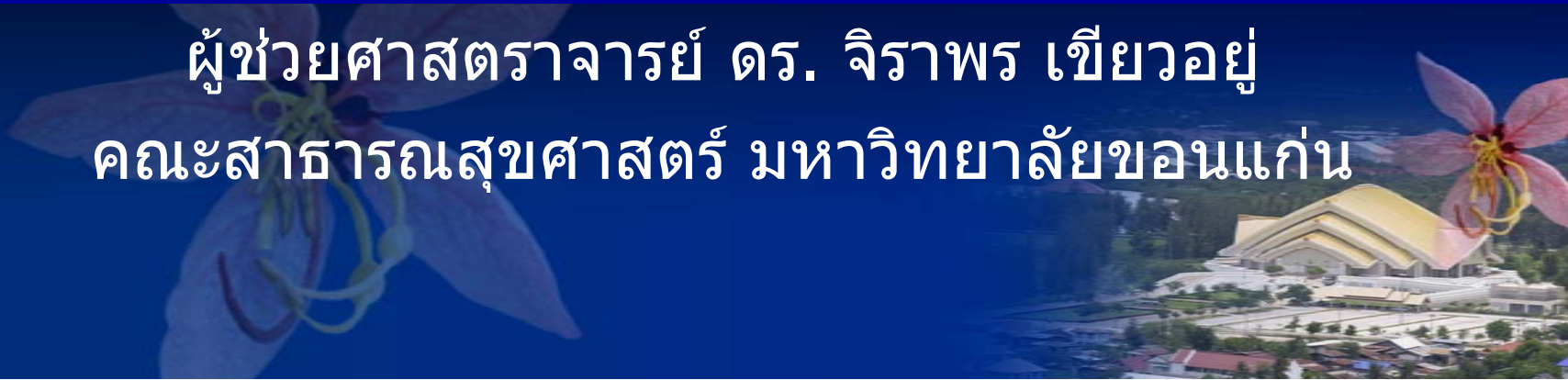




การพัฒนาเครื่องมือและ การทดสอบเครื่องมือการวิจัย

ผู้ช่วยศาสตราจารย์ ดร. จิราพร เขียวอยู่
คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น





ข้อพิจารณาก่อนสร้างเครื่องมือ

- ข้อมูลอะไร? ที่ต้องการทราบอย่างแท้จริง
- จะเก็บข้อมูลจากใคร? และจะใช้วิธีใด?
- ระดับการศึกษาของผู้ให้ข้อมูลส่วนใหญ่เป็นอย่างไร?
- ขนาดตัวอย่างเท่าใด?
- จะจัดการกับข้อมูล (data processing) นั้นอย่างไร?





วิธีการเก็บรวบรวมข้อมูลและเครื่องมือที่ใช้

วิธีการเก็บข้อมูล	เครื่องมือที่ใช้
<input type="checkbox"/> ข้อมูลที่มีอยู่แล้ว	<input type="checkbox"/> แบบรวบรวมข้อมูล, Checklist
<input type="checkbox"/> สังเกต	<input type="checkbox"/> แบบบันทึกการสังเกต, ประสาทสัมผัส
<input type="checkbox"/> การสัมภาษณ์	<input type="checkbox"/> แบบสัมภาษณ์, Checklist, Tape recorder
<input type="checkbox"/> การสอบถาม	<input type="checkbox"/> การสอบถาม



ข้อดี-ข้อด้อยของวิธีการเก็บข้อมูลจากข้อมูลที่มีอยู่แล้ว

ข้อดี

- ไม่แพง

ข้อด้อย

- เข้าถึงข้อมูลยาก อาจขัดจริยธรรม
- ข้อมูลอาจไม่สมบูรณ์
- ไม่ตรงกับความต้องการ



ข้อดี-ข้อด้อยของวิธีการเก็บข้อมูลด้วยการสังเกต

ข้อดี

- ได้รายละเอียด
มาก ลึก

ข้อด้อย

- อาจขัดจริยธรรม หรือละเมิดสิทธิส่วนบุคคล
- ข้อมูลที่ได้อาจบิดเบือนเนื่องจากการปรากฏตัว
ของผู้เก็บข้อมูล
- ต้องการ การฝึกอบรมสูง และตลอดเวลา



ข้อดี-ข้อด้อยของวิธีการเก็บข้อมูลโดย การสัมภาษณ์แบบตัวต่อตัว (Face to face)

ข้อดี

- เหมาะสำหรับผู้ที่อ่าน-เขียนไม่ได้
- Response rate สูงกว่าการใช้แบบสอบถาม

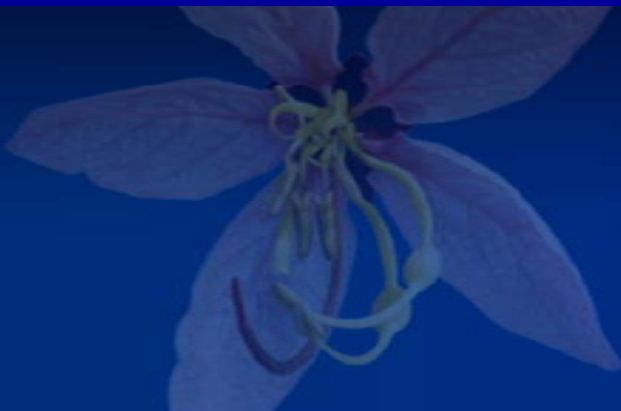
ข้อด้อย

- การปรากฏตัวของผู้สัมภาษณ์ อาจมีผลต่อข้อมูลที่ได้รับ
- ข้อมูลที่ได้มีความสมบูรณ์น้อยกว่า การได้ข้อมูลจากการสังเกต



ข้อดี-ข้อด้อยของวิธีการเก็บข้อมูลโดย การสัมภาษณ์แบบเจาะลึก (In-depth interview)

ข้อดี	ข้อด้อย
<input type="checkbox"/> ได้รายละเอียดมากขึ้น	<input type="checkbox"/> ใช้เวลