

การรายงานสถิติ

Bloomberg
Philanthropies



DATA FOR
HEALTH INITIATIVE



ข้อความ ยกมา

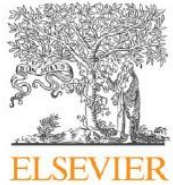
ความจริงก็คือ ปัญหาของการรายงานสถิติที่ไม่ได้มาตรฐานนั้น เป็นสิ่งที่เกิดขึ้นมาอย่างยาวนาน เกิดในวงกว้าง อาจกลายเป็น ปัญหาหนัก โดยส่วนส่วนใหญ่ที่เกี่ยวข้อกับสถิติในระดับพื้นฐาน แต่สิ่งนี้กลับถูกมองข้ามไปโดยผู้อ่านส่วนใหญ่ของ วรณกรรมชีวการแพทย์ (Lang และ Secic 2549)



วัตถุประสงค์ของการเรียนรู้

- หัวข้อนี้เกี่ยวกับแนวทางในการรายงานสถิติเพื่อหลีกเลี่ยงข้อผิดพลาดสำคัญที่พบเป็นประจำ
- เมื่อสิ้นสุดการฝึกอบรมในหัวข้อนี้ ท่านจะได้เรียนรู้วิธีการหลีกเลี่ยงข้อผิดพลาดประจำของการรายงานสถิติ เพื่อ
 - ยกระดับคุณภาพของบทความของท่าน
 - เอื้อต่อความเข้าใจของผู้อ่าน

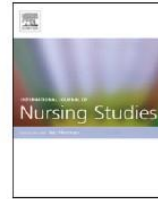
International Journal of Nursing Studies 52 (2015) 5–9



Contents lists available at [ScienceDirect](#)

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns



Basic statistical reporting for articles published in Biomedical Journals: The “Statistical Analyses and Methods in the Published Literature” or the SAMPL Guidelines[☆]



Thomas A. Lang^a, Douglas G. Altman^b

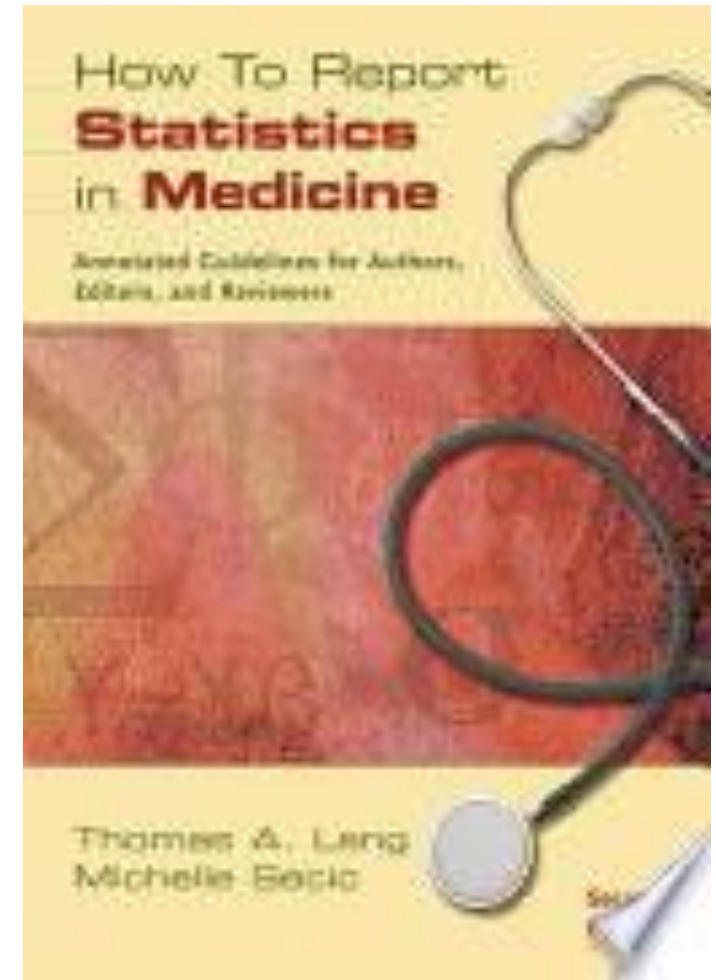
^aTom Lang Communications and Training International, United States

^bCentre for Statistics in Medicine, Oxford University, United Kingdom

Have they reflected that the sciences founded on observation can only be promoted by statistics?... If medicine had not neglected this instrument, this means of progress, it would possess a greater number of positive truths, and stand less liable to the accusation of being a science of unfixed principles, vague and conjectural.

Jean-Etienne Dominique Esquirol, an early French

Although errors have been found in more complex statistical procedures (Burton and Altman, 2004; Mackinnon, 2010; Schwarzer et al., 2000), paradoxically, many errors are in basic, not advanced, statistical methods (George, 1985). Perhaps advanced methods are suggested by consulting statisticians, who then competently perform the analyses, but it is also true that authors are far more





ความแม่นยำในทางตัวเลข

- ข้อมูลที่เป็นตัวเลขนั้น เวลาคำเสนอ ควรบดเศษเป็นตัวเลขกลมๆ แต่ไม่ควรบดเศษเวลาทำการวิเคราะห์ข้อมูล
- ตัวเลขที่มีทศนิยมตั้งแต่สามจุดขึ้นไปนั้น ควรพิจารณาว่าจำเป็นที่จะต้องละเอียดขนาดนั้นหรือไม่



Round data to 2 significant or effective figures

Factor	Cases	Rate	Rate Ratio	P
None	27451	2.345	1.000	Ref
A	34211	3.433	1.464	0.1011
B	11002	5.661	2.414	0.0133
C	5643	6.001	2.559	0.0005

p = p-value

Ref = reference exposure category



Round data to 2 significant or effective figures (2)

Factor	1,000 Cases	Rate	Rate ratio	p
None	27	2.3	1.0	Ref
A	34	3.4	1.5	>0.10
B	11	5.7	2.4	<0.05
C	6	6.0	2.6	<0.001

p = p-value

Ref = reference exposure category



Round data to 2 significant or effective figures (3)

Before

Factor	Cases	Rate	Rate Ratio	P
None	27451	2.345	1.000	Ref
A	34211	3.433	1.464	0.1011
B	11002	5.661	2.414	0.0133
C	5643	6.001	2.559	0.0005

After

Factor	Cases	Rate	Rate ratio	p
None	27	2.3	1.0	Ref
A	34	3.4	1.5	>0.10
B	11	5.7	2.4	<0.05
C	6	6.0	2.6	<0.001

p = p-value

Ref = reference exposure category



ระบุขนาดของตัวอย่างและขนาดของกลุ่มทั้งหมดของการ วิเคราะห์ในแต่ละกรณี

- n = จำนวนของผู้เข้าร่วมการศึกษาวิจัยจากตัวอย่างหนึ่งกลุ่มที่ผู้ศึกษาวิจัยต้องการศึกษาวิจัย
 - ในการใช้ทั่วไป (ถึงแม้จะไม่ถูกต้อง) n คือ ขนาดของกลุ่มต่าง ๆ หรือตัวอย่างทั้งหมดในการศึกษาวิจัย
- N = จำนวนของผู้เข้าร่วมจากประชากรหนึ่งกลุ่มที่ผู้ศึกษาวิจัยต้องการศึกษาวิจัย
 - ในการใช้ทั่วไป (ถึงแม้จะไม่ถูกต้อง) N คือ ขนาดของตัวอย่างทั้งหมดในการศึกษาวิจัยหนึ่ง ๆ



การนำเสนอร้อยละ

- หลีกเลี่ยงการใช้ร้อยละเพื่อการนำเสนอตัวอย่างขนาดเล็ก
 - จุดตัดที่ $n = 20$ เพื่อสื่อถึงตัวอย่างขนาดเล็กนั้นเป็นตัวเลขที่เหมาะสม แต่ไม่มีหลักการรองรับ
- ระบุเศษและตัวหารในการคำนวณเสมอ
 - ร้อยละ 25 (650/2,598)
 - ร้อยละ 33 (ผู้ป่วยจำนวน 30 รายจากจำนวน 90 ราย)
 - กระจายจำนวน 12 ตัวจากจำนวน 16 ตัว (ร้อยละ 75)



