definite operation for anorectal malformation in 1982.[2-5] Divided colostomy is generally preferred over loop colostomy due to the higher rate of complications associated with the latter, which includes prolapse, risk of incomplete diversion of feces that causes subsequent distension of distal rectal pouch, and possible contamination

Corresponding author: Ravit Ruangtrakool
E-mail: sisuped@mahidol.ac.th
Received 14 February 2022 Revised 6 July 2022 Accepted 31 July 2022
ORCID ID:http://orcid.org/0000-0001-8162-2941
http://dx.doi.org/10.33192/Smj.2022.81

of feces into the urinary tract in those with a connecting fistula between the rectum and genitourinary system.[2-6] However, controversy still surrounds the higher risk of complications associated with loop colostomies compared to divided colostomies with previously published studies showing debatable results.[7-9] Thus, this study was conducted to describe and compare outcomes and complications of colostomies in children with anorectal malformations according to type and level of colostomy. Outcomes related to colostomy closure with respect to type of colostomy were also compared.

MATERIALS AND METHODS

Following approval by the Siriraj Institutional Review Board (Si 175/2019) a retrospective study was conducted in children with anorectal malformations who underwent a colostomy at Siriraj Hospital between December 2003 to June 2018. Children with cloacal exstrophy and major chromosomal anomalies incompatible with life and those with incomplete medical information were excluded from the study. Patients' demographics, type of malformation, location and type of colostomy was collected. First, a colostomy was performed and this was followed by definitive repair. Following achieving an adequate neo-anus size as dilated by the parents, colostomy closure was performed. Loop colostomy was the preferred option in our division at Siriraj Hospital. Complications during colostomy were recorded, including prolapse, retraction, parastomal hernia, urinary tract infection, bleeding, and skin excoriation. Upon colostomy closure, operative time and complications were noted. Complications during colostomy closure included wound infection, wound dehiscence, and anastomosis leakage. The collected data was analyzed using SPSS software version 18 (SPSS Inc. Released 2009. PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc.). Continuous data was expressed as median and IQR and categorical data expressed as numbers and percentages. A Chi square test or Fisher's exact test was used to compare outcomes in type and location of colostomy. Non-inferiority test for difference in overall complications between loop and divided colostomy was conducted. Non-inferiority was demonstrated when lower bound of the 95% one-sided CI for difference in overall complications was lower than pre-specified non-inferior margin of 10%. A p-value of <0.05 indicated statistical significance.

RESULTS

Of the 178 patients whose medical records were reviewed, 11 were excluded, which meant 167 patients were included in the study. Out of the 167 patients included, 159 underwent loop colostomies while eight had a divided colostomy for fecal diversion. One hundred and four out of 159 participants were male, of which 98 had a loop colostomy. Sixty-three patients were female, of which 61 underwent a loop colostomy. For colostomy level, there were four locations in total; ascending colon, transverse colon, descending colon, and sigmoid colon. The majority of patients underwent a (134 out of 167) sigmoid loop colostomy. There was a wide distribution of malformation types ranging from imperforate anus without fistula to complex defects without significant differences between the loop and divided colostomy group (Table 1).

The differences in complications found in the loop and divided colostomy groups were not significant (Table 2). Overall complication rates were 33.3% in the loop colostomy group and 62.5% in the divided colostomy group ($p = 0.100$). Urinary tract infections were the most frequently observed complications in both the loop (23.3%) and divided (50%) colostomy group. The prolapse rate was 8.8% in the loop colostomy group and 0% in the divided colostomy group, while skin excoriation was 6.3% in the loop colostomy group and 12.5% in the divided colostomy group. When comparing complications according to colostomy location, there were no difference in overall or individual complications (Table 3). Interestingly, no statistically significant difference was noted in overall complications or prolapse rates between the transverse and sigmoid colostomy.

The median operative time for colostomy closures was 160 minutes for loop colostomy and 195 minutes for divided colostomy. The difference was not statistically significant ($p = 0.128$). The incidence of complications such as wound infection, wound dehiscence, and leakage following closures in loop colostomy and divided colostomy were not statistically significant (Table 4).

When statistics for non-inferiority were performed, with pre-specified non inferior margin of 10% between loop and divided colostomies, non-inferiority was demonstrated as the difference in overall complications ($p = 0.008$).

DISCUSSION

Colostomy with subsequent definite repair is the standard treatment in people with non-low type malformation. Divided colostomies have been proposed over loop colostomies due to reports of an increase in complications associated with loop colostomy.[2-6] In fact, loop colostomy was condemned by Pena A,[2,4,5] a world authority in anorectal malformation management, due to an increased prolapse rate, risk of incomplete diversion of feces causing subsequent distension of distal rectal pouch,

TABLE 1. Comparison of the patient characteristics between the loop colostomy and the divided colostomy.

Variable	Loop Colostomy (n = 159)	Divided Colostomy (n = 8)	Total (n = 167)	*P*-value
Gender, n (%)				0.711
Male	98 (61.6%)	6 (75.0%)	104 (62.3%)	
Female	61 (38.4%)	2 (25.0%)	63 (37.7%)	
Level of Colostomy, n (%)				0.137
Ascending	3 (1.9%)	0 (0%)	3 (1.8%)	
Transverse	18 (11.3%)	1 (12.5%)	19 (11.4%)	
Descending	2 (1.3%)	1 (12.5%)	3 (1.8%)	
Sigmoid	134 (84.3%)	6 (75.0%)	140 (83.8%)	
Type of Malformation, n (%)				0.819
Imperforate Anus Without Fistula	30 (18.9%)	3 (37.5%)	33 (19.8%)	
Perineal Fistula	10 (6.3%)	0 (0%)	10 (6.0%)	
Vestibular Fistula	14 (8.8%)	0 (0%)	14 (8.4%)	
Rectovaginal Fistula	6 (3.8%)	0 (0%)	6 (3.6%)	
Rectobulbar Urethral Fistula	26 (16.4%)	0 (0%)	26 (15.6%)	
Rectoprostatic Urethral Fistula	19 (11.9%)	1 (12.5%)	20 (12.0%)	
Rectobladder Neck Fistula	11 (6.9%)	1 (12.5%)	12 (7.2%)	
Rectovesicle Fistula	9 (5.7%)	0 (0%)	9 (5.4%)	
Persistent Cloaca < 3 Cm	13 (8.6%)	1 (12.5%)	14 (8.4%)	
Persistent Cloaca > 3 Cm	7 (4.4%)	0 (0%)	7 (4.2%)	
Rectal Atresia	3 (1.9%)	0 (0%)	3 (1.8%)	
Complex Defect	3 (1.9%)	0 (0%)	3 (1.8%)	

TABLE 2. Complications from colostomy, comparing between loop colostomy and divided colostomy.

	Loop Colostomy (n = 159)	Divided Colostomy (n = 8)	Total (n = 167)	*P*-value
Overall Complication, n (%)	53 (33.3%)	5 (62.5%)	58 (34.7%)	0.100
Prolapse	14 (8.8%)	0 (0%)	14 (8.4%)	0.376
Retraction	2 (1.3%)	0 (0%)	2 (1.2%)	0.747
Parastomal Hernia	1 (0.6%)	0 (0%)	1 (0.6%)	0.820
Urinary Tract Infection	37 (23.3%)	4 (50.0%)	41 (24.6%)	0.094
Bleeding	5 (3.1%)	0 (0%)	5 (3.0%)	0.607
Skin Excoriation	10 (6.3%)	1 (12.5%)	11 (6.6%)	0.502

TABLE 3. Complications from colostomy, comparing among different sites.

Variables	Ascending colostomy (n = 3)	Transverse colostomy (n = 19)	Descending colostomy (n = 3)	Sigmoid colostomy (n = 140)	*P*-value
Overall Complication, n (%)	2 (66.7%)	6 (31.6%)	1 (33.3%)	50 (35.7%)	0.706
Prolapse	1 (33.3%)	3 (15.8%)	0 (0%)	10 (7.1%)	0.231
Retraction	0 (0%)	0 (0%)	0 (0%)	2 (1.4%)	0.948
Parastomal Hernia	0 (0%)	0 (0%)	0 (0%)	1 (0.7%)	0.981
Urinary Tract Infection	2 (66.7%)	5 (26.3%)	1 (33.3%)	34 (24.3%)	0.409
Bleeding	0 (0%)	0 (0%)	0 (0%)	5 (3.6%)	0.820
Skin Excoriation	1 (33.3%)	0 (0%)	0 (0%)	10 (7.1%)	0.168

TABLE 4. Outcomes at colostomy closure, comparing between loop colostomy and divided colostomy.

Variable	Loop Colostomy (n = 159)	Divided Colostomy (n = 8)	Total (n = 167)	*P*-value
Operative time (min)				0.128
Median (min. max)	160 (35, 457)	195 (120, 215)	160 (35, 157)	
Complications, n (%)	12 (7.5%)	1 (12.5%)	13 (7.7%)	0.672
Wound Infection	8 (5.0%)	1 (12.5%)	9 (5.39%)	0.410
Wound Dehiscence	1 (0.6%)	0 (0%)	1 (0.6%)	0.814
Leakage	1 (0.6%)	0 (0%)	1 (0.6%)	0.814
Gut Obstruction	3 (1.9%)	0 (0%)	3 (1.8%)	0.682
Incisional Hernia	1 (0.6%)	0 (0%)	1 (0.6%)	0.814

TABLE 5. Overall complications in loop and divided colostomy.

	Loop colostomy (n=159)	Divided colostomy (n=8)	Difference (95% one-sided CI)	Non-inferiority test (*P*-value)
Overall complications	53 (33.3%)	5 (62.5%)	-28.5% (-1.7, ∞)*	0.008**

*Non-inferiority was demonstrated (lower bound of the 95% one-sided CI for difference in overall complications between loop and divided colostomy was lower than pre-specified non-inferiority margin of 10%)

**Non-inferiority was demonstrated and *p*-value of non-inferiority test was less than significant level of 0.05

and possible contamination of feces into the urinary tract in patients with a connecting fistula between the rectum and genitourinary system.[2-5] However, in our study, there was no statistically significant difference between the loop and divided colostomy group regarding overall complications. The complication rate associated with loop colostomies in this study was consistent with previously published studies at about 20%-30%.[3,7,9] However, this study elicited a higher rate of complications in divided colostomies compared to other studies (62.5% vs. 8%-30%).[3,7,9]

The prolapse rate of loop colostomies was quite low in our study (8.8%) compared to other published studies which reported rates of up to 18%.[3,9] A low prolapse rate in loop colostomies in our institution might be the result of the stoma creation technique used at our center where loop colostomy was performed at the descending-sigmoid colonic junction and the proximal and distal limb of colostomy site were sutured together prior to exteriorization and fixation at sheath and skin. The suturing of the proximal and distal limb may have decreased the mobility of colon.

Divided colostomies were preferred over loop colostomies due to risk of fecal contamination into the distal rectourinary fistula in the latter.[2,4] Although urinary tract infection was the most common complication, there was no significant difference between the loop and divided colostomy group in this study. This finding was consistent with previously published studies.[3,7] A loop colostomy conducted in the proper way was able to complete fecal diversion and was not different from divided colostomy.

Regarding location of colostomy, no statistically significant difference in complication rates was elicited in this study. However, our study had higher rates of overall complications for both transverse and sigmoid colostomies compared to results published by van den Hondel et al[9] and Demirogullari et al.[10] This might be the result of including urinary tract infection as a complication in our study while other studies did not include it. Previous literatures have revealed that transverse colostomies have a higher prolapse rate than other colostomy locations.[9-11] Regarding transverse colostomies, our study had a lower prolapse rate than others. As mentioned previously, this might be due to our surgical technique of placing sutures between the proximal and distal limb of colon prior to exteriorization of stoma at sheath and skin. Also, we had more experience performing a loop colostomy regardless of location when compared to other studies.

Since there were no significant differences in complications in loop and divided colostomies, we attempted to determine whether loop colostomy was non-inferior compared to divided colostomy in respect to complication rates. The non-inferior margin was pre-determined to be 10%. Interestingly, we found that non-inferiority, which was shown as a *p*-value in the non-inferiority test, to be less than the significant level of 0.05. This had not been shown in previously published studies.

Since there was no difference in complications between loop and divided colostomies during the stoma creation period, outcomes during and after colostomy closure were investigated to demonstrate the advantage of one stoma over the other. The operative time for colostomy closure seemed shorter for loop colostomy at 160 minutes compared to 195 minutes for divided colostomy. However, there was no significant difference to suggest easier closure in loop colostomy. Complications such as wound infection, wound dehiscence and anastomotic leakage were also not significantly different.

The limitation of this study was its retrospective design which means some information might be missing. Moreover, the number of subjects was relatively small at 167 patients. Also, there were a smaller number of divided colostomy patients compared to the loop colostomy group as it is our division's preference to perform the latter. This made comparison between the two groups difficult in the study. However, our results were similar to previously published studies in which loop colostomy had good results compared to divided colostomy. A multicenter study may be performed in the future to increase the number of patients and data of divided colostomy cases.

CONCLUSION

Loop colostomy is non-inferior to divided colostomy in terms of overall complications and is a feasible diversion procedure for anorectal malformation. Proper technique and experience with loop colostomy helps achieve complete diversion of feces with outcomes similar to that of divided colostomy.

ACKNOWLEDGEMENTS

We would like to thank Dr. Sasima Tongsai from the Division of Clinical Epidemiology, Department of Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University for her continuous help with data processing and statistical analysis.

Conflicts of interest: The authors have no conflicts of interest to declare.

REFERENCES

1. Sahoo S, Ganguly R, Dash M, Pradhan A, Priya TG, Mohanty. Study of Congenital Malformation in a Tertiary Care Teaching Hospital. Siriraj Med J. 2021;73(9):609-13.
2. Pena A, Migotto-Krieger M, Levitt MA. Colostomy in anorectal malformations: a procedure with serious but preventable complications. J Pediatr Surg. 2006;41(4):748-56.
3. Oda O, Davies D, Colapinto K, Gerstle JT. Loop versus divided colostomy for the management of anorectal malformations. J Pediatr Surg. 2014;49(1):87-90.
4. Levitt MA, Peña A. Imperforate anus and cloacal malformations. In: Holcomb III GW, Murphy JA, ed. Ashcraft's Pediatric Surgery, 5th edition, Philadelphia: Saunders Elsevier; 2010.p.468-90.
5. Wilkins S, Peña A. The role of colostomy in the management of anorectal malformations. Pediatr Surg Int. 1988;3:105-9.
6. Gardikis S, Antypas S, Mamoulakis C, Demetriades D, Dolatzas T, Tsalkidis A, et al. Colostomy type in anorectal malformations: 10-years experience. Minerva Pediatr 2004; 56(4):425-9.
7. Liechty ST, Barnhart DC, Huber JT, Zobell S, Rollins MD. The morbidity of a divided stoma compared to a loop colostomy in patients with anorectal malformation. J Pediatr Surg. 2016; 51(1):107-10
8. Patwardhan N, Kiely EM, Drake DP, Spitz L, Pierro A. Colostomy for anorectal anomalies: high incidence of complications. J Pediatr Surg. 2001;36(5):795-8.
9. van den Hondel D, Sloots C, Meeussen C, Wijnen R. To split or not to split: colostomy complications for anorectal malformations or hirschsprung disease: a single center experience and a systematic review of the literature. Eur J Pediatr Surg. 2014;24(1): 61-9.
10. Demirogullari B, Yilmaz Y, Yildiz GE, Ozen IO, Karabulut R, Turkyilmaz Z, et al. Ostomy complications in patients with anorectal malformations. Pediatr Surg Int. 2011;27(10):1075-8.
11. Almosallam OI, Aseeri A, Shanafey SA. Outcome of loop versus divided colostomy in the management of anorectal malformations. Ann Saudi Med. 2016;4:352-5.

Lidocaine Reducing Pain from Benzathine Penicillin Injection: A Controlled Trial

Sukhum Jiamton, M.D., Ph.D.*, Nuttagarn Jantanapornchai, M.D*, Ya-Nin Nokdhes, M.D.*, Poramin Patthamalai, M.D., Ph.D.*, Ploypailin Tantrapornpong, M.D.*, Yanisorn Nanchaipruek, M.D. *, Pichaya Limpoka, M.D.*

**Sexually Transmitted Disease and HIV Division, Department of Dermatology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Objective: This study aimed to investigate the efficacy and safety of benzathine penicillin G (BPG) injection compared between dilution with 1% lidocaine hydrochloride and dilution with sterile water to reduce pain in Thai male syphilis patients.

Materials and Methods: This randomized, split-buttock, double-blind controlled trial was conducted at the Sexually Transmitted Disease and HIV Division, Department of Dermatology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand during September 2018 to July 2019. After randomization, 20 left and 20 right buttocks were injected with 1.2 million-unit BPG (half dose) with lidocaine as the diluent. The other 20 left and 20 right buttocks were then injected with 1.2 million-unit BPG (half dose) with sterile water as the diluent. Pain at each buttock was measured by numeric rating scale during and immediately after the injection, and at 5 minutes, 20 minutes, and 24 hours after injection.

Results: Forty males (mean age: 30.6±10.3 years) were included. Compared to sterile water diluent, we found that dilution with 1% lidocaine significantly reduced pain during and immediately after injection, and at 5-minutes and 20-minutes post-injection (all $p<0.001$). There was no significant improvement in pain at 24-hours post-injection. Minor adverse events were observed in 37.5% of patients, including generalized rash, pruritus, and fever. One patient experienced minor drug allergy.

Conclusion: One percent lidocaine as a diluent of BPG was found to be effective for reducing pain during and after BPG injection.

Keywords: Efficacy; 1% lidocaine hydrochloride; sterile water; pain; benzathine penicillin G injection; syphilis (Siriraj Med J 2022; 74: 699-704)

INTRODUCTION

Syphilis is a bacterial infectious disease that is caused by *Treponema pallidum* subspecies *pallidum*.[1] Effective treatment is intramuscular (IM) penicillin G for all stages of the disease. In adults, one IM injection of benzathine penicillin G (BPG) 2.4 million units is recommended for early syphilis treatment, including primary, secondary, and early latent syphilis. Three doses (once a week for 3 consecutive weeks) of BPG 2.4 million units is suggested for late latent and tertiary syphilis without neurosyphilis treatment.[1]

BPG that is diluted with sterile water causes pain during and after injection at the injection sites. The local pain and discomfort associated with the injection tend to decrease compliance, especially in children and adolescents.[2,3] To reduce pain and improve patient compliance in those

Corresponding author: Sukhum Jiamton
E-mail: sukhum.jia@mahidol.ac.th
Received 14 March 2022 Revised 11 September 2022 Accepted 13 September 2022
ORCID ID:http://orcid.org/0000-0003-1068-1586
http://dx.doi.org/10.33192/Smj.2022.82

requiring 3 injections, we set forth to investigate the efficacy and safety of BPG injection compared between dilution with 1% lidocaine hydrochloride and dilution with sterile water to reduce pain in Thai male syphilis patients.

MATERIALS AND METHODS

This randomized, split-buttock, double-blind controlled trial in syphilis patients aged 18 years or older was conducted at the Sexually Transmitted Disease and HIV Division of the Department of Dermatology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand during September 2018 to July 2019. Patients with allergy to lidocaine or penicillin were excluded. The protocol for this study was approved by the Siriraj Institutional Review Board (SIRB) (Si 635/2017), and written informed consent was obtained from all study participants.

Both buttocks of each patient were randomized using block randomization with a block size of four. It was used to determine which side of the buttock was intervention or control. Thus, for instance, BPG diluted with lidocaine mean intervention or BPG diluted with sterile water mean control (Fig 1). The dilution was 1.2 million-unit BPG whether with 4.0 ml of lidocaine or with 4.0 ml of sterile water. One dermatologist prepared medication for administration. However, only the first dose of injection was examined in those required three doses of injection.

A 20-gauge 1.5-inch-long needle was used for injection. One blinded dermatologist performed drug injection at each buttock intramuscularly. All participants were blinded. Pain was assessed using a numeric rating scale, with a zero indicating no pain, and a ten indicating the most severe pain. Pain was assessed during and immediately after the injection, and at 5 minutes, 20 minutes, and 24 hours after the injection. Adverse effects and events were recorded. The same blinded dermatologist assessed pain score and adverse effect.

Statistical methods

The data were analyzed using Statistical Package for the Social Sciences (SPSS, Inc., Chicago, IL, USA) version 18. Categorical data are reported as number and percentage, and continuous data are reported as mean plus/minus standard deviation. Paired *t*-test was used to compare the pain score between the two diluent formulations at during and immediately after the injection, and at 5 minutes, 20 minutes, and 24 hours after the injection. A *p*-value less than 0.05 was considered statistically significant for all tests.

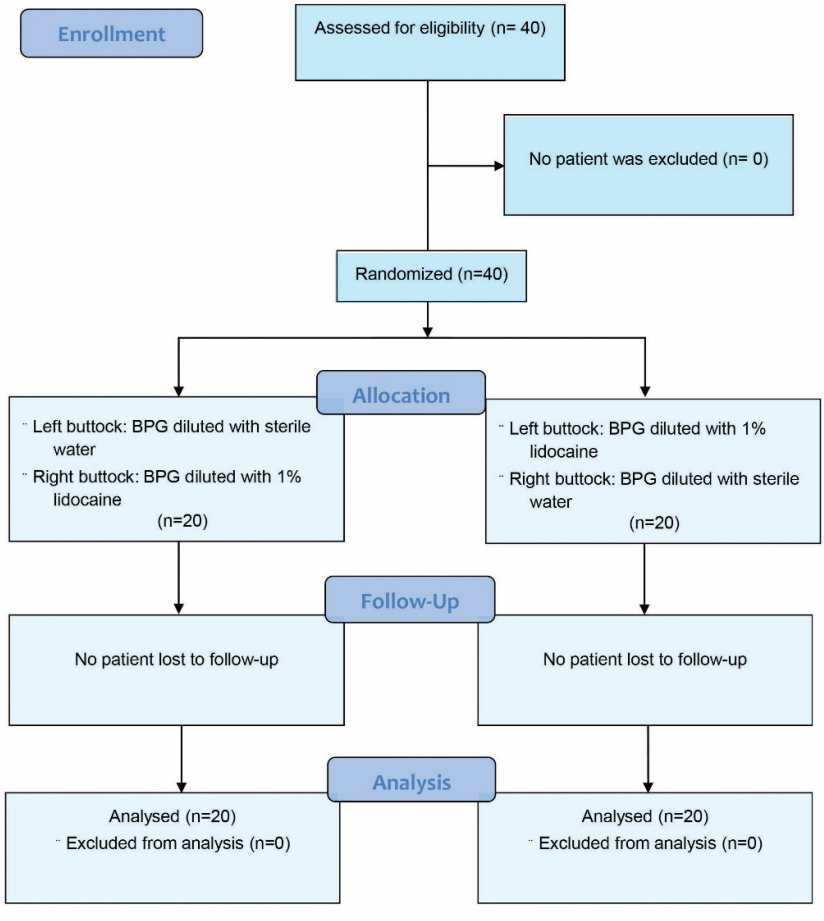


Fig 1. Flow diagram of the randomized clinical trial of benzathine penicillin G (BPG).

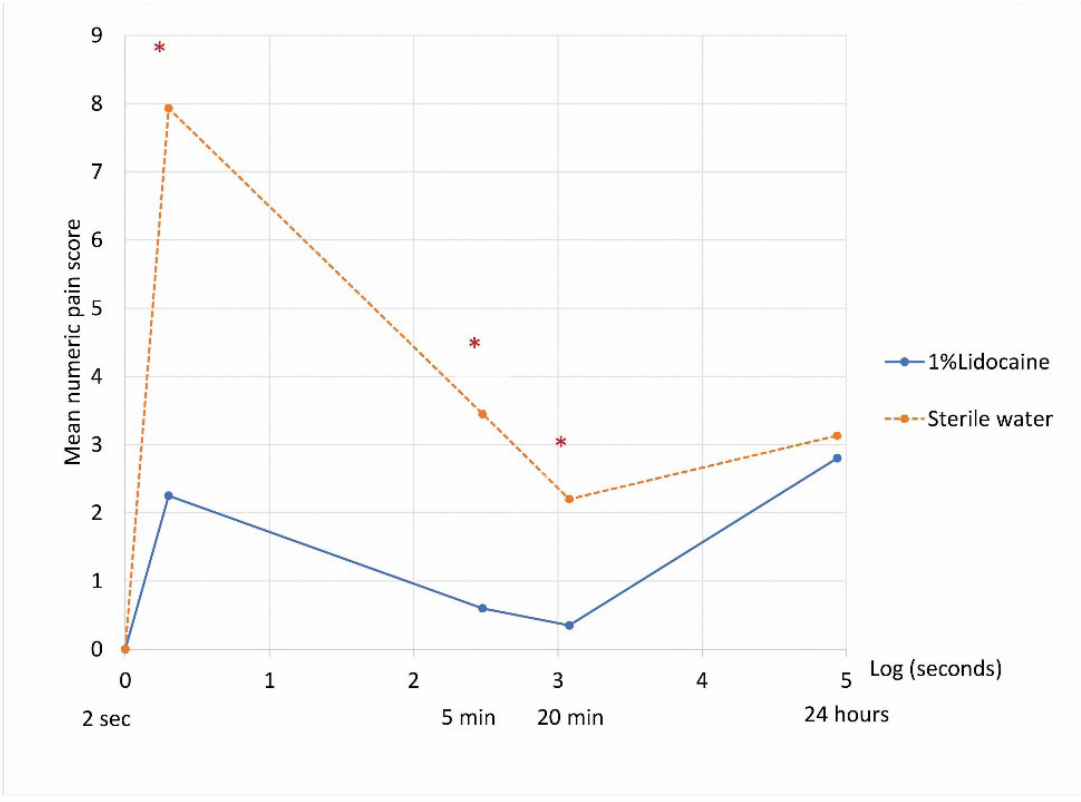


Fig 2. Mean numeric pain score over time compared between injection of BPG diluted with 1% lidocaine and injection of BPG diluted with sterile water.

RESULTS

Forty Thai males that were recruited from our outpatient dermatology clinic were included. The age of patients ranged from 18 to 59 years, and the mean age was 30.6±10.3 years. Half of patients reported being homosexual, and 15% were bisexual. No patients reported that they always use a condom during intercourse. Fifty-five and forty percent of patients were in the late latent stage or secondary stage of syphilis, respectively. Thirty-three of 37 patients had a Venereal Disease Research Laboratory (VDRL) titer higher than 1:16, and all patients had a positive Treponema pallidum particle agglutination assay (TPHA) result (Table 1). Compared to sterile water diluent, we found that dilution with 1% lidocaine significantly reduced pain during and immediately after injection, and at 5-minutes and 20-minutes post-injection (all $p<0.001$) (Table 2). There was no significant improvement in pain at 24-hours post-injection. Minor adverse events were observed in 37.5% of patients, including generalized rash, pruritus, and fever. One patient experienced minor drug allergy (Table 3). However, there was no local adverse event so the local adverse event between study and control group could not done.

DISCUSSION

Our study showed that BPG injection using 1% lidocaine as the diluent significantly reduces pain at injection sites during and immediately after the injection, and at 5 minutes and 20 minutes after the injection. Interestingly, however, the pain score in the lidocaine group increased by the 24-hour time point to very near the pain score in the sterile water group. This study also found intramuscular injection with BPG to be associated with two distinct episodes of pain during and immediately after injection, and at 24 hours after injection. This reemergence of pain after a decrease in pain at 5-minutes and 20-minutes post-injection is likely due to the fact that the elimination half-life of lidocaine is biphasic within a range of 90-120 minutes in most patients.[4]

Pain during intramuscular injection is a problem in clinical practice.[5] Reducing intramuscular injection pain may be achieved by combining drugs with local anesthetics, such as lidocaine.[6] Lidocaine is a sodium-channel blocking drug with rapid onset of action, and it has a minimal toxicity profile.[2,7] Four milliliters of 1% lidocaine was used in this study for a total dose of 2 mg. This is less than the maximum dose of lidocaine without

TABLE 1. Demographic and clinical characteristics of the 40 included syphilis patients.

Characteristics	n	%
Age (years)		
Mean ± SD	30.6 ± 10.3	-
Range	18-59	-
Sexual orientation		
Homosexual	20	50.0
Heterosexual	14	35.0
Bisexual	6	15.0
Frequency of condom use		
Always	0	0.0
Sometimes	37	92.5
Never	3	7.5
Underlying medical condition		
Yes	7	17.5
No	33	82.5
HIV infection	23	57.5
Stage of syphilis		
Primary syphilis	0	0.0
Secondary syphilis	16	40.0
Early latent syphilis	2	5.0
Late latent syphilis	22	55.0
Tertiary Syphilis	0	0.0
Presence history of known contact disease	11	30.8
Recurrent syphilis	9	19.2
TPHA reactive at diagnosis (titer)	38	100.0
>1:80	38	100.0
VDRL reactive at diagnosis (titer)	37	92.5
1:1	2	5.4
1:2	1	2.7
1:4	1	2.7
1:16	3	8.1
1:32	7	18.9
1:64	8	21.6
1:128	11	29.7
1:256	4	10.8

TABLE 2. Mean pain visual analogue scale (VAS) at different time points compared between injection of benzathine penicillin G diluted with 1% lidocaine and injection of benzathine penicillin G diluted with sterile water.

	1% Lidocaine side (mean±SD)	Sterile water side (mean±SD)	*P*-value
During and immediately after injection	2.25±1.85	7.93±1.95	**1.45×10^{-18}***
5 minutes after injection	0.60±1.43	3.45±2.64	**4.55×10^{-8}***
20 minutes after injection	0.35±1.21	2.20±2.15	**7.12×10^{-6}***
24 hours after injection	2.80±1.94	3.13±2.08	0.156

*A *p*-value<0.05 indicates statistical significance

TABLE 3. Adverse events and complication at 24 hours after treatment among the 40 included syphilis patients.

Adverse events and complication	n	(%)
Adverse events		
No	25	62.5%
Yes	15	37.5%
Generalized rash	9	60.0%
Generalized pruritus	8	53.3%
Fever	8	53.3%
Complication		
No	39	97.5%
Yes	1	2.5%
Minor drug allergy	1	100.0%

epinephrine, which was reported to be 300 mg.[8] All patients in this study were observed for lidocaine toxicity, and no signs of toxicity were observed. Previous studies reported no adverse pharmacokinetic effects of lidocaine on IM penicillin.[2,9] It was also reported that the BPG combined with lidocaine did not change the concentration of penicillin in body fluid.[2] No local adverse event was observed in our study. The adverse events in our study were generalized symptoms such as rash, pruritus, and fever. Because participants received both interventions and all of them had no history of allergy from penicillin[10] or lidocaine[11], so these recorded generalized symptoms might be from penicillin or lidocaine.

CONCLUSION

One percent lidocaine as a diluent of BPG was found to be effective for significantly reducing pain during and immediately after BPG injection, and at 5-minutes and 20-minutes post-injection. We recommend the use of 1% lidocaine as a diluent for BPG in this clinical setting.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the patients that generously agreed to participate in this study, and the statistician for assistance with statistical analysis.

Funding disclosure: This study was funded by a grant from the Siriraj Research Development Fund (Managed by Routine to Research: R2R) (COA no. [IO] R01613501).

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

REFERENCES

1. Workowski KA, Bachmann LH, Chan PA, Johnston CM, Muzny CA, Park I, et al. Sexually Transmitted Infections Treatment Guidelines, 2021. MMWR Recomm Rep. 2021;70(4):1-187.
2. Amir J, Ginat S, Cohen YH, Marcus TE, Keller N, Varsano I. Lidocaine as a diluent for administration of benzathine penicillin G. Pediatr Infect Dis J. 1998;17(10):890-3.
3. Matthapan L, Bunyaratavej S, Leeyaphan C, Rujitharanawong C, Limphoka P, Prasong W, et al. Clinical Application of Toothbrush Technique for Specimen Collection of Tinea Capitis. Siriraj Med J [Internet]. 2019 Mar 28. Available from: https://he02.tcithaijo.org/index.php/sirirajmedj/article/view/180764
4. Thomson PD, Melmon KL, Richardson JA, Cohn K, Steinbrunn W, Cudihee R, et al. Lidocaine Pharmacokinetics in Advanced Heart Failure, Liver Disease, and Renal Failure in Humans. Ann Intern Med. 1973;78(4):499-508.

5. Morsy MMF, Mohamed MA, Abosedira MM, Al-Harbi K, Abdelaziz NA, Hal SQK, et al. Lidocaine as a diluent for benzathine penicillin G reduces injection pain in patients with rheumatic fever: A prospective, randomized double-blinded crossover study. Australian Journal of Basic and Applied Sciences. 2012;6:236-40.
6. Russell K, Nicholson R, Naidu R. Reducing the pain of intramuscular benzathine penicillin injections in the rheumatic fever population of Counties Manukau District Health Board. J Paediatr Child Health. 2014;50(2):112-7.
7. Razavi BM, Fazly Bazzaz BS. A review and new insights to antimicrobial action of local anesthetics. Eur J Clin Microbiol Infect Dis. 2019;38(6):991-1002.
8. Physicians' desk reference. Montvale, N.J. Medical Economics Co., 1996.
9. Schichor A, Bernstein B, Weinerman H, Fitzgerald J, Yordan E, Schechter N. Lidocaine as a diluent for ceftriaxone in the treatment of gonorrhea. Does it reduce the pain of the injection? Arch Pediatr Adolesc Med. 1994;148(1):72-5.
10. Yip DW, Gerriets V. Penicillin. [Updated 2021 Sep 30]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK554560/
11. Cherobin ACFP, Tavares GT. Safety of local anesthetics. Anais Brasileiros de Dermatologia 2020;95:82-90.

An Overview on Postoperative Cognitive Dysfunction; Pathophysiology, Risk Factors, Prevention and Treatment

Thanathip Suenghataiphorn, M.D.*, Sakdipat Songwisit, M.D.*, Surapa Tornsatitkul, M.D., Pawit Somnuke, M.D., Ph.D.*****

**Faculty of Medicine Siriraj Hospital, **Department of Pharmacology, Faculty of Pharmacy, ***Department of Anesthesiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.*

ABSTRACT

Postoperative cognitive dysfunction (POCD) is an event that alarms medical personnel owing to its adverse effects, including heightened morbidity and mortality rates, prolonged recovery times, and increased lengths of hospital stay and healthcare expenditure. The populations at high risk are elderly, critical patients, or complicated cases that need prolonged surgery in which the hemodynamics are not stable. Although guidelines have been established to facilitate the early diagnosis of POCD, its prevention is recommended for good patient outcomes. A preoperative assessment is a prerequisite for patient optimization before surgery. Intraoperative, enhanced-recovery protocols have been widely adopted to promote recovery following surgery. Frequent, postoperative assessments of patients' vital signs and cognitive functions are required for early POCD detection. Patients diagnosed with POCD need regular follow-up, and proper patient counselling is paramount.

Keywords: Diagnostic and Statistical Manual of Mental Disorders (DSM-5); Enhanced Recovery After Surgery (ERAS); Montreal Cognitive Assessment Scale (MoCA); Neurocognitive disorder, Postoperative cognitive dysfunction (POCD) (Siriraj Med J 2022; 74: 705-713)

INTRODUCTION

Postoperative cognitive dysfunction (POCD) is a condition that can occur during the postoperative or postanesthetic periods.[1] Previous study reported that upon discharge, 41.4% of patients aged over 60 years developed POCD and, notably, up to 12.7% of those patients were detected with POCD at 3 months.[1] Morbidity of the patients with POCD at 3 months and 1 year after surgery was 25.8% and 10%, respectively.[2] The etiologies can be categorized into (1) patient factors: age > 60 years old, low education level, American Society of Anesthesiologists (ASA) physical status ≥ 3, and comorbidities, for example, cerebrovascular disease, anemia (preoperative hemoglobin ≤ 11 g/dl), preexisting cognitive dysfunction, poor functional capacity, severe illness, postoperative respiratory complications and postoperative infection (2) surgical factors: complex and complicated surgeries or complications during surgery, type of surgery and time of surgery ≥ 4 hours and (3) anesthetic factors: intraoperative use of benzodiazepines and Isoflurane volatile anesthetic agent, intraoperative hypotension and oxygen desaturation during anesthetic induction.[3-7] POCD contributes to declined general health, longer length of hospitalization, longer length

Corresponding author: Pawit Somnuke
E-mail: Pawit-pup@hotmail.com
Received 5 June 2022 Revised 5 July 2022 Accepted 6 July 2022
ORCID ID:http://orcid.org/0000-0002-6773-1882
http://dx.doi.org/10.33192/Smj.2022.83

of postsurgical recovery, and an increase in 1-year post-surgical mortality rate.[8,9] The prevention of POCD requires cooperation between medical specialties throughout the preoperative, perioperative and postoperative periods.[9] The best strategy for combatting POCD is to prevent before it occurs. It is paramount that medical personnel have a thorough understanding about POCD so that they are capable of planning how to prevent POCD and also to enhance patients' postoperative recovery and quality of life.

This review discusses and summarizes the details of POCD by using evidence-based medicine covering POCD definitions, clinical symptoms, diagnosis, pathophysiology, risk factors, prevention, treatment, and prognosis.

Definition

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) has yet to issue a formal definition of POCD. However, the International Society of Postoperative Cognitive Dysfunction (ISPOCD) defines it as a condition that can develop when ≥ 1 abnormality in a discrete area of mental state such as memory, consciousness or attention is discovered, which can occur anytime from immediate postoperative period to 6 months later. Usually, the onset of impaired memory and intellectual disability ranges from weeks to months postoperatively while recovery is within days to weeks.[1,10] POCD can be diagnosed by comparing the differences in the results of baseline preoperative and postoperative psychometric testing.[9]

In 2018, the new consensus among international medical doctors and scientists was published in the British Journal of Anesthesia.[11] The clearer definition of POCD was introduced to facilitate research and education endeavors: (1) delayed neurocognitive recovery (within 30 days postoperatively) and (2) postoperative neurocognitive disorder (between 30 days and 12 months postoperatively). POCD can be differentiated from other diseases, such as delirium or dementia, as outlined in Table 1.[1,12-14]

Pathophysiology

Although the pathophysiology of POCD is still not elucidated, it is, by evidence-based, associated with neuroinflammation, disruption of blood-brain barrier (BBB) integrity, neurosynaptic damage, mitochondrial dysfunction and oxidative stress.[15-19] Other mechanisms including hyperventilation, hypotension or cerebral microemboli were also proposed to involve POCD.[10] Surgical stimuli induce the expression of the inflammatory mediator, high mobility group box-1 (HMGB1), which interacts with the pattern recognition receptor (PRR) on macrophages causing downstream activation of innate immunity. Also, the S100 calcium-binding protein A12 (S100A12) level increases after the operation and thus resulting in intracephalic signal transduction and inflammation.[20] Upon surgical stimulation, intracellular RNA released from the damaged tissue is detected by the immune cells therefore inflammatory process is initiated. Proinflammatory cytokines from macrophages including tumor necrosis factor alpha (TNF-α), interleukin-6 (IL-6)

TABLE 1. Differential diagnoses of postoperative cognitive dysfunction.

Parameters	Delirium	POCD	Dementia
Onset	Within 3 days	Within a few months	Months to years
Duration	Days to weeks	Weeks to months	Months till death
Attention	Decreased	Decreased	Decreased
Consciousness	Altered	Normal	Normal
Symptoms	Fluctuation within the day; alteration of consciousness; can be hypoactive or hyperactive	Memory decline or cognitive decline	Memory decline; executive function decline; changes in behavior and abilities Increased risk of functional decline
Activities of daily living	Increased risk of functional decline	No risk of functional decline	

Abbreviation: POCD, Postoperative cognitive dysfunction

and IL-1β are upregulated in the blood circulation.[21-24] These cytokines can breach through the BBB via vagus nerve or paraventricular areas of the BBB leading to activation of cerebrovascular endothelial cells which will produce secondary messenger to secrete more proinflammatory cytokines.[25] An increase in brain-specific inflammatory markers such as serum S100 calcium-binding protein B (S100B) and neuron-specific enolase (NSE) after surgery also supports that brain inflammation could lead to POCD.[10,12,26,27] Albeit playing smaller role than surgery, anesthesia is involved in the pathophysiology of POCD. Past study demonstrated that isoflurane caused apoptosis in human neuronal cell lines and mouse brain potentially through the accumulation of amyloid β peptide. Randomized controlled studies also reported higher incidence of POCD and level of proinflammatory markers in groups anesthetized under volatile anesthesia implying that volatile might involve in POCD mechanism.[28,29]

Risk factors

The risk factors of POCD can be divided into patient, surgical, and anesthetic factors.

1. Patient factors

Old age[1,30]

Age is a major factor of POCD especially in the elderly aged > 60 years old. Studies have shown that older age has various effects on the brain, for instance, decreased brain volume, decreased BBB density, decreased neurogenesis, reduced cognitive reserve, increased brain inflammation, and increased brain-vessel degeneration. The medial temporal lobe atrophy as well as the white matter hyperintensity as seen by magnetic resonance imaging (MRI) in the elderly were well correlated with clinical cognitive decline.[31]

Low level of education[9,32]

Many higher-educated people are prone to engage in greater levels of complex thinking, leading to heightened usage of the brain neural network. This extra utilization may result in the prevention of brain decay due to a corresponding increase in the cognitive reserve and improvements to the efficiency of neuronal replacement. Education levels may therefore be employed to indicate cognitive reserves as each additional year of study has been demonstrated to result in around a 10% reduction in the incidence of POCD.

Preexisting cerebrovascular disease[1,33]

The patients with preexisting cerebral infarction were reported to be 18.2% at risk of POCD compare with 4.9% in the control group. Therefore, cerebrovascular disease was considered the potential non-modifiable risk factors of POCD.

Preexisting Systemic Lupus Erythematosus (SLE)[34]

A correlation between cognitive impairment and underlying SLE had been reported previously. However, the incidence of cognitive dysfunction was not dependent on SLE duration, activity or evidence of preexisting neuropsychiatric involvement.

Presence of insulin resistance[35]

Preexisting insulin resistance was independently associated with the incidence of POCD. It has been shown that insulin resistance reflects metabolic disease which is related to neuropathological process regarding aging and cognitive function. A reduction in insulin receptor on the BBB according to insulin resistance results in decreased insulin transport into the brain causing POCD and Alzheimer's disease.

Genetics[36]

The Human Apolipoprotein E (ApoE) gene is located on chromosome 19. E4 allele of the APOE gene is evidenced to account for Alzheimer's disease, cognitive dysfunction and atherosclerosis. To date there are no studies confirming the effect of sex difference on POCD. However, men who are the carriers of *APOE4* alleles were reported to have higher risk of POCD than women with *APOE4*.

Alcoholism[37]

The elderly with history of alcohol abuse could pose a higher risk for POCD especially in the domains related to visuospatial and executive functions.

2. Surgical factors[1,6,38]

Complicated and prolonged surgeries > 4 hours, complications during the peri- or postoperative periods, and procedures needing multiple surgeries are all risk factors which could result in POCD. Cardiac surgery with the application of cardiopulmonary bypass pump (CPB) is a predisposing factor to POCD. Prolonged arterial cross clamping time in cardiac surgery plays an important role in POCD according to poor cerebral hypoperfusion.

3. Anesthetic factors

Factors such as a prolonged anesthetic period causing disequilibrium of fluids and electrolytes, acute blood loss, oxygen desaturation, and peri- or postoperative

anesthetic complications account for POCD. Studies have also demonstrated that various anesthetic agents can affect POCD; for example, midazolam may lead to memory impairment than propofol or remifentanil.[39] Medications affecting the cholinergic system can increase the POCD risk. On the other hand, previous studies have found that the perioperative usage of dexmedetomidine may result in a lower POCD risk by reducing the levels of IL-6 and S100B.[40-42] Intraoperative use of volatile anesthesia especially isoflurane and sevoflurane had been reported to influence higher risk of POCD when compared with intravenous propofol.[29]

Assessment

Assessment of POCD is not straightforward. Variations among assessors, different POCD definitions or diagnostic tools used, the timing of evaluation, emotion, degree of pain, medication profile, and environmental setting are common factors contributing to different assessment results. 43 Many studies have assessed POCD by observing changes in patients' neuropsychological signs. The Mini-Mental State Examination (MMSE) assesses orientation (time and place), memory (immediate and short-term), calculation, language (naming, repetition, listening, reading comprehension, and writing), visuospatial awareness, concentration, and attention while the Montreal Cognitive Assessment tool (MoCA) focuses on visuospatial and executive function (alternate trail-making test, copy the cube, and clock drawing), language ability, attention and calculation, delayed recall, and abstract thinking.[44] These tools are the most common clinical screening tests for POCD. However, they are not suitable for cognitive follow-up evaluation.[45] Other test batteries that are designed to evaluate cognitive status include various neuropsychological tests (NPT) which determine specified cognitive domains such as Digit span test, Trail Making Test, Groove Pegboard Test, etc.

Several screening tests for POCD with comparable sensitivity and specificity at differentiating mild cognitive impairment (MCI) from dementia are Addenbrooke's Cognitive Exam (ACE-III), Quick Mild Cognitive Impairment Screening (Qmci), Saint Louis University Mental Status (SLUMS), Mini-cog, Rowland Universal Dementia Assessment Scale (RUDAS) and Abbreviate Mental Test (AMT) (Table 2).[3,46-55]

The new consensus for POCD diagnosis recommends applying the diagnostic criteria for a neurocognitive disorder from DSM-5.[11] Neurocognitive assessment relies on a subjective test (based on the responses of the patient or close relatives), an objective test (standardized NPT) as well as an assessment of the patient's ability to perform the activities of daily living (ADL). This new approach provides a more accurate POCD diagnosis compared with the previous recommendation where only an objective test was considered. Recent publication reported the use of the Thai version of RUDAS to screen for POCD at postoperative day 5-9 through real-time video stream over mobile phone internet connection. Even though the test consumed longer time, almost 30 minutes per each patient, than usual face-to-face evaluation, this method encouraged the use of telemedicine in geriatric patients especially who were not well complied with clinical follow-up.[56]

Prevention

There are currently 2 main strategies for POCD prevention:[1,9]

1. Patient factors

Comprehensive geriatric assessment and preoperative assessment to stratify and optimize risks before proceeding to operations are recommended during the preoperative period.

2. Surgical and anesthetic factors

The Enhanced Recovery after Surgery (ERAS) protocol has been reported to improve postoperative recovery and to reduce rate of postoperative hospitalization and morbidity. The general principles of ERAS involve limited fluid intake, preference of laparoscopic surgery, appropriate anesthetic agent administration, adequate pain medications, early feeding, and early mobilization. The ERAS protocol covers the preoperative, intraoperative, and postoperative states, as outlined below.[3,42,57-64]

a. Preoperative state

- Controlling patients' underlying diseases, optimizing the risk factors, and providing preoperative counselling and prehabilitation.
- Recognizing the risks contributing to worsening outcomes such as major surgery, immobilization, and prolonged hospitalization.
- Encouraging social and moderate physical activities to improve cognitive function.
- Implementation of prehabilitation program at 6-8 weeks preoperatively which involves processes designed to improve the preoperative functional status of the patients by:

- Adequate and appropriate exercise including breathing exercises or resistance training.
- Supplementary dietary intake to improve malnutrition.
- Giving education and advice especially to elderly patients.
- Treatment of comorbidities e.g., atrial fibrillation which could relate to POCD.

b. Immediate preoperative state

TABLE 2. Neuropsychological tests for the assessment of postoperative cognitive dysfunction.

Parameters	MMSE	MoCA	ACE-III	Qmci
Total score	30	30	100	100
Cut off Score for MCI	< 24	< 26	< 82–88	< 62
Average time to complete	10 min	15 min	16 min	5 min
Sensitivity (%)	79.8%	90	84–93	90
Specificity (%)	81.3%	87	100	87
Advantages	- Less time consuming - Easy to use	- High sensitivity - Can identify MCI and cognitive dysfunction in Alzheimer's and Parkinson's diseases	- Can differentiate MCI from early dementia - Provides scores for different cognitive domains with correlation to NPT	- Less time consuming - High sensitivity - Useful test to detect MCI and dementia
Disadvantages	- Low sensitivity and not suitable to screen for MCI	- Designed for MCI rather than dementia	- Cannot differentiate dementia subtypes	- Maybe inaccurate when used in certain subgroups e.g., post stroke patients
Parameters	**SLUMS**	**Mini-cog**	**RUDAS**	**AMT**
Total score	30	5	100	10
Cut off Score for MCI	< 27	< 4	< 25	< 9
Average time to complete	7 min	3 min	10 min	3-5 min
Sensitivity (%)	98	85.7	76.2	91.5
Specificity (%)	98	79.4	75	82.4
Advantages	- High sensitivity and specificity	- Less time consuming - Can be used in primary care setting	- Can differentiate MCI from dementia and normal cognition	- Less time consuming - Can be used as first line screening in acute setting
Disadvantages	- May be affected in patients with ≤ 6 years of education and non-white ethnicity - New tool, not widely used	- Cannot be used in patients with visual impairment or difficulty to hold the pen/pencil	- May be affected in patients with ≤ 6 years of education	- Normal AMT may not exclude MCI therefore it cannot be used as a rule-out test

Abbreviations: ACE-III, Addenbrooke's Cognitive Examination III; AMT, Abbreviated Mental Test; MCI, Mild cognitive impairment; MOCA, The Montreal Cognitive Assessment tool; NPT, Neuropsychological test; Qmci, The Quick Mild Cognitive Impairment; RUDAS, Rowland Universal Dementia Assessment Scale; SLUMPS, Saint Louis University Mental Status

- Reducing the fasting time is beneficial since prolonged starvation stimulates stress and anxiety. Maintaining patients in an euvolemic state is recommended.
- Oral carbohydrate preloading is suggested to stimulate the neuroendocrine response to stress.
- Thromboprophylaxis is recommended for all patients either by intermittent pneumatic compression devices, compression stockings or encouraging early mobilization. The risk of bleeding needs to be evaluated before anticoagulant administration.
- Antibiotic prophylaxis should be given 60 minutes before skin incisions.

c. Intraoperative state
- Preanesthetic sedatives and anxiolytics are not routinely administered.
  Short-acting opioids are recommended.
- Sevoflurane, desflurane, intravenous thiopental, and propofol infusions are recommended to reduce risks for POCD.
- Monitoring of anesthetic depth and cerebral oxygenation during surgery.
- Anesthetic agents with small molecular structures, isoflurane and desflurane, could create amyloid β-oligomerization which involve POCD. A large molecular agent like propofol, if administered with smaller molecular agents, could also result in amyloid β-oligomerization production.
- Avoidance of prolonged nitric oxide usage.
- Laparoscopic surgeries are recommended, given the decreased levels of bowel distention and lower incidences of postoperative nausea and vomiting.
- Low tidal volume ventilation (5–7 mL/kg) and real-time hemodynamic monitoring via an esophageal doppler are recommended.
- A high level of oxygen can increase blood flow to the anastomotic site, lessen the risk of delayed wound healing, and lower the incidences of postoperative nausea and vomiting.
- Epidural analgesia is recommended for open surgical procedures.
- Avoidance of drains or nasogastric tubes as they hinder early mobilization.
- Administration of a local anesthetic around the wound helps reduce pain and urinary retention.
- Fluids overloading should be avoided according to poor intestinal anastomosis, prolonged bowel ileus, and systemic edema.
- Balanced salt solutions are recommended because saline overload leads to metabolic acidosis and hyperchloremia.
- Hypotension from central neuraxial blockade and general anesthesia should be resolved by vasoconstrictors rather than fluid resuscitation.
- Keeping normothermia as hypothermia may result in increased metabolic demand, altered drug metabolism, and impaired immunity and coagulation.

d. Postoperative state
- Opioids should be used rationally. Paracetamol and NSAIDs are encouraged for pain control.
- Epidural anesthesia should be given continuously.
- Oral fluids should be given 2 hours after surgery, and intravenous fluids should be administered cautiously to reduce anastomotic dehiscence and infection.
- Maintenance of blood glucose between 180–200 mg/dl.
- Early mobilization is encouraged. Nasogastric tubes, abdominal drains, opioids, and epidural catheters should therefore not be used from the second day onwards unless necessary.
- The patients should be settled in a calm environment and properly advised before their discharge.

Treatment

POCD treatment is usually based on 2 major principles:[1,10]

1. POCD can be mimicked by POCD-like conditions (e.g., myocardial infarction, septic shock, medication or toxic-substance abuse, electrolyte imbalance, a hypo- or hyperglycemic state, endocrine or liver dysfunction, and neurological deficits). Hypoglycemia can be treated with an intravenous glucose solution. Thiamine is the proper treatment if Wernicke encephalopathy is suspected.

2. The general principles of surgical patient care can be adapted for POCD patients. Some essential methods include the provision of adequate ventilation and oxygenation, hemodynamic support, or adequate postoperative pain control. Pain control is critical and consequently requires sufficient patient counselling. Monitoring of vital signs, electrolytes, and the cardiovascular and pulmonary functions is suggested for enhanced recovery. When POCD is resolved, patients may recall fragments of memories during their POCD episodes which may create stress. It is therefore necessary that medical personnel give appropriate advice and counselling.

CONCLUSION

POCD is a condition that warrants early detection and treatment. The best measure to is to prevent it before the resulting cognitive dysfunction develops. Proper pre-, intra- and postoperative patient care should be considered to reduce the risks of POCD. Preoperative assessments are necessary to identify patients at risk and to optimize patients' condition for particular surgical procedures. According to the ERAS protocol, hemodynamic stabilization, adequate fluid administration, avoidance of excessive anesthetic agents, and appropriate management of hypotension or hypothermia are measures for POCD risk reduction. Adequate postoperative pain control, breathing exercise, and early mobilization are essential to prevent POCD and improve patients' outcomes. Additionally, appropriate environmental settings and frequent postoperative POCD assessments are encouraged. Once POCD develops, patients should be treated promptly, scheduled for regular follow-ups, and given proper counselling.

ACKNOWLEDGMENTS

The authors thank Dr. Arunotai Siriussawakul, M.D., for her advice and support with the content review and development of the final revision. We also gratefully acknowledge the administrative assistance provided by Miss Chanita Janoonsong.

Conflicts of interest: The authors declare that there are no conflicts of interest.

REFERENCES

1. Rundshagen I. Postoperative cognitive dysfunction. Dtsch Arztebl Int. 2014;111(8):119-25.
2. Moller JT, Cluitmans P, Rasmussen LS, Houx P, Rasmussen H, Canet J, et al. Long-term postoperative cognitive dysfunction in the elderly ISPOCD1 study. ISPOCD investigators. International Study of Post-Operative Cognitive Dysfunction. Lancet. 1998 351(9106):857-61.
3. Needham MJ, Webb CE, Bryden DC. Postoperative cognitive dysfunction and dementia: what we need to know and do. Br J Anaesth. 2017;119(Suppl_1):i115-i125.
4. Li D, Liu H. Cognitive function assessment should be included in preoperative evaluation. J Biomed Res. 2017;32(3):161-3.
5. Lertkovit S, Siriussawakul A, Suraarunsumrit P, Lertpipopmetha W, Manomaiwong N, Wivatdechakul W, et al. Polypharmacy in older adults undergoing major surgery: prevalence, association with postoperative cognitive dysfunction and potential associated anesthetic agents. Front Med (Lausanne). 2022;9:811954.
6. Wang M, Wang J, Li X, Xu X, Zhao Q, Li Y. A predictive model for postoperative cognitive dysfunction in elderly patients with gastric cancer: a retrospective study. Am J Transl Res. 2022;14(1):679-86.
7. Fan Y, Liu X, Wu S, Liu Y. The risk factors for cognitive dysfunction in elderly patients after laparoscopic surgery: A retrospective analysis. Medicine (Baltimore). 2021;100(2):e23977.
8. Monk TG, Weldon BC, Garvan CW, Dede DE, van der Aa MT, Heilman KM, et al. Predictors of cognitive dysfunction after major noncardiac surgery. Anesthesiology. 2008;108(1):18-30.
9. Kotekar N, Shenkar A, Nagaraj R. Postoperative cognitive dysfunction - current preventive strategies. Clin Interv Aging. 2018;13:2267-73.
10. Pappa M, Theodosiadis N, Tsounis A, Sarafis P. Pathogenesis and treatment of post-operative cognitive dysfunction. Electron Physician. 2017;9(2):3768-75.
11. Evered L, Silbert B, Knopman DS, Scott DA, DeKosky ST, Rasmussen LS; Nomenclature Consensus Working Group. Recommendations for the nomenclature of cognitive change associated with anaesthesia and surgery-2018. Acta Anaesthesiol Scand. 2018;62(10):1473-80.
12. Zhu M, Li Y, Wan Z, Zhang D, Wang X. Effects of small-dose lidocaine combined with ketamine on early postoperative cognitive function in elderly patients undergoing gastrointestinal tumor surgery. Nan Fang Yi Ke Da Xue Xue Bao. 2015;35(7):1076-8
13. Suraarunsumrit P, Pathonsmith C, Srinonprasert V, Sangarunakul N, Jiraphorncharas C, Siriussawakul A. Postoperative cognitive dysfunction in older surgical patients associated with increased healthcare utilization: a prospective study from an upper-middle-income country. BMC Geriatr. 2022;22(1):213.
14. Weng CF, Lin KP, Lu FP, Chen JH, Wen CJ, Peng JH, et al. Effects of depression, dementia and delirium on activities of daily living in elderly patients after discharge. BMC Geriatr. 2019;19(1):261.
15. Luo A, Yan J, Tang X, Zhao Y, Zhou B, Li S. Postoperative cognitive dysfunction in the aged: the collision of neuroinflammaging with perioperative neuroinflammation. Inflammopharmacology. 2019;27(1):27-37.
16. Netto MB, de Oliveira Junior AN, Goldim M, Mathias K, Fileti ME, da Rosa N, et al. Oxidative stress and mitochondrial dysfunction contributes to postoperative cognitive dysfunction in elderly rats. Brain Behav Immun. 2018;73:661-9.
17. Zhang S, Dong H, Zhang X, Li N, Sun J, Qian Y. Cerebral mast cells contribute to postoperative cognitive dysfunction by promoting blood brain barrier disruption. Behav Brain Res. 2016;298(Pt B):158-66.
18. Zhu H, Liu W, Fang H. Inflammation caused by peripheral immune cells across into injured mouse blood brain barrier can worsen postoperative cognitive dysfunction induced by isoflurane. BMC Cell Biol. 2018;19(1):23.
19. Xiao JY, Xiong BR, Zhang W, Zhou WC, Yang H, Gao F, et al. PGE2-EP3 signaling exacerbates hippocampus-dependent cognitive impairment after laparotomy by reducing expression levels of hippocampal synaptic plasticity-related proteins in aged mice. CNS Neurosci Ther. 2018;24(10):917-29.
20. Li QH, Yu L, Yu ZW, Fan XL, Yao WX, Ji C, et al. Relation of postoperative serum S100A12 levels to delirium and cognitive dysfunction occurring after hip fracture surgery in elderly patients. Brain Behav. 2019;9(1):e01176.
21. Noll F, Behnke J, Leiting S, Troidl K, Alves GT, Müller-Redetzky H, et al. Self-extracellular RNA acts in synergy with exogenous danger signals to promote inflammation. PLoS One. 2017;12(12):e0190002.

22. Li JJ, Wang B, Kodali MC, Chen C, Kim E, Patters BJ, et al. In vivo evidence for the contribution of peripheral circulating inflammatory exosomes to neuroinflammation. J Neuroinflammation. 2018;15(1):8.
23. Ma G, Chen C, Jiang H, Qiu Y, Li Y, Li X, et al. Ribonuclease attenuates hepatic ischemia reperfusion induced cognitive impairment through the inhibition of inflammatory cytokines in aged mice. Biomed Pharmacother. 2017;90:62-8.
24. Liu Y, Yin Y. Emerging Roles of Immune Cells in Postoperative Cognitive Dysfunction. Mediators Inflamm. 2018;2018:6215350.
25. Lim A, Krajina K, Marsland AL. Peripheral inflammation and cognitive aging. Mod Trends Pharmacopsychiatry. 2013;28: 175-87.
26. Hudetz JA, Iqbal Z, Gandhi SD, Patterson KM, Byrne AJ, Hudetz AG, et al. Ketamine attenuates post-operative cognitive dysfunction after cardiac surgery. Acta Anaesthesiol Scand. 2009;53(7):864-72.
27. Androsova G, Krause R, Winterer G, Schneider R. Biomarkers of postoperative delirium and cognitive dysfunction. Front Aging Neurosci. 2015;7:112.
28. Fodale V, Santamaria LB, Schifilliti D, Mandal PK. Anaesthetics and postoperative cognitive dysfunction: a pathological mechanism mimicking Alzheimer's disease. Anaesthesia. 2010;65(4):388-95.
29. Geng YJ, Wu QH, Zhang RQ. Effect of propofol, sevoflurane, and isoflurane on postoperative cognitive dysfunction following laparoscopic cholecystectomy in elderly patients: A randomized controlled trial. J Clin Anesth. 2017;38:165-71.
30. Strøm C, Rasmussen LS, Sieber FE. Should general anaesthesia be avoided in the elderly? Anaesthesia. 2014;69 Suppl 1(Suppl 1): 35-44.
31. Agarwal P, Panda AK, Jena S, Mohapatra SSG. Correlation of cerebral atrophy and white matter hyperintensity burden in MRI with clinical cognitive decline. SMJ. 2022;74(5):323-9.
32. Feinkohl I, Winterer G, Spies CD, Pischon T. Cognitive Reserve and the Risk of Postoperative Cognitive Dysfunction. Dtsch Arztebl Int. 2017;114(7):110-7.
33. Berger M, Nadler JW, Browndyke J, Terrando N, Ponnusamy V, Cohen HJ, et al. Postoperative Cognitive Dysfunction: Minding the Gaps in Our Knowledge of a Common Postoperative Complication in the Elderly. Anesthesiol Clin. 2015;33(3):517-50.
34. Katchamart W, Koolvisoot, A. Cognitive dysfunction in Systemic Lupus Erythematosus patients at Siriraj Hospital. SMJ. 2004;56(9):485-92.
35. He X, Long G, Quan C, Zhang B, Chen J, Ouyang W. Insulin Resistance Predicts Postoperative Cognitive Dysfunction in Elderly Gastrointestinal Patients. Front Aging Neurosci. 2019; 11:197.
36. Schenning KJ, Murchison CF, Mattek NC, Kaye JA, Quinn JF. Sex and genetic differences in postoperative cognitive dysfunction: a longitudinal cohort analysis. Biol Sex Differ. 2019;10(1):14.
37. Hudetz JA, Iqbal Z, Gandhi SD, Patterson KM, Hyde TF, Reddy DM, et al. Postoperative cognitive dysfunction in older patients with a history of alcohol abuse. Anesthesiology. 2007;106(3): 423-30.
38. Xu T, Bo L, Wang J, Zhao Z, Xu Z, Deng X, et al. Risk factors for early postoperative cognitive dysfunction after non-coronary bypass surgery in Chinese population. J Cardiothorac Surg. 2013;8:204.
39. Yu DN, Zhu Y, Ma J, Sun Q. Comparison of post-anesthesia delirium in elderly patients treated with dexmedetomidine and midazolam maleate after thoracic surgery. Biomedical Research (India). 2017;28(2):6852-5.
40. Yang W, Kong LS, Zhu XX, Wang RX, Liu Y, Chen LR. Effect of dexmedetomidine on postoperative cognitive dysfunction and inflammation in patients after general anaesthesia: A PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore). 2019;98(18):e15383.
41. Li Y, He R, Chen S, Qu Y. Effect of dexmedetomidine on early postoperative cognitive dysfunction and peri-operative inflammation in elderly patients undergoing laparoscopic cholecystectomy. Exp Ther Med. 2015;10(5):1635-42.
42. Zhang J, Liu G, Zhang F, Fang H, Zhang D, Liu S, et al. Analysis of postoperative cognitive dysfunction and influencing factors of dexmedetomidine anesthesia in elderly patients with colorectal cancer. Oncol Lett. 2019;18(3):3058-64.
43. Rasmussen LS, Larsen K, Houx P, Skovgaard LT, Hanning CD, Moller JT; ISPOCD group. The International Study of Postoperative Cognitive Dysfunction. The assessment of postoperative cognitive function. Acta Anaesthesiol Scand. 2001;45(3):275-89.
44. Qiao Y, Feng H, Zhao T, Yan H, Zhang H, Zhao X. Postoperative cognitive dysfunction after inhalational anesthesia in elderly patients undergoing major surgery: the influence of anesthetic technique, cerebral injury and systemic inflammation. BMC Anesthesiol. 2015;15:154.
45. Liu J, Huang K, Zhu B, Zhou B, Ahmad Harb AK, Liu L, et al. Neuropsychological Tests in Post-operative Cognitive Dysfunction: Methods and Applications. Front Psychol. 2021;12:684307.
46. Potts C, Richardson J, Bond RB, Price RK, Mulvenna MD, Zvolsky P, et al. Reliability of Addenbrooke's Cognitive Examination III in differentiating between dementia, mild cognitive impairment and older adults who have not reported cognitive problems. Eur J Ageing. 2021:1-13.
47. Bruno D, Schurmann Vignaga S. Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. Neuropsychiatr Dis Treat. 2019;15:441-7.
48. O'Caoimh R, Gao Y, McGlade C, Healy L, Gallagher P, Timmons S, et al. Comparison of the quick mild cognitive impairment (Qmci) screen and the SMMSE in screening for mild cognitive impairment. Age Ageing. 2012;41(5):624-9.
49. Xu Y, Yi L, Lin Y, Peng S, Wang W, Lin W, et al. Screening for Cognitive Impairment After Stroke: Validation of the Chinese Version of the Quick Mild Cognitive Impairment Screen. Front Neurol. 2021;12:608188.
50. Emery A, Wells J, Klaus SP, Mather M, Pessoa A, Pendlebury ST. Underestimation of Cognitive Impairment in Older Inpatients by the Abbreviated Mental Test Score versus the Montreal Cognitive Assessment: Cross-Sectional Observational Study. Dement Geriatr Cogn Dis Extra. 2020;10(3):205-15.
51. Pendlebury ST, Klaus SP, Mather M, de Brito M, Wharton RM. Routine cognitive screening in older patients admitted to acute medicine: abbreviated mental test score (AMTS) and subjective memory complaint versus Montreal Cognitive Assessment and IQCODE. Age Ageing. 2015;44(6):1000-5.
52. Storey JE, Rowland JT, Basic D, Conforti DA, Dickson HG. The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale. Int Psychogeriatr. 2004;16(1):13-31.
53. Manjavong M, Limpawattana P, Sawanyawisuth K. Performance

of the Rowland Universal Dementia Assessment Scale in Screening Mild Cognitive Impairment at an Outpatient Setting. Dement Geriatr Cogn Dis Extra. 2021;11(2):181-8.

54. Li X, Dai J, Zhao S, Liu W, Li H. Comparison of the value of Mini-Cog and MMSE screening in the rapid identification of Chinese outpatients with mild cognitive impairment. Medicine (Baltimore). 2018;97(22):e10966.
55. Yang YP, Huang YC, Chen CS, Yang YC, Wang JJ. Sensitivity and Specificity of the Saint Louis University Mental Status Examination to Detect Mild Cognitive Impairment and Dementia in Chinese Population. Gerontology. 2021;67(2):152-9.
56. Suesat H, Srinonprasert V, Limpawattana P, Kakyos S, Poontananggul J, Jiraphorncharas C, et al. Detection of postoperative cognitive dysfunction by telemedicine among octogenarian patients who underwent minor elective surgery: Prospective cohort study. SMJ. 2022;74(2):126-33.
57. Sethi AK, Kochhar A, Ahmad Z. Enhanced Recovery After Surgery. In: Sehgal R, Trikha A, editors. Yearbook of Anesthesiology-6. New Delhi (India): Jaypee Brothers Medical Publishers (P) Ltd; 2017. p. 152-632.
58. Khan S, Gatt M, Horgan A, Anderson I, MacFie J. ASGBI: Issues in professional practise guidelines for implementation of enhanced recovery protocols. ASGBI. 2009. [cited 2022 Apr 28]. Available from: https://www.asgbi.org.uk/en/publications/issues_in_ professional_practice.cfm.
59. Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. Arch Phys Med Rehabil. 2004;85(10):1694-704.
60. Moran J, Guinan E, McCormick P, Larkin J, Mockler D, Hussey J, et al. The ability of prehabilitation to influence postoperative outcome after intra-abdominal operation: A systematic review and meta-analysis. Surgery. 2016;160(5):1189-201.
61. Mandal PK, Bhavesh NS, Chauhan VS, Fodale V. NMR investigations of amyloid-β peptide interactions with propofol at clinically relevant concentrations with and without aqueous halothane solution. J Alzheimers Dis. 2010;21(4):1303-9.
62. Mandal PK, Simplaceanu V, Fodale V. Intravenous anesthetic diazepam does not induce amyloid-beta peptide oligomerization but diazepam co-administered with halothane oligomerizes amyloid-beta peptide: an NMR study. J Alzheimers Dis. 2010;20(1):127-34.
63. Futier E, Constantin JM, Paugam-Burtz C, Pascal J, Eurin M, Neuschwander A; IMPROVE Study Group. A trial of intraoperative low-tidal-volume ventilation in abdominal surgery. N Engl J Med. 2013;369(5):428-37.
64. Wuethrich PY, Burkhard FC, Thalmann GN, Stueber F, Studer UE. Restrictive deferred hydration combined with preemptive norepinephrine infusion during radical cystectomy reduces postoperative complications and hospitalization time: a randomized clinical trial. Anesthesiology. 2014;120(2):365-77.

Serum Neurofilament Light Chain: A Potential Biomarker for Peripheral Neuropathy

I Putu Eka Widyadharma, M.D., Ph.D., Eric Hartono Tedyanto, M.D.

Department of Neurology, Faculty of Medicine, Universitas Udayana, Bali, Indonesia.

ABSTRACT

In some neurological diseases, advanced examinations can be used as diagnostic tools. Several indicators have also been discovered that can be used to assess the severity of neuronal damage and neurological disease progression. Neurofilament light chain (NfL) is a cytoskeleton protein that makes up the structure of neuron axons and is released when a neuron is injured, allowing it to assess neuronal injury severity. NfL was first used to diagnose central nervous system disorders like dementia, multiple sclerosis, and other neurodegenerative diseases. But, NfL levels have also been elevated in peripheral nervous system disorders, like in several neuropathic conditions, including amyloid neuropathy, HIV-associated neuropathy, diabetic peripheral neuropathy, leprosy neuropathy, and other neuropathy, according to various investigations. Theoretically, all abnormalities induced by axonal injury will increase blood NfL levels, allowing NfL testing to be utilized as a measurement tool. NfL levels can also be a predictive indicator to monitor treatment efficacy and peripheral neuropathy progression.

Keywords: Biomarker; neurofilament light chain; peripheral neuropathy; prognostic (Siriraj Med J 2022; 74: 714-720)

INTRODUCTION

In recent years, neurology has made significant technological innovations. A variety of neuroimaging methods can generate accurate images of the brain. Moreover, various biomarkers have been developed which may be used in clinical trials to estimate the level of neuronal damage. One of them is the neurofilament light chain (NfL). NfL is released into the CSF and bloodstream whenever there is damage to neurons.[1]

Since it is an indicator of axonal damage, the serum neurofilament light chain (NfL) is a potential diagnostic in neurological diseases. Previously, NfL has only been detected in CSF. The NfL can still be detected in the blood due to the new advanced technologies, making it simpler to detect and avoiding traumatic procedures like a lumbar puncture. A neurologist can use NfL as a prospective diagnostic as an accurate sign of nerve injury. If the cardiologist has troponin, the neurologist has the neurofilament light chain (NfL).[2]

NfL concentrations in the normal population are rarely reported. Tobias et al revealed that NfL levels in normal populations are 7.3 (±3) pg/mL in serum and 416 (±191) pg/mL in CSF. In patients with Multiple Sclerosis, NfL levels are 16.4 (±14.4) pg/mL in serum and 2368 (±1947) pg/mL in CSF. The levels of NfL are affected by age, BMI, and renal function. The association between age and NfL concentration was positive ($r = 0.325$; p-value <0.0001), while the correlation between BMI and NfL concentration was negative ($r = 0.227$; p-value <0.0001). No significant differences exist between NfL concentration and gender. In addition, there was a strong correlation between NF-L levels and renal function. NfL concentration and eGFR were also found to have a very strong connection ($r = -0.492$; p-value <0.0001).[3-5]

Corresponding author: Eric Hartono Tedyanto
E-mail: erichartonoo@gmail.com
Received 28 April 2022 Revised 15 August 2022 Accepted 16 August 2022
ORCID ID:http://orcid.org/0000-0002-4554-0348
http://dx.doi.org/10.33192/Smj.2022.84

CSF and serum NfL levels were higher in patients with a central or peripheral nervous system injury. This increase has been linked to neurological diseases, according to certain studies. The NfL can also be used to predict future outcomes. Because it can be easily detected and non-invasively in the blood, NfL is a promising biomarker for monitoring the progression of neurological diseases and evaluating the efficacy of therapy.[6]

Peripheral neuropathy affects approximately 2.4 percent of the population, with symptoms varying depending on what type of nerve fiber is affected, the type of neuron injury, and the severity of the injury. Peripheral neuropathy is most commonly caused by diabetes. However, HIV can also directly or indirectly induce peripheral neuropathy through antiretroviral (ARV) medications. Systemic disease, infection, and malnutrition are also all potential causes of peripheral neuropathy.[7]

A neurologist might conduct an electrophysiological evaluation of nerve conduction velocity to diagnose peripheral neuropathy. However, nerve conduction velocity may not be able to accurately assess the severity and progression of neuropathy in some conditions, requiring the use of additional biomarkers to determine prognostic value. On the other hand, NfL has lately undergone massive research and can be utilized as a biomarker for peripheral nerve injury. Serum NfL levels are known to be elevated in cases of peripheral neuropathy and correlate with disease severity.[8] In this review, we provide the role of the neurofilament light chain (NfL) as a biomarker of peripheral neuropathy.

