



บทความพิเศษ

ข้อแนะนำในการรายงานผลการศึกษา

ผศ.ดร.จุฬาลักษณ์ โกมลตรี, ปร.ด.

การรายงานผลการศึกษาที่ไม่ชัดเจน ไม่ครบถ้วนเป็นปัญหาที่พบบ่อยทั่วโลก เกิดขึ้นในการศึกษาทุกรูปแบบ เช่น observational study (cross-sectional, case-control, cohort), clinical trial, diagnostic test study, cross-over study และ quality improvement study การรายงานผลการศึกษาที่ชัดเจนและครบถ้วนจะทำให้ผู้อ่านมีความเข้าใจในวัตถุประสงค์ของการศึกษา, รูปแบบการศึกษา, การดำเนินการศึกษา, ผลและข้อสรุปที่ได้จากการศึกษา ผู้อ่านสามารถตัดสินใจในเรื่องคุณภาพของการศึกษา (internal validity) และการนำผลการศึกษาไปใช้ในประชากรกลุ่มอื่น (external validity)

คณะทำงานซึ่งประกอบด้วยบุคคลต่าง ๆ เช่น นักระบาดวิทยา, นักสถิติ, เจ้าหน้าที่ของวารสารชั้นนำ เช่น Annals of Internal Medicine, BMJ, International Journal of Epidemiology, JAMA, Preventive Medicine, Lancet ได้ร่วมกันกำหนดข้อแนะนำในการเขียนผลการศึกษาที่เป็น clinical trial เป็นครั้งแรกในปี 1996 ซึ่งเรียกว่า CONSORT (Consolidated Standards of Reporting Trials) statement หลังจากนั้นก็ได้มี

การทบทวน แก้ไข เพิ่มเติม และครั้งสุดท้ายคือปี 2008¹⁻³ CONSORT statement ประกอบด้วย checklist ในด้านต่าง ๆ 5 ด้าน รวม 22 ข้อ คือ Title, abstract (1 ข้อ), Introduction (1 ข้อ), Methods (10 ข้อ), Results (7 ข้อ) และ Discussion (3 ข้อ)

หลังจาก CONSORT statement ได้มีข้อแนะนำในการเขียนผลการศึกษาจากรูปแบบการศึกษาคืออื่น ๆ เช่น STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement สำหรับการศึกษาแบบ observational study ซึ่งรวมถึง cross-sectional, case-control และ cohort study⁴⁻⁵, STARD (Standards for Reporting of Diagnostic Accuracy)⁶ และ SQUIRE (Standards for Quality Improvement Reporting Excellence) สำหรับ health care study⁷ บทความนี้จะได้กล่าวเฉพาะ checklist ใน STROBE และ CONSORT เนื่องจากเป็นลักษณะการศึกษาคือพบบ่อยในงานวิจัยทางสุขภาพ และยกตัวอย่างการเขียนที่ดีจากวารสารต่าง ๆ

ตารางที่ 1 และ 2 แสดง checklist ใน STROBE และ CONSORT statement ผู้นิพนธ์ควรรายงานในหัวข้อต่าง ๆ ของ checklist โดยมีรายละเอียดที่เพียงพอและถูกต้องตามความเป็นจริง ส่วนลำดับและรูปแบบการเขียนขึ้นกับผู้นิพนธ์ธรรมเนียมปฏิบัติ และรูปแบบของวารสาร checklist นี้จะช่วยให้มีการรายงานที่ดี แต่ไม่ได้หมายความว่า จะเป็นการศึกษาที่มีคุณภาพดี checklist ใน STROBE มี 22 หัวข้อ โดยที่ 18 หัวข้อใช้ได้กับทุกรูปแบบการศึกษาของ observational study ส่วน 4 หัวข้อที่เหลือ (หัวข้อ 6, 12, 14 และ 15) มีความแตกต่างกันในแต่ละรูปแบบการศึกษา ส่วน checklist ใน CONSORT มีจำนวน 22 หัวข้อเช่นเดียวกันแต่เนื้อหาบางอย่างแตกต่างกัน เช่น ในหัวข้อ randomization (หัวข้อ 8, 9, 10), blinding (หัวข้อ 11), baseline data (หัวข้อ 15) และ adverse events (หัวข้อ 19) ทั้ง STROBE และ CONSORT statement จะแนะนำให้เขียน flow chart ของการเข้ามาในการศึกษาของผู้ป่วย ซึ่งจะทำให้เข้าใจได้ง่ายและมีรายละเอียด ครบถ้วนสมบูรณ์ รายละเอียดต่าง ๆ ของ STROBE และ CONSORT statement (เช่น เหตุผลที่มา, ความสำคัญ และตัวอย่างในแต่ละหัวข้อ) จะดูได้ที่ www.strobe-statement.org และ www.consort-statement.org ตามลำดับ

ในที่นี้จะขอยกตัวอย่างการรายงานที่ดี ซึ่งอาจจะไม่ครอบคลุมครบถ้วนในทุกหัวข้อและทุกรูปแบบของการศึกษา

1) Title and Abstract

ชื่อเรื่องควรสอดคล้องกับการศึกษา ไม่ยาวเกินไปและระบุรูปแบบการศึกษาด้วย ส่วนบทคัดย่อควรระบุที่มาของการศึกษา วัตถุประสงค์ วิธีการ ผลและสรุป

ตัวอย่าง : Title

ตัวอย่างที่ 1

ความสัมพันธ์ระหว่างความฉลาดทางอารมณ์กับพฤติกรรมทางเพศของวัยรุ่นในเขตกรุงเทพมหานคร^๑

ผลการรักษา zuclopenthixol accuphase กับ haloperidol และ diazepam ในการรักษาอาการถอนพิษสุรา^๑

Efficacy of donepezil in early-stage Alzheimer disease. A randomized placebo-controlled trial¹⁰

ตัวอย่าง : Abstract

ตัวอย่างที่ 2

วัตถุประสงค์ เพื่อประเมินผลของการเจริญสติในชีวิตประจำวันและพัฒนาแบบวัดสุขภาวะทางจิต

วัตถุประสงค์และวิธีการ การศึกษาทดลองในบุคลากร โรงพยาบาลศรีธัญญาจำนวน 30 คนที่เข้ารับการฝึกปฏิบัติเพื่อเจริญสติในชีวิตประจำวันตามแนวมหาสติปัฏฐานสูตร ระหว่างวันที่ 21-23 พฤษภาคม 2550 ประเมินสุขภาวะทางจิตก่อน

การอบรม และติดตามประเมินผลหลังการอบรม 4 สัปดาห์ด้วยแบบวัดสุขภาวะทางจิตที่ผู้วิจัยสร้างขึ้นเองเป็นแบบประเมินค่า วิเคราะห์ค่าเฉลี่ย ความถี่ และเปรียบเทียบค่าเฉลี่ยของสุขภาวะทางจิตระหว่างก่อนและหลังการฝึกอบรม

ผล มีผู้ตอบแบบสอบถามครบถ้วนจำนวน 25 คน เป็นเพศหญิง 17 คน อายุเฉลี่ย 45.71 ปี (SD=10.43) ระยะเวลาการศึกษาเฉลี่ย 12 ปี (SD=3.85) เกือบทั้งหมดไม่เคยปฏิบัติ

เจริญสติ (21 คน) ได้ข้อคำถามของแบบวัดสุขภาวะทางจิตจำนวน 16 ข้อ ประกอบด้วย 2 ส่วน คือ ส่วนคุณลักษณะจำนวน 8 ข้อ ได้แก่ ความสงบทางใจ ความเบิกบานใจ เป็นต้น และส่วนสมรรถนะทางจิต 8 ข้อ ได้แก่ ความอดทนในการทำงาน ความคาดหวังคนอื่น เป็นต้น ซึ่งมีค่าความสอดคล้องในระดับดี ความแตกต่างของคะแนนเฉลี่ยรวมก่อนและหลังการทดลองของแบบประเมินทั้งฉบับและแต่ละส่วนมีค่าต่างกันอย่างมีนัยสำคัญ

ตารางที่ 1 checklist ใน STROBE statement

Item No.	Recommendation
Title and abstract	1 (a) Indicate the study design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary sheet of what was done and what was found
Introduction	
Background/rationale	2 Explain the scientific background and rationale for the investigation being reported
Objectives	3 State specific objectives, including any prespecified hypotheses
Methods	
Study design	4 Present key elements of study design early in the paper
Setting	5 Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6* (a) Cohort study – Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study – Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls

ตารางที่ 1 checklist ใน STROBE statement (ต่อ)

	Item No.	Recommendation
		Cross sectional study – Give the eligibility criteria, and the sources and methods of selection of participants
	(b)	Cohort study – For matched studies, give matching criteria and number of exposed and unexposed
		Case-control study – For matched studies, give matching criteria and number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12*	<p>(a) Descriptive all statistical methods, including those used to control for confounding</p> <p>(b) Describe any methods used to examine subgroups and interactions</p> <p>(c) Explain how missing data were addressed</p> <p>(d) Cohort study – If applicable, explain how loss to follow-up was addressed</p> <p>Case-control study – If applicable, explain how matching of cases and controls was addressed</p> <p>Cross sectional study – If applicable, describe analytical methods taking account of sampling strategy</p> <p>(e) Describe any sensitivity analyses</p>

ตารางที่ 1 checklist ใน STROBE statement (ต่อ)

	Item No.	Recommendation
Results		
Participants	13	(a) Report numbers of individuals at each stage of study – e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Cohort study – Summarise follow-up time (e.g., average and total amount)
Outcome data	15*	Cohort study – Report numbers of outcome events or summary measures over time Case-control – Report numbers in each exposure category, or summary measures of exposure Cross-sectional study – Report numbers of outcome events or summary measures
Main results	16	(a) Report the numbers of individuals at each stage of the study – e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Other analyses	17	Report other analyses done – e.g., analyses of subgroup and interactions, and sensitivity analyses

ตารางที่ 1 checklist ใน STROBE statement (ต่อ)

	Item No.	Recommendation
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study, if applicable, for the original study on which the present articles is based

ตารางที่ 2 checklist ใน CONSORT statement

Section	Item	Descriptor
Title and abstract	1	How participants were allocated to interventions (e.g., “random allocation,” “randomised,” or “randomly assigned”)
Introduction		
Background	2	Scientific background and explanation of rationale
Methods		
Participants	3	Eligibility criteria for participants; settings and locations where the data were collected
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered
Objectives	5	Specific objectives and hypotheses
Outcome	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors)
Sample size	7	How sample size was determined; explanation of any interim analyses and stopping rules when applicable
Randomisation–sequence generation	8	Methods used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification)
Randomisation–allocation concealment	9	Methods used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until intervention were assigned
Randomisation–implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups
Blinding (masking)	11	Whether participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment
Statistical methods	12	Statistical methods used to compare groups for primary outcome; methods for additional analyses, such as subgroup analyses and adjusted analyses

ตารางที่ ๒ checklist ใน CONSORT statement (ต่อ)

Section	Item	Descriptor
Results		
Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended) – specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome; describe deviations from planned study protocol, together with reasons
Recruitment	14	Dates defining the periods of recruitment and follow-up
Baseline data	15	Baseline demographic and clinical characteristics of each group
Number analysed	16	Number of participants (denominator) in each group included in each analysis and whether analysis was by “intention-to-treat”; state the results in absolute numbers when feasible (e.g., 10/20, not 50%)
Outcome and estimation	17	For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (e.g., 95% CI)
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating which are prespecified and which are exploratory
Adverse events	19	All important adverse events or side effects in each intervention group
Discussion		
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes
Generalisability	21	Generalisability (external validity) of the trial findings
Overall evidence	22	General interpretation of the results in the context of current evidence

2) Introduction

บทนำควรกล่าวถึงเหตุผลที่ทำการศึกษานี้ วัตถุประสงค์ของการศึกษาและสมมติฐานที่ต้องการทดสอบ

3) Methods

วัตถุประสงค์และวิธีการควรกล่าวถึงสิ่งที่ได้ทำในการศึกษานี้ โดยมีรายละเอียดที่เพียงพอเพื่อให้ผู้อื่นเข้าใจ, ตัดสินใจได้ว่าวิธีการศึกษาเหมาะสมกับคำถามการวิจัยหรือไม่ เช่น รูปแบบการศึกษา สถานที่ที่ทำการศึกษา ลักษณะของกลุ่มตัวอย่าง เกณฑ์การคัดเข้าและเกณฑ์การคัดออก การสุ่มตัวอย่าง (sampling) การสุ่มแบ่งกลุ่ม (randomization) วิธีการให้การรักษา ตัวแปรผลการรักษาและตัวแปรอื่น ๆ เครื่องมือที่ใช้ การกำหนดขนาดตัวอย่าง การปกปิด (blinding) และการวิเคราะห์ผลทางสถิติ

ตัวอย่าง : รูปแบบการศึกษา,
วัตถุประสงค์, กลุ่มตัวอย่าง

ตัวอย่างที่ 3

การศึกษานี้เป็นการศึกษาแบบย้อนหลัง เพื่อวิเคราะห์ต้นทุนการบริการสุขภาพจิตของสถานพยาบาล สังกัดกรมสุขภาพจิตทั่วประเทศ จำนวน 17 แห่งของปี 2547 ในทัศนะของผู้ให้บริการ เก็บข้อมูลจากแบบฟอร์มบันทึกข้อมูลที่แก้ไขปรับปรุงมาจากการศึกษาต้นทุนต่อหน่วยบริการสุขภาพจิต ปี 2543 และการศึกษาต้นทุนบริการต่อหน่วยโรงพยาบาลสวนปรุงปีงบประมาณ 2543¹²

ตัวอย่างที่ 4

กลุ่มตัวอย่างคือพนักงานโรงงานอุตสาหกรรมอาหารแห่งหนึ่ง จังหวัดปทุมธานี สุ่มตัวอย่างแบบเจาะจง เลือกระดับการศึกษาไม่สูงกว่าประกาศนียบัตรวิชาชีพ (ปวช.) และสมัครใจตอบแบบสอบถามที่ประกอบด้วยข้อมูลส่วนบุคคล แบบวัดความเครียดฉบับศรีธัญญา (ST-5) แบบสอบถาม Hospital Anxiety Depression scale (HAD) และแบบประเมินความเครียดรายด้าน ประเมินภาวะเครียดภายใน 4 สัปดาห์ เก็บข้อมูลระหว่างเดือนกันยายน - ตุลาคม 2551¹³

ตัวอย่างที่ 5

This 240 week, randomized, double-blind, placebo-controlled study was conducted at 17 sites in the United States. Patients were randomized in an approximately 2:1 ratio to donepezil, 5 mg/d, for the first 6 weeks, with a forced escalation to 10 mg/d thereafter, or placebo. Patients unable to tolerate 10 mg of donepezil were discontinued from the study. Assessments were conducted at screening, at baseline, and at 6-week intervals thereafter through week 24 (final study visit) or at unscheduled termination visits.¹⁰

ตัวอย่างที่ ๖

This randomized, double-blind, double-dummy, placebo-controlled study was conducted at 71 medical centers in 14 countries from January to September 2003. The study compared the efficacy and safety of MFNS 200 µg once daily with MFNS 200 µg twice daily to establish the minimally effective MFNS dose.