การนำเสนอร้อยละ (2)

- ใช้สูตรนี้ในการคำนวณการเปลี่ยนแปลงต่างๆ ที่เกิดขึ้นในข้อมูล

(ค่าสุดท้าย – ค่าเบื้องต้น) / ค่าเบื้องต้น

- การที่อุณหภูมิของร่างกายเพิ่มขึ้น 10 องศาเซลเซียสจาก 30 องศาเซลเซียสเป็น 40 องศาเซลเซียสนั้นเป็นการเพิ่มขึ้นร้อยละ 33

$$(40-30)/30=0.33$$

- การที่อุณหภูมิของร่างกายลดลง 10 องศาเซลเซียสจาก 40 องศาเซลเซียสเป็น 30 องศาเซลเซียสนั้นเป็นการลดลงร้อยละ 25

$$(30-40)/40=-0.25$$



การนำเสนอร้อยละ (3)

- ในการรายงานการเปลี่ยนแปลงต่างๆ ที่เกิดขึ้นในข้อมูลเป็นร้อยละนั้น จงใช้คำอย่างระมัดระวัง
 - จำนวนผู้เสียชีวิตโดยประมาณ...ได้ลดลงร้อยละ 12 จาก 20,356 ราย เหลือ 17,849 ราย
 - จำนวนผู้เสียชีวิตโดยประมาณ ... ได้ลดลงเหลือร้อยละ 12



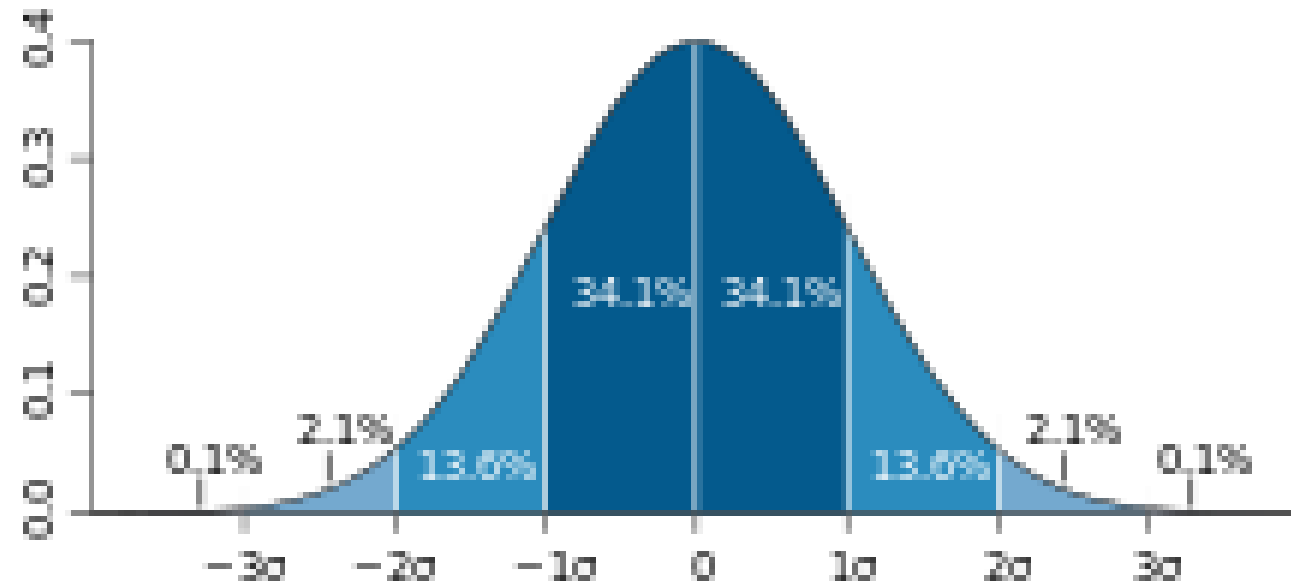
การสรุปข้อมูล

- นำเสนอข้อมูลเป็นตารางและแผนภาพ
 - ตารางจะนำเสนอค่าที่วัดได้จริง (exact values)
 - แผนภาพจะนำเสนอการประเมินโดยรวมจากข้อมูล (ทำหน้าที่เล่าเรื่อง)
- ห้ามกล่าวข้อมูลซ้ำลงในเนื้อความ
- เนื้อความควรมีเพียง 1-2 ประโยคที่ชี้ไปถึงส่วนสำคัญของบทความ สรุปผลที่พบ หรือแบบแผนสำคัญที่อยู่ในตารางหรือแผนภาพ
- เราจะอภิปรายเรื่องตารางและแผนภาพในสัปดาห์ที่ 3 และ 4 ของการฝึกอบรม