Neurofilament light chain (NfL)

The essential features of neurons are neurofilaments, built up of protein triplets and present on nerve axons (Fig 1). The neurofilament core cannot functionally work without the neurofilament light chain (NfL) subunit. Almost every neuron component contains the protein neurofilament light chain (NfL). The diameter and speed of nerve conduction from peripheral nerves are determined by NfL accumulation, linked to axon growth during myelination.[9] Depending on the severity of axonal damage in peripheral nerves, NfL can be released into the extracellular space and bloodstream. An apophagocytotic process releases NfL into the CSF and bloodstream when neurons in the central nervous system are damaged. NfL will enter the CSF through direct drainage and then enter the bloodstream through arachnoid granulation and lymphatic flow in the subarachnoid space, making it detectable in both the CSF and the blood.[10] According to some studies, the amount of NfL in CSF is 500 times higher than in blood because CSF is directly related to the central nervous system. NfL concentrations in the blood are too low to be detected by an ELISA test. Thanks to recent technological advancements, a new method, SIMOA (Single-Molecule Assay (SiMoA), has been developed to detect NfL down to a single-digit picogram per millimeter unit.[11]

NfL in neurological cases

Axonal damage in the central and peripheral nervous systems, such as stroke, head trauma, multiple sclerosis, ALS, Alzheimer's disease, frontotemporal dementia,

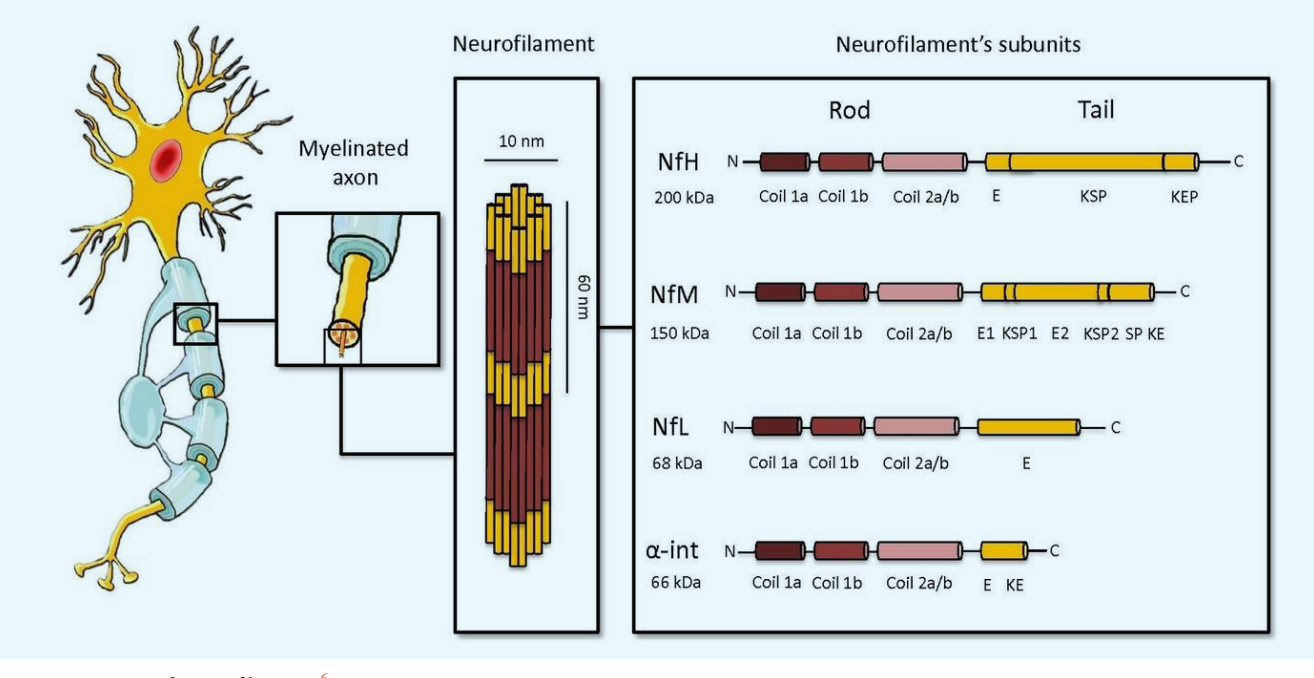


Fig 1. Structure of neurofilament.[6]

and peripheral neuropathy, can be identified by NfL measurement. Previous studies linked increased NfL levels in CSF and serum to neurodegenerative and neuroinflammatory processes, indicating demyelination and axonal damage.[12-16]

Neurological patients have much greater NfL levels in their CSF and blood than healthy or non-neurological patients. NfL is a test that can identify neurological problems caused by axonal damage. It can tell the difference between varying degrees of axonal damage, progression, and whether or not it is a neurodegenerative condition. As a result, the NfL examination is utilized as a biomarker to validate the diagnosis after a full neurological examination or other biomarker and neuroimaging procedures.[6] In the event of peripheral neuropathy, NfL can be utilized as a non-invasive diagnostic technique to determine therapy success and progression.[8]

It is unclear how long the duration of NfL levels increases in patients with peripheral neuropathy. In studies on multiple sclerosis, traumatic brain injury, and stroke, NfL levels peak 3–4 weeks after a clinical relapse and remain elevated for 6–12 months. Further studies on how long NfL levels increase in peripheral neuropathy need to be done for prognostic purposes.[17]

NfL in peripheral neuropathy

Peripheral neuropathy has become a global health concern, affecting 2.4 percent of the world's population, or around 10 million people in the European Union and 7 million in the United States. Measurement of nerve conduction velocity is the gold standard for diagnosing peripheral neuropathy. However, it cannot be used as a monitor for the success of therapy or disease progression in some cases. As a result, a peripheral nerve damage biomarker is required. On the other hand, NfL has recently undergone extensive research and can be used as a biomarker for peripheral nerve damage. NfL levels have been shown to increase in peripheral neuropathy patients' blood and correlate with disease severity, implying that the NfL is involved in disease progression and can be used as a prognostic factor in peripheral nerve damage.[8,18]

Neuronal neurofilament breakdown is thought to use a combination of ubiquitin-mediated proteasomal and apophagocytotic mechanisms. Based on the transport of other CNS-degraded proteins, it is likely that neurofilament fragments drain directly into CSF and blood via numerous pathways. These include lymphatic outflow into subarachnoid and perivascular regions and direct draining into CSF and blood via arachnoid granulations. Once NfL enters the bloodstream, the half-life is a crucial factor with consequences for disease activity monitoring frequency. In a longitudinal study of NfL levels before and after implantation of an intrathecal catheter, NfL levels in both CSF and serum peaked one month after surgery and returned to baseline six to nine months later.[2]

Other biomarkers besides NfL can be used to diagnose peripheral neuropathy, including Brain Derived Neurotrophic Factor (BDNF), Nerve Growth Factor (NGF), and other inflammatory markers such as IL-1, 6,10, 18, and TNF-alpha. Low BDNF levels were correlated with CIPN in 91 multiple myeloma patients treated

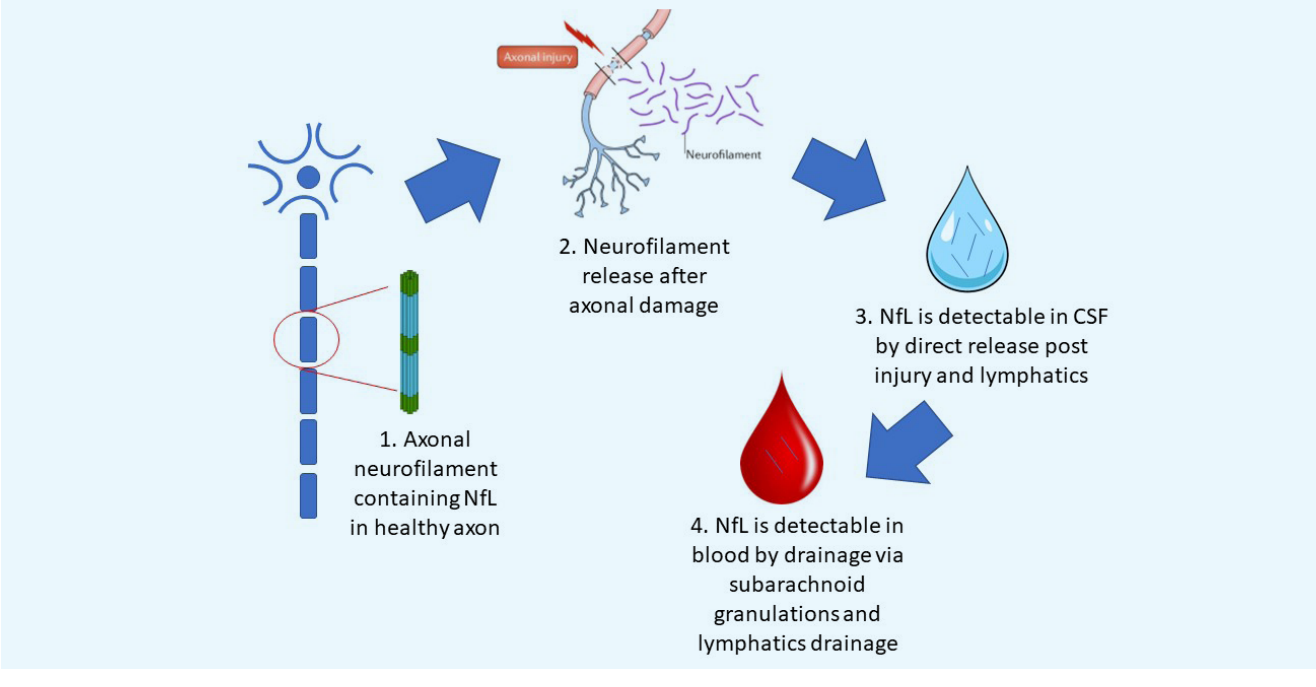


Fig 2. Pathophysiology of neurofilament light chain in cerebrospinal fluid (CSF) and blood.

with bortezomib, and a cut-point of 9.11 ng/ml was 76% sensitive and 71% specific for identifying Chemotherapy-induced Peripheral Neuropathy (CIPN). Nonetheless, another study found no correlation between BDNF and the incidence of CIPN. In one investigation, a correlation was shown between decreasing NGF and the severity of neuropathy as measured by nerve conduction velocity testing. High levels of IFN-, IL-1, and IL-8, but low levels of IL-10 and IL-6, were linked to peripheral neuropathy symptoms. Due to the inconsistency and expense of these biomarkers, NfL testing is recommended to monitor peripheral neuropathy.[19]

Sandelius et al. suggested that the cut-off value of NfL for peripheral neuropathy was 20 pg/mL with a sensitivity of 71% and specificity of 75%. Increased serum NfL concentration is not specific to peripheral neuropathy because other neurological disorders such as multiple sclerosis, Alzheimer's disease, stroke, and Amyotrophic Lateral Sclerosis (ALS) also reported increases. NfL is not useful for diagnosis, but it may be useful to measure axonal damage and could serve as a biomarker of progressivity of the disease for monitoring and response to treatment. NfL is sensitive to detecting axonal damage and correlates with disease severity and progressivity.[18]

Several studies have reported elevated levels of NfL in amyloid neuropathy, HIV-associated neuropathy, diabetic peripheral neuropathy, chemotherapy-induced peripheral neuropathy, and pyridoxine-induced sensory neuropathy.

Amyloid neuropathy

Amyloidosis patients with polyneuropathy experience axonal degeneration, which results in elevated serum NfL levels. Axonal degeneration is caused by the accumulation of amyloid fibrils in the endoneurium and direct toxicity to the nerve's prefibrillar oligomers.[20-22] Patients with symptomatic polyneuropathy, as well as those who are asymptomatic, have elevated serum NfL levels. Serum NfL levels can be used as a marker for early-stage axonal damage in asymptomatic or subclinical amyloidosis, making it essential to diagnose, treat, and monitor the progress and success of amyloidosis therapy.[23] The AUC between asymptomatic and symptomatic amyloid neuropathy patients was 0.99 (p .001), and a NfL concentration of 10.6 pg/mL distinguished these individuals with a sensitivity of 96.2% and a specificity of 93.8%.[20] Serum NfL levels increase the most in patients with abnormal EMG results. This demonstrates that serum NfL is a sensitive marker for early detection of polyneuropathy and is strongly associated with the disease.[14]

HIV-associated neuropathy

HIV-associated neuropathy manifests as distal symmetrical polyneuropathy and toxic antiretroviral neuropathy (ATN), which is difficult to distinguish clinically and electrophysiologically regardless of the use of antiretroviral drugs or the onset of symptoms. HIV-associated neuropathy is linked to the patient's viral load and CD4+ cell count. The use of dNRTIs like stavudine, didanosine, or zalcitabine has been linked to ATN. After antiretroviral therapy, the symptoms of HIV-related polyneuropathy improve as the viral load decreases. After a year of ARV treatment, the symptoms of ATN will worsen.[24]

NFL is a structural component of myelinated axons that have been used as a marker of axonal damage in neurodegenerative diseases in several studies. Axonal damage also occurs in HIV-associated neuropathy, but research on elevated serum NfL levels in HIV-associated neuropathy is uncommon. The HIV in Dementia study is the most widely conducted. Compared to HIV patients without dementia, NfL levels were significantly higher in HIV patients with dementia. The levels of plasma NfL and CSF NfL did not differ significantly. Damian et al. conducted a study to see if NfL levels were elevated in HIV-associated neuropathy patients. The researchers discovered an increase in NfL levels in both CSF and serum in 26 of 54 patients with neuropathy, which correlates to the severity of the neuropathy.[25] NfL levels are not only used as markers of damage to the central nervous system but also in the peripheral nervous system, such as neuropathy, according to these studies.[26]

Chemotherapy induced peripheral neuropathy (CIPN)

CIPN is a side effect of chemotherapy in some cancers. Proper diagnosis, treatment, and dosage adjustments are required to avoid permanent nerve damage. Because CIPN is an axonopathy, it can mimic the symptoms of polyneuropathy. Previous research has discovered that elevated serum NfL levels are linked to peripheral neuropathy and the severity of nerve damage, allowing NfL levels to be measured in CIPN patients. In a mouse model given the cytostatic drug vincristine (VCR) 0.2 mg/kg intravenously four times per week, serum NfL levels increased fourfold, with signs of axonopathy on neurophysiological and pathological examinations. The presence of the NfL in the blood can determine the severity of CIPN.[27-29]

Other chemotherapy drugs, such as oxaliplatin, can cause neuronal cell death and neuropathy in the dorsal ganglion. One study found a link between serum NfL levels and changes in nerve amplitude after treatment with

oxaliplatin. Serum NfL levels were significantly higher in 5 patients with grade 3 OIPN (oxaliplatin-induced peripheral neuropathy) than in grades 0-2 (80 percent sensitivity and 86 percent specificity with a cut-off value of 195 pg/mL). Based on the findings of these studies, serum NfL can be used as a monitor for the severity of OIPN.[30]