The study was conducted in accordance with the Declaration of Helsinki and guidelines on Good Clinical Practice. Subjects who were eligible at screening visit (visit 1) and baseline visit (visit 2) were randomized in a 1:1:1:1 ratio to 1 of 4 treatment arms: Randomization was performed according to a computer-generated code, stratified on the basis of duration of rhinosinusitis symptoms before baseline (7–14 days and 15–28 days).¹⁴

ตัวอย่าง : ตัวแปรผลการศึกษา,
การวัดตัวแปรและเครื่องมือ
ที่ใช้

ตัวอย่างที่ ๗

To determine the relationship between the lung volumes and the estimated fetal weight, we measured the biparietal diameter, head circumference, abdominal

circumference and femur length, from which estimated fetal weight was derived using the Hadlock formula.

To assess intraobserver variability the stored right and left lung volumes of each examination was remeasured on two further separate occasions by the same examiner at each gestational week (18–34 weeks). To assess interobserver variability, 24 scans, selected randomly, were measured three times by a second examiner.¹⁵

ตัวอย่างที่ ๘

The main outcome was change in hemoglobin A_{1c} levels from baseline to 12 months from randomization. We also measured hemoglobin A_{1c} levels at quarterly intervals. The secondary outcomes were changes from baseline to 12 months in the Patient Health Questionnaire–9 score, Hypoglycemia Fear Survey score, Summary of Diabetes Self-Care Activities score, body mass index, and the Diabetes Quality of Life score.¹⁶

ตัวอย่างที่ ๙

Our primary outcome was change in pain level as assessed on a VAS (0 to 100 mm) from baseline to 1 month. Sec-

ondary outcomes were change in measures of hand disability at 1 month, as well as change in pain level and measures of disability at 12 months.

Two independent, trained physicians of physical medicine and rehabilitation (who were not involved in patients' treatment) assessed several clinical and radiographic (structural) variables at enrollment (baseline) and at 1, 6, and 12 months.¹⁷

ตัวอย่างที่ 10

The primary efficacy endpoint was mean AM/PM MSS over days 2 to 15 of the treatment phase.

Secondary efficacy endpoints included mean MSS, TSS, and individual scores (average of AM/PM scores) for each symptom averaged weekly and for days 2 to 15 and 16 to 29. Time to onset of action was assessed and defined as the first day of active treatment on which MSS was statistically significantly different from placebo and sustained thereafter. ...¹⁴

ตัวอย่างที่ 11

เครื่องมือที่ใช้ในการวิจัยได้แก่ แบบสอบถามเกี่ยวกับลักษณะทั่วไปของกลุ่มตัวอย่าง แบบ

ทดสอบกลไกทางจิต Defense Style Questionnaire - 40 ฉบับภาษาไทยโดย ชัยยศ จิตติรังสรรค์ ประกอบด้วย 40 ข้อคำถามได้แก่ mature defense 8 ข้อคำถาม neurotic defense 8 ข้อคำถาม และ immature defense 24 ข้อคำถาม รวมกลไกทางจิต 20 ชนิด (คำถาม 2 ข้อต่อการประเมินกลไกทางจิต 1 ชนิด) และแบบสอบถาม The Five Scale of Self-Esteem for Children (FSC) ฉบับภาษาไทยโดยสุวรรณี พุทธิศรี และ ชัชวาลย์ ศิลปกิจ ประกอบด้วยคำถาม 36 ข้อ ประเมินใน 5 ด้านได้แก่ ด้านมุมมองรวม (global scale) ด้านการศึกษา (academic scale) ด้านภาพลักษณ์ (body image scale) ด้านครอบครัว (family scale) และด้านสังคม (social scale) โดยให้กลุ่มตัวอย่างตอบแบบสอบถามด้วยตนเอง¹⁸

ตัวอย่างที่ 12

A questionnaire was designed for GPs and primary care practice nurses (Table 1). Free-text comments were invited at the end of the survey. A draft questionnaire was piloted among a selection of health professionals including two GP registrars, two consultant gynaecologists, two public health/microbiology laboratory consultants and two nurse practitioners with specialist experience in the provision of sexual health. The final questionnaire was a result of feedback and opinion from these health profes-

sionals. The questionnaire was then sent to all GPs within the county. A total of 281 questionnaires reached their destination. ... The response rate to the successfully sent questionnaires was 61.2%.¹⁹

ตัวอย่าง : Randomization, blinding ตัวอย่างที่ 13

We randomly assigned patients who met the inclusion criteria and agreed to participate. The randomization process was centralized at the coordinating office which had no involvement in enrollment, follow-up, or assessment of participants. A statistician made a computer-generated randomization list (with a block size of 6 and stratified by center) at the coordinating office. Once the screening process was complete, the investigator sent a fax to the coordinating office. The coordinating office randomly assigned the patient to a treatment and faxed the investigator the allocated treatment.¹⁷

ตัวอย่างที่ 14

We prepared a computer-generated randomization list, stratified by hospital site, which used blocks of random sizes (3, 6, 9, and 12) in advance. The

database manager of the Clinical Trials Unit, who was independent of the study, held the list electronically in a password-protected database. Allocation could only be revealed to the recruiting researcher, who then assigned each participant after obtaining consent. We collected all baseline measures before randomization. The laboratory technicians, who were blinded to the allocation throughout the study and the self-reported psychological measures, assessed hemoglobin A_{1c} levels, which was the main outcome. The researchers, study participants, and nurse therapists were not blinded to allocation after randomization because of the nature of delivering psychological treatments.¹⁶

ตัวอย่างที่ 15

The PROSPECT was conducted in 20 primary care practices located in greater New York, ... , from May 1999 to August 2001, with individual patients followed for 2 years. After being paired by urban location, academic affiliation, size, and population type, practices were randomly assigned to the intervention or to usual care by coin flip (cluster randomization by practice). Patients were recruited from an

age-stratified (60–74 years and ≥ 75 years), random sample of patients with upcoming appointments. Research associates confirmed study eligibility (age ≥ 60 years, Mini-Mental State Examination score ≥ 18 and English-speaking) of consenting patients and screened patients for depression by using the Centers for Epidemiologic Studies Depression scale.²⁰

ตัวอย่าง : การคำนวณขนาดตัวอย่าง

ตัวอย่างที่ 16

The study was conducted in Mushin, a high-density inner city area of Lagos in Nigeria with a population of 11,689 school entrants in 76 public primary schools. EPI Info (6.04) statistical package was used to compute the initial sample size of 256 based on the following formula:

Sample size (n) = $k / [1 + (k/\text{population})]$;

where, $k = z^2 p(1-p) / d^2$

z = Z-score corresponding to 95% CI (i.e. 1.96),

p = available local prevalence rate in the target population (which was 2.8% obtained from a comparable local study),

d = margin of error allowed (2%)

By building in an attrition rate of 20% the sample size was increased to 306 as a baseline. The sample was drawn

from 8 (10%) schools chosen by simple random sampling. The first child was randomly selected from the class register and every third child thereafter. This process yielded a total of 361 children. However, parental consent was obtained for 359 who were enlisted for the study. ...²¹

ตัวอย่างที่ 17

To achieve a 15-mm clinical difference in VAS score between control and intervention groups, with a slightly overestimated SD of 26, an α risk for 0.05, and power ($1-\beta$) of 0.80, the number of participants needed in each group was 54. With an estimated 10% of patients lost to follow-up, we sought to include 120 patients but recruited only 112.¹⁷

ตัวอย่างที่ 18

It was established that a total target sample size of ~ 940 subjects (235 subjects/treatment group) would provide 90% power at a 2-sided α level of 0.049 to detect a difference of ≥ 0.7 points in mean AM/PM MSS over days 1 to 15 between treatment groups. Joint power for detecting differences for MFNS μg twice daily versus placebo and versus amoxicillin was at least 81%.¹⁴

ตัวอย่างที่ 19

Sample size was estimated by analyzing previous data from studies comparing verbal rating-scale pain scores between patients receiving a COX-2 selective inhibitor and those receiving placebo after spinal fusion surgery. With 90% power, a mean difference of 2.9, a standard deviation of 1.0, and $\alpha = 0.05$, a power analysis of ANOVA testing on four independent means would require 16 patients per group. To be conservative, we planned to initially enroll 20 patients per group. At the end of the study, if there was a drop-out in any group due to a protocol violation, new patients would be enrolled and randomly assigned to treatment groups to increase the total to 20 per group.²²

ตัวอย่างที่ 20

Sample size calculation was based on a hypothesized 0.8% difference in hemoglobin A_{1c} levels in the motivational enhancement therapy plus cognitive behavior therapy group (or motivational enhancement therapy alone group) compared with usual diabetes care. We assumed that the SD of the changes was approximately 1.65. At a power of 90%, a type I error of 0.05

(2-tailed), a randomization ratio of 1:1:1, and a 20% withdrawal rate, we estimated a sample size of 339 participants ($n=113$ in each group).¹⁶

ตัวอย่างที่ 21

We designed the study as a noninferiority study. A difference of 0.5 has been recognized as the minimum clinically important difference to distinguish treatments on the dyspnea subscale of the CRQ. By using this value in the sample size calculation for a noninferiority study, with an α level of 0.025, $1-\beta$, 0.90, and SD of 1.1 (slightly greater than previous similar studies), the required sample size was 204 (102 patients per group). On the basis of our previous study in a similar patient sample, we anticipated 15% attrition, so we planned to randomly assign 240 patients.²³

ตัวอย่าง : วิเคราะห์ทางสถิติ

ตัวอย่างที่ 22

Efficacy analyses were performed on the intent-to-treat and fully evaluable (FE) populations, whereas safety analyses were performed on the safety population. Study end point was the final visit obser-

vation (i.e., if a patient was missing a week 24 observation, then the last observed value was carried forward and used as the end point visit observation).