การสรุปข้อมูล (2)

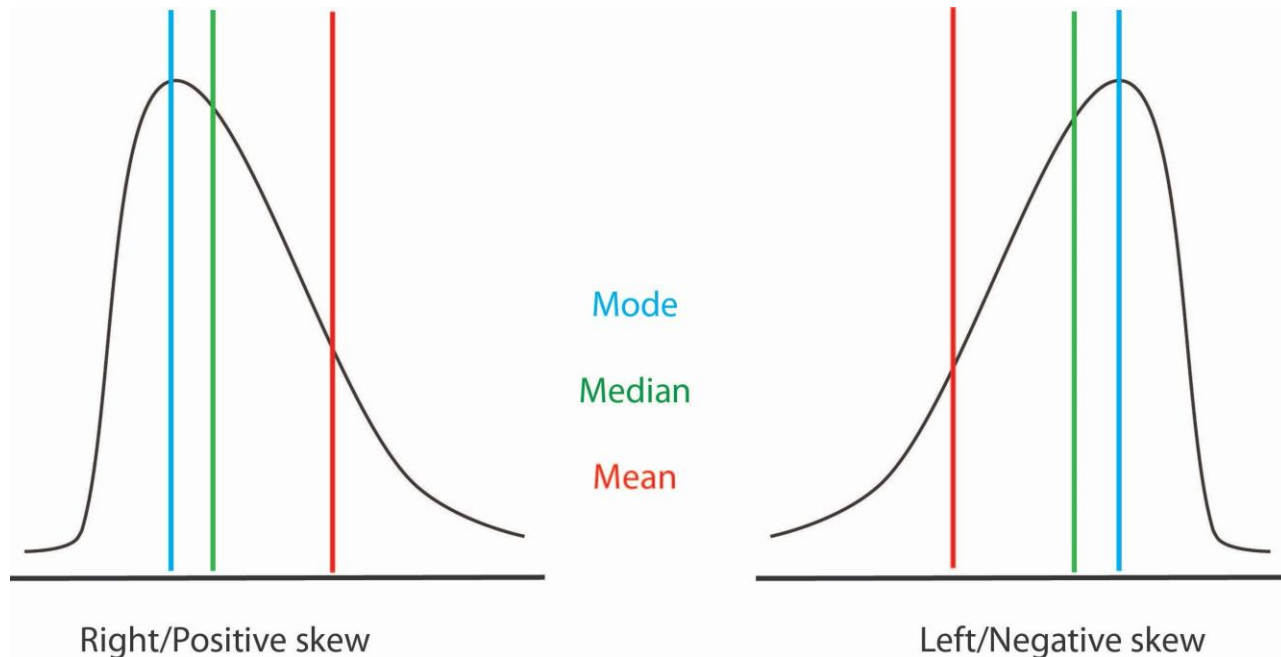
- สำหรับข้อมูลที่มีการแจกแจงแบบปกติโดยประมาณนั้น
 - ใช้ค่าเฉลี่ยและการเบี่ยงเบนมาตรฐาน (SD)
 - ใช้สูตรนี้ ค่าเฉลี่ย (SD) ไม่ใช่ค่าเฉลี่ย \pm SD (เครื่องหมาย " \pm " นั้นไม่จำเป็น)
 - ค่าเฉลี่ย (SD) คือ 12 มล (2 มล)
 - ห้ามใช้ค่าความคลาดเคลื่อนมาตรฐาน (SEM)





การสรุปข้อมูล (3)

- สำหรับข้อมูลที่มีการแจกแจงแบบไม่เป็นปกติ นั้น ให้ใช้
 - ค่ามัธยฐาน (ค่ากลางของข้อมูลชุดนั้นๆ)
 - ค่าพิสัย (ค่าต่ำสุดและสูงสุด)
 - ค่าพิสัยระหว่างควอไทล์ (ค่าที่เปอร์เซ็นต์ไทล์ที่ 25 และ 75)





การสรุปข้อมูล (4)

- ตัวอย่างของสรุปข้อมูลที่มีการแจกแจงแบบไม่เป็นปกติ
 - น้ำหนักมัธยฐานคือ 72 กก (ค่าที่เปอร์เซ็นต์ไทล์ที่ 25 คือ 60 กก ค่าที่เปอร์เซ็นต์ไทล์ที่ 75 คือ 87 กก)
 - น้ำหนักมัธยฐานคือ 72 กก (ค่าพิสัยระหว่างควอไทล์คือ 60 ถึง 87 กก)
 - ภายหลัง 8 สัปดาห์ น้ำหนัก (ค่ามัธยฐานและค่าพิสัยระหว่างควอไทล์) เท่ากับ 72 กก (60 ถึง 87 กก)



สัดส่วน อัตราและอัตราส่วน

- โปรดอ้างอิงถึงบทที่ 3 ของ *Principles of Epidemiology in Public Health Practice* เพื่อการคำนวณความเสี่ยง อัตราและอัตราส่วน
- ใช้เส้นเอียง (/) เพื่อสื่อถึงสัดส่วนและอัตราต่างๆ
 - อัตราการติดเชื้อเฉลี่ยที่ 50/100,000 คน
- ใช้เครื่องหมายทวิภาค (:) เพื่อสื่อถึงอัตราส่วนต่างๆ
 - อัตราส่วนระหว่างผู้ชายกับผู้หญิงนั้นเท่ากับ 3:4
- ใช้ตัวเขียนเมื่อสื่อถึงเศษส่วนที่กำหนดที่ขยายคำนาม
 - ครึ่งหนึ่งของจำนวนผู้ป่วย
 - ส่วนใหญ่ในจำนวนสองในสาม



การรายงานความเสี่ยง อัตราและอัตราส่วน

- ระบุประเภทของอัตรา (อาทิ อัตราอุบัติเหตุ อัตราการรอดชีวิต) อัตราส่วน (อาทิ อัตราส่วนของ การจะเกิดขึ้นและไม่เกิดขึ้น) หรือความเสี่ยง (อาทิ ความเสี่ยงสัมบูรณ์)
- ระบุปริมาณที่น่าเสนอเป็นพิเศษและส่วนหาร
- ระบุช่วงเวลาที่นำแต่ละอัตรามาใช้
- ระบุหน่วยประชากรใด ๆ (อันได้แก่ ตัวคูณหน่วย อาทิ $\times 100$ $\times 10,000$)
- พิจารณาที่จะรายงานหน่วยวัดความแม่นยำ (ช่วงความเชื่อมั่น) สำหรับความเสี่ยง อัตราและ อัตราส่วนต่าง ๆ โดยประมาณ



เรื่องกวนใจของ Kim

- การไม่ใช้เครื่องหมายลูกน้ำสำหรับหลักพันขึ้นไป
 - ห้ามใช้ช่องว่าง (12 345)
 - ห้ามใช้จุด (12.345)
 - จงอย่าละเลยเครื่องหมายลูกน้ำ (12345)
- การไม่นำเสนอสถิติให้สอดคล้องต่อเนื่องกัน
- การเสนอข้อมูลซ้ำกันทั้งในเนื้อความ ในตารางและในแผนภาพ
- การไม่นำเสนอเศษหรือตัวหาร
- การไม่ระบุขนาดของตัวอย่าง



ขั้นตอนต่อไป

1. ทบทวนบทความเรื่องการรายงานสถิติขั้นพื้นฐานโดย Lang และ Altman (ซึ่งอยู่ในเอกสารประกอบการเข้าร่วมการฝึกอบรมของท่าน)
2. ทบทวนบทความของท่านและตรวจสอบว่ามี การปฏิบัติตามแนวทางต่างๆ ที่ได้นำเสนอไว้ในหัวข้อนี้