Diabetic peripheral neuropathy (DPN)

An observational study used Serum NfL as a non-invasive diagnostic tool to detect diabetic peripheral neuropathy and its progression. NfL levels are related to the neuropathy disability score (NDS) and decreased nerve conduction velocity in some nerves. The AUC for serum NfL was 0.564, and the DPN cut-off point was 12.6 pg/ml. NfL is also associated with the hyperalgesia phenotype and is positively correlated with the severity of DPN.[31,32]

NfL mRNA levels have also been elevated in prediabetic patients with peripheral neuropathy. This supports the hypothesis that NfL mRNA levels are significantly higher in prediabetic patients when small-diameter nociceptive afferent C fibers are interfered with in hyperglycemic conditions, causing axon damage and neuropathic pain symptoms. This level is positively correlated with DN4 questionnaire score.[9,33]

Leprosy neuropathy

Mycobacterium leprae can damage both myelinated and unmyelinated nerve fibers. Patients with leprosy may develop painful neuropathy symptoms. The pathogenesis of neuropathy in leprosy includes infection of the Schwann cells, demyelination, and damage to the axons, leading to atrophy. In tuberculoid and borderline leprosy, axon damage is caused by inflammation of the endoneurial membrane, which destroys nerve structures and causes nerve damage. According to the results of a nerve biopsy, axons and myelin are lost in patients with leprosy. Electromyography also revealed axonal polyneuropathy. Axon damage is a focus of research into the mechanism of leprosy neuropathy. Biomarkers such as NfL can be used with other tests to help determine prognosis and treatment success.[35,36]

Post-herpetic neuralgia

After the reactivation of the varicella-zoster virus, which damages the cell body and axons, post-herpetic neuralgia (PHN) develops. The pathology of PHN is associated with peripheral axonal damage, sensory neuron degeneration, and dorsal horn atrophy. However, several theories suggest that after viral reactivation, axonal damage occurs due to inflammation in PHN. The role of NfL in post-herpetic neuralgia has rarely been studied. More research on the NfL as a biomarker of post-herpetic neuralgia is needed.[37]

Pyridoxine-induced sensory neuropathy (PISN)

According to a study, NfL levels in the CSF and blood increased on day four after rats received pyridoxine therapy. Pyridoxine's primary target is the cell body of

TABLE 1. Summarize how NfL is used in various peripheral neuropathy diseases.

Disease	NfL levels	Indication	Sensitivity	Specificity	AUC	Reference
Peripheral neuropathy	20 pg/mL	Prognostic	71%	75%	0.755	18
Amyloid neuropathy	10.6 pg/mL	Prognostic	96.2%	93.8%	0.99	20
CIPN	195 pg/mL	Prognostic, treatment response	80%	86%	N/A	30
Diabetic peripheral neuropathy	12.6 pg/ml	Prognostic, treatment response	77.6%	86.3%	0.564	32
ALS	93 pg/ml	Prognostic	80.5%	90.9%	0.85	38
Post-stroke cognitive impairment	46.12 pg/ml	Prognostic	71%	81.5%	0.785	39

DRG neurons, which is followed by secondary nerve fiber degeneration. NfL is released directly from the DRG to the CSF via the subarachnoid space from the neuronal cell body and surrounding nerve fibers.[34]

Summarize

As described above, NfL can help determine the progression and response to peripheral neuropathy treatment. Here we provide a table summarizing how NfL is used in various peripheral neuropathy diseases discussed in the manuscript (Table 1). We also compared diseases such as ALS and post-stroke cognitive impairment. No studies determine the cut-off value for some types of peripheral neuropathy. Further research on cut-off NfL levels needs to be done.

CONCLUSION

Serum NfL can be used as a diagnostic tool for peripheral neuropathy after a careful history and physical examination. In addition, NfL levels can also be used as a monitor for the success of therapy and the progression of peripheral neuropathy to be used as a prognostic value. Further studies regarding when serum NfL levels begin to elevate, how long they last, and clear cut-off points for each type of peripheral neuropathy are needed to strengthen the diagnostic value and specificity of serum NfL.

Conflict of interest: No conflict of interest

REFERENCES

1. Barro C, Chitnis T, Weiner HL. Blood neurofilament light: a critical review of its application to neurologic disease. Ann Clin Transl Neurol. 2020;7(12):2508-23.
2. Thebault S, Booth RA, Freedman MS. Blood neurofilament light chain: The neurologist's troponin? Biomedicines. 2020;8(11): 1-11.
3. Sejbaek T, Mendoza JP, Penner N, Madsen JS, Olsen DA, Illes Z. Comparison of neurofilament light chain results between two independent facilities. BMJ Neurol Open. 2020;2(2):e000063.
4. Ladang A, Kovacs S, Lengelé L, Locquet M, Reginster JY, Bruyère O, et al. Neurofilament light chain concentration in an aging population. Aging Clin Exp Res. 2022;34(2):331-9.
5. Bäckström D, Linder J, Jakobson Mo S, Riklund K, Zetterberg H, Blennow K, et al. NfL as a biomarker for neurodegeneration and survival in Parkinson disease. Neurology. 2020;95(7):e827-38.
6. Gaetani L, Blennow K, Calabresi P, Di Filippo M, Parnetti L, Zetterberg H. Neurofilament light chain as a biomarker in neurological disorders. J Neurol Neurosurg Psychiatry. 2019; 90(8):870-81.
7. Benn M. Peripheral Neuropathy-Time for Better Biomarkers? Clin Chem. 2020;66(5):638-40.
8. Remiche G, Kadhim H, Maris C, Mavroudakis N. Peripheral neuropathies, from diagnosis to treatment, review of the literature and lessons from the local experience. Rev Med Brux. 2013;34(4):211-20.
9. Celikbilek A, Tanik N, Sabah S, Borekci E, Akyol L, Ak H, et al. Elevated neurofilament light chain (NFL) mRNA levels in prediabetic peripheral neuropathy. Mol Biol Rep. 2014;41(6): 4017-22.
10. Thebault S, Tessier DR, Lee H, Bowman M, Bar-Or A, Arnold DL, et al. High serum neurofilament light chain normalizes after hematopoietic stem cell transplantation for MS. Neurol Neuroimmunol Neuroinflammation. 2019;6(5):e598.
11. Wilson DH, Rissin DM, Kan CW, Fournier DR, Piech T, Campbell TG, et al. The Simoa HD-1 Analyzer: A Novel Fully Automated Digital Immunoassay Analyzer with Single-Molecule Sensitivity and Multiplexing. J Lab Autom. 2016;21(4):533-47.
12. Jordanova A, De Jonghe P, Boerkoel CF, Takashima H, De Vriendt E, Ceuterick C, et al. Mutations in the neurofilament light chain gene (NEFL) cause early onset severe Charcot-Marie-Tooth disease. Brain. 2003;126(3):590-7.
13. Kuchel GA, Poon T, Irshad K, Richard C, Julien JP, Cowen T. Decreased neurofilament gene expression is an index of selective axonal hypotrophy in ageing. Neuroreport. 1996;7(8):1353-9.
14. Perrot R, Berges R, Bocquet A, Eyer J. Review of the multiple aspects of neurofilament functions, and their possible contribution to neurodegeneration. Mol Neurobiol. 2008;38(1):27-65.
15. Disanto G, Barro C, Benkert P, Naegelin Y, Schädelin S, Giardiello A, et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. Ann Neurol. 2017;81(6):857-70.
16. Landqvist Waldö M, Frizell Santillo A, Passant U, Zetterberg H, Rosengren L, Nilsson C, et al. Cerebrospinal fluid neurofilament light chain protein levels in subtypes of frontotemporal dementia. BMC Neurol. 2013;13.
17. Ferreira-Atuesta C, Reyes S, Giovanonni G, Gnanapavan S. The Evolution of Neurofilament Light Chain in Multiple Sclerosis. Front Neurosci. 2021;15:642384.
18. Sandelius Å, Zetterberg H, Blennow K, Adiutori R, Malaspina A, Laura M, et al. Plasma neurofilament light chain concentration in the inherited peripheral neuropathies. Neurology. 2018;90(6): e518-24.
19. Rodwin RL, Siddiq NZ, Ehrlich BE, Lustberg MB. Biomarkers of Chemotherapy-Induced Peripheral Neuropathy: Current Status and Future Directions. Front Pain Res. (Lausanne). 2022;3:864910.
20. Maia LF, Maceski A, Conceição I, Obici L, Magalhães R, Cortese A, et al. Plasma neurofilament light chain: an early biomarker for hereditary ATTR amyloid polyneuropathy. Amyloid. 2020;27(2): 97-102.
21. Kapoor M, Foiani M, Heslegrave A, Zetterberg H, Lunn MP, Malaspina A, et al. Plasma neurofilament light chain concentration is increased and correlates with the severity of neuropathy in hereditary transthyretin amyloidosis. J Peripher Nerv Syst. 2019;24(4):314-9.
22. Reilly MM, King RHM. Familial Amyloid Polyneuropathy. Brain Pathol. 1993;3(2):165-76.
23. Shin SC, Robinson-Papp J. Amyloid neuropathies. Mt Sinai J Med. 2012;79(6):733-48.
24. Amruth G, Praveen-kumar S, Nataraju B, Nagaraja BS. HIV associated sensory neuropathy. J Clin Diagn Res. 2014;8(7): MC04-7.
25. Damian A, Skolasky R, Demsky C, McArthur J, Sacktor N, Zetterberg H, et al. Neurofilament light chain protein levels in HIV positive subjects with neuropathy. Neurology. 2019;92(15

Supplement 1).
26. Gisslén M, Price RW, Andreasson U, Norgren N, Nilsson S, Hagberg L, et al. Plasma Concentration of the Neurofilament Light Protein (NFL) is a Biomarker of CNS Injury in HIV Infection: A Cross-Sectional Study. EBioMedicine. 2015;3:135-40.
27. Cavaletti G, Marmiroli P. Chemotherapy-induced peripheral neurotoxicity. Nat Rev Neurol. 2010;6(12):657–66.
28. Cavaletti G, Cornblath DR, Merkies ISJ, Postma TJ, Rossi E, Frigeni B, et al. The chemotherapy-induced peripheral neuropathy outcome measures standardization study: From consensus to the first validity and reliability findings. Ann Oncol. 2013;24(2):454-62.
29. Meregalli C, Fumagalli G, Alberti P, Canta A, Carozzi VA, Chiorazzi A, et al. Neurofilament light chain as disease biomarker in a rodent model of chemotherapy induced peripheral neuropathy. Exp Neurol. 2018;307:129-32.
30. Kim SH, Choi MK, Park NY, Hyun JW, Lee MY, Kim HJ, et al. Serum neurofilament light chain levels as a biomarker of neuroaxonal injury and severity of oxaliplatin-induced peripheral neuropathy. Sci Rep. 2020;10(1):7995.
31. Cai L, Huang J. Neurofilament light chain as a biological marker for multiple sclerosis: A meta-analysis study. Neuropsychiatr Dis Treat. 2018;14:2241-54.
32. Morgenstern J, Groener JB, Jende JME, Kurz FT, Strom A, Göpfert J, et al. Neuron-specific biomarkers predict hypo- and hyperalgesia in individuals with diabetic peripheral neuropathy. Diabetologia. 2021;64(12):2843-55.
33. Singleton JR, Smith AG, Bromberg MB. Increased prevalence of impaired glucose tolerance in patients with painful sensory neuropathy. Diabetes Care. 2001;24(8):1448-53.
34. Sano T, Masuda Y, Yasuno H, Shinozawa T, Watanabe T, Kakehi M. Blood Neurofilament Light Chain as a Potential Biomarker for Central and Peripheral Nervous Toxicity in Rats. Toxicol Sci. 2021;185(1):10-8.
35. Nascimento OJM. Leprosy neuropathy: Clinical presentations. Arq Neuropsiquiatr. 2013;71(9 B):661-6.
36. Hassan R, Hui M, Sireesha Y, Afshan J, Meena AK, Uppin MS. Understanding demyelination in leprosy neuropathy: A nerve biopsy analysis. Ann Indian Acad Neurol. 2020;23(6):829-32.
37. Bennett GJ, Peter CPN. Herpes zoster and postherpetic neuralgia: Past, present and future. Pain Res Manag. 2009;14(4):275-82.
38. Verde F, Otto M, Silani V. Neurofilament Light Chain as Biomarker for Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. Front Neurosci. 2021;15:679199.
39. Wang Z, Wang R, Li Y, Li M, Zhang Y, Jiang L, et al. Plasma Neurofilament Light Chain as a Predictive Biomarker for Post-stroke Cognitive Impairment: A Prospective Cohort Study. Front Aging Neurosci. 2021;13:631738.