The intent-to-treat population included all patients randomized to treatment regardless of follow-up evaluations or compliance with treatments. The FE population comprised patients who completed 24 weeks of treatment with 80% or more compliance at the week 24 visits and at 2 other clinic visits, with no significant protocol violations. ...¹⁰

ตัวอย่างที่ 23

วิเคราะห์ค่าความถี่ ค่าเฉลี่ยของคะแนนภาวะซึมเศร้า เปรียบเทียบค่าคะแนนภาวะซึมเศร้า ก่อนและหลังการให้คำปรึกษาโดยสถิติ Wilcoxon Signed Ranks test การวิเคราะห์ข้อมูลเชิงคุณภาพ โดยการวิเคราะห์เนื้อหาจากคำพูดของผู้รับการปรึกษา จำแนกตามหมวดหมู่ของหลักจริยศาสตร์²⁴

ตัวอย่างที่ 24

Continuous variables were summarized with means and standard deviations, while categorical variables were summarized with frequencies and percentages. Continuous data were analyzed using analysis

of variance if normally distributed, or the nonparametric Wilcoxon rank sums test if nonnormally distributed. The chi-squared test or Fisher's exact test was used for comparison of frequency data where appropriate. ...²⁵

ตัวอย่างที่ 25

For the analysis of differences between independent groups involving Likert scales (e.g., PR residents versus NR residents with regard to general impression of patient care), the evaluation was carried out using the nonparametric Mann-Whitney test for ordinal differences between independent groups. For comparisons of independent groups involving categorical scales (e.g., "yes", "no", "don't know"), the χ^2 test was used. The nonparametric Wilcoxon test for ordinal differences between paired samples was used to compare paired samples involving Likert scales paired data (before the 80-h workweek versus after the 80-h workweek), while McNemar test was used to compare paired samples involving categorical data.²⁶

ตัวอย่างที่ 26

Intra- and inter-observer agreement between measurements was assessed using Bland-Altman plots. Eighteen scans were selected at random and the outlining process repeated by both the same observer (intra-observer agreement) and a different observer (inter-observer agreement).

The relationship between lung volume and gestational age was examined using multilevel models, to account for the correlation between multiple observations on the same subject. Multilevel models were fitted to the data using MLn software (Institute of Education, London, UK). ...²⁷

ตัวอย่างที่ 27

Morphine consumption, pain at rest, pain with movement, heart rate, mean arterial blood pressure, and respiratory rate were compared among the four groups over the postoperative time points, with repeated measures linear fixed model. If group differences were significant ($P < 0.05$), then treatment groups were compared at each time point with Tukey-Kramer post hoc testing.

Sedation scores were compared among the four groups by repeated mea-

sures analysis of variance for assessment of time and treatment effects. If differences were found, a Bonferroni post hoc test was performed.

Total PCA morphine use, duration of surgery, and intraoperative blood loss were compared among the four groups with ANOVA, and Tukey-B post hoc testing. Demographic data were analyzed using ANOVA or chi-square test, as appropriate. The incidence of each side effect was compared with an omnibus chi-square test, and if significant ($P < 0.05$), pairwise post hoc comparisons of each of the three drug groups to the placebo group were made with Fisher's exact test ($P < 0.0167$ for significance).²²

ตัวอย่างที่ 28

Where appropriate, the following tests were used: chi-square for comparing proportions, t tests for comparison of means, and crude odds ratios (ORs) to estimate relative risks. For the more common contact lens (CL)-related complications, binomial logistic regression (for binary outcomes) was used to estimate ORs as measures of association. The ORs are reported as estimates of relative risk

throughout this study. Analyses where OR estimates were adjusted for effect of other covariate (potential confounders) are referred to throughout this paper as multi-variable analysis.

Interactions between CL type and covariates were assessed through classical methods (testing homogeneity of ORs across the covariate strata), and through likelihood ratio tests comparing appropriate logistic models with and without the interaction term. These were limited to interactions thought to be biologically feasible. The regression diagnostics included the Hosmer–Lemeshow goodness-of-fit test, and calculation of the area under the receiver operating characteristic curve. The statistical software used was STATA (Stata Corporation, College Station, TX). The study was approved by the Local Ethics Committee.²⁸

4) Results

ผลการศึกษาจะประกอบไปด้วยส่วนต่าง ๆ เช่น (1) จำนวนคนในแต่ละขั้นตอนของการศึกษา เช่น จำนวนคนที่มีคุณสมบัติเหมาะสม จำนวนคนที่ได้รับการสุ่ม จำนวนคนที่ใช้ในการวิเคราะห์ทางสถิติ (2) flow diagram ของการศึกษา (3) ระยะเวลาที่ติดตามกลุ่มตัวอย่าง เหตุผล

ของ drop-out, loss follow-up (4) การบรรยายและเปรียบเทียบลักษณะของกลุ่มตัวอย่างเมื่อเริ่มต้นก่อนให้การรักษา (5) การรายงานและเปรียบเทียบตัวแปรผลการศึกษาระหว่างกลุ่ม ระหว่างเวลาต่าง ๆ และ (6) ผลการวิเคราะห์ทางสถิติต่าง ๆ

ตัวอย่างการเขียนผลการศึกษาก็ได้กล่าวในบทความอื่นต่อไป

5) Discussion

ในบทวิจารณ์จะมีการสรุปถึงผลที่สำคัญของการศึกษานี้ เปรียบเทียบผลที่ได้จากการศึกษานี้กับการศึกษาในอดีต อภิปรายสาเหตุของความเหมือนและความแตกต่างระหว่างผลที่ได้จากการศึกษานี้และการศึกษาในอดีต ข้อจำกัดของการศึกษานี้และการนำผลจากการศึกษานี้ไปใช้ในประชากรกลุ่มอื่น

ตัวอย่างที่ 29

ข้อจำกัดของการศึกษารั้งนี้ไปด้วยเรื่องความปลอดภัยในพื้นที่ในการเก็บข้อมูล และโรงเรียนหยุดบอยอันเป็นผลจากเหตุการณ์ความไม่สงบ ทำให้มีข้อจำกัดในเรื่องการเป็นตัวแทนของนักเรียนในเขตจังหวัดชายแดนใต้และเครื่องมือที่ใช้เป็นแบบประเมินที่ไม่ใช่เครื่องมือวิจัยโรคและการสัมภาษณ์ทางคลินิก²⁹

ตัวอย่างที่ 30

เนื่องจากไม่มีกลุ่มเปรียบเทียบหรือเปรียบเทียบกับ การปฏิบัติสร้างสุขภาวะทางจิต

ด้วยการปฏิบัติตามแนวอื่น ไม่สามารถสรุปได้ว่า การเจริญสติจะเป็นการปฏิบัติที่ดีที่สุดกว่าแบบอื่น ๆ อย่างไรก็ตามก็ย่อมเป็นไปตามหลักคำสอนว่าต้องเลือกให้เหมาะสมกับความชอบของแต่ละบุคคล ปฏิบัติแล้วสบายมีความสุขจึงจะเป็นวิธีการที่ถูกต้อง ส่วนการติดตามผลในระยะยาวนั้นขึ้นกับความสม่ำเสมอในการปฏิบัติ¹¹

ตัวอย่างที่ 31

การวิจัยครั้งนี้เป็นการศึกษาเฉพาะใน เพศหญิง และไม่ได้แบ่งประเภทและระยะเวลาของการถูกทารุณกรรม นอกจากนี้ข้อมูลได้จากการตอบ ด้วยตนเอง ซึ่งอาจมีข้อจำกัดในข้อความที่เกี่ยวเนื่องกับภาพลักษณ์ของวัยรุ่น เช่น บรรยายภายในครอบครัว มีแนวโน้มที่เด็กบางรายอาจตอบในด้านบวก ทำให้คำตอบเชื่อถือได้ต่ำ¹⁸

เอกสารอ้างอิง

1. Moher D, Schulz KF, Altman DG. For the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Lancet* 2001;357:1191-4.
2. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. For the CONSORT Group. The revised CONSORT Statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
3. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. For the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008; 337:a2390.
4. Vandembroucke JP, Elm EV, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. For the STROBE initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* 2007;147:w-163-w-194.
5. Elm EV, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandembroucke JP and STROBE Initiative. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007; 335:806-8.

6. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clinical Chemistry* 2003;49:1-6.
7. Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney SE, for the SQUIRE development group. Publication guidelines for quality improvement studies in health care: evolution of the SQUIRE project. *BMJ* 2009; 338:a3152.
8. ภัทรา จิรลาภ, นุชนาถ เมฆสีทอง. ความสัมพันธ์ระหว่างความฉลาดทางอารมณ์กับพฤติกรรมทางเพศของวัยรุ่นในเขตกรุงเทพมหานคร. *วารสารสุขภาพจิตแห่งประเทศไทย* 2550;15:22-32.
9. อังกูร ภัทรากร. ผลการรักษา zuclopenthixol accuphase กับ haloperidol และ diazepam ในการรักษาอาการถอนพิษสุรา. *วารสารสุขภาพจิตแห่งประเทศไทย* 2550;15:162-9.
10. Seltzer B, Zolnouni P, Nunez M, Goldman R, Kumar D, Ieni J, Richardson S. For the Donepezil “402” Study Group. Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial. *Arch Neurol* 2004;61:1852-6.
11. ชัชวาลย์ ศิลปกิจ, อรวรรณ ศิลปกิจ. การเจริญสติและแบบวัดสุขภาวะทางจิต. *วารสารสุขภาพจิตแห่งประเทศไทย* 2552;17:252-60.
12. เอนก สุภรินนท์, สุวัฒน์ มหัตนรินทร์กุล. นุปรอ-วรรณ พัวพันธ์ประเสริฐ. ต้นทุนบริการสุขภาพจิตของกรมสุขภาพจิตปี 2547. *วารสารสุขภาพจิตแห่งประเทศไทย* 2550;15:1-9.
13. อรวรรณ ศิลปกิจ, เอมวดี เกียรติศิริ. ความเครียดของพนักงานโรงงานอุตสาหกรรม. *วารสารสุขภาพจิตแห่งประเทศไทย* 2552;17:37-44.
14. Meltzer EO, Bachert C, Staudinger H. Treating acute rhinosinusitis: comparing efficacy and safety of mometasone furoate nasal spray, amoxicillin, and placebo. *J Allergy Clin Immunol* 2005;116:1289-95.
15. Gerards FA, Engels MAJ, Twisk JWR, Van Vugt JMG. Normal fetal lung volume measured with three-dimensional ultrasound. *Ultrasound Obstet Gynecol* 2006;27:134-44.
16. Ismail K, Thomas SM, Maissi E, Chalder T, Schmidt U, Bartlett J, et al. Motivational enhancement therapy with and without cognitive behavior therapy to treat type 1 diabetes. *Ann Intern Med* 2008;149:708-19.
17. Rannou F, Dimet J, Boutron I, Baron G, Fayad F, Mace Y, et al. Splint for base-of-thumb osteoarthritis. *Ann Intern Med* 2009;150:661-9.
18. รัชนี ฉลองเกื้อกุล. ความรู้สึกมีคุณค่าในตนเองและการใช้กลไกป้องกันทางจิตของเด็กที่ถูกทารุณกรรม. *วารสารสุขภาพจิตแห่งประเทศไทย* 2551;16:78-89.
19. Shefras J, Edmondson S, McNulty C. Countywide survey of the current practice of Chlamydia detection in primary care. *J Fam Plann Reprod Health Care* 2002; 28(3):145-8.
20. Gallo JJ, Bogner HR, Morales KH, Post EP, Lin JY, Bruce ML. The effect of a primary care practice-based depression

- intervention on mortality in older adults. A randomized trial. *Ann Intern Med* 2007; 146:689–98.
21. Olusanya BO, Okolo AA, Adeosun AA. Predictors of hearing loss in school entrants in a developing country. *J Postgrad Med* 2004;50:173–9.
22. Reuben SS, Buvanendran A, Kroin JS, Raghunathan K. The analgesic efficacy of celecoxib, pregabalin, and their combination for spinal fusion surgery. *Anesth Analg* 2006;103:1271–7.
23. Ismail K, Thomas SM, Maissi E, Chalder T, Schmidt U, Bartlett J, et al. Effects of home-based pulmonary rehabilitation in patients with chronic obstructive pulmonary disease. *Ann Intern Med* 2008;149:869–78.
24. อุทธีรงค์ หาญรินทร์, สมพร รุ่งเรืองกลกิจ. การให้การปรึกษารายบุคคลแนวพุทธศาสนาในผู้ป่วยเรื้อรังที่มีภาวะซึมเศร้า. *วารสารสุขภาพจิตแห่งประเทศไทย* 2551;16:242–51.
25. Charlton M, Angulo P, Chalasani N, Merriman R, Viker K, Charatcharoenwittaya P, et al. Low circulating levels of dehydroepiandrosterone in histologically advanced nonalcoholic fatty liver disease. *Hepatology* 2008;47:484–92.
26. Biller CK, Antonacci AC, Pelletier S, Homel P, Spann C, Cunningham MJ, et al. The 80-hour work guidelines and resident survey perceptions of quality. *J Surg Res* 2006;135:275–81.
27. Bahmaie A, Hughes SW, Clark T, Milner A, Saunders J, Tilling K, Maxwell DJ. Serial fetal lung volume measurement using three-dimensional ultrasound. *Ultrasound Obstet Gynecol* 2000;16:154–8.
28. Radford CF, Minassian D, Dart J KG, Stapleton F, Verna S. Risk factors for nonulcerative contact lens complications in an ophthalmic accident and emergency department: a case-control study. *Ophthalmology* 2009;116:385–92.
29. เบญจพร ปัญญา, สุลินดา จันทระเสนา. ปัญหาสุขภาพจิตนักเรียนในพื้นที่เกิดเหตุการณ์ความไม่สงบจังหวัดชายแดนใต้. *วารสารสุขภาพจิตแห่งประเทศไทย* 2552;17:1–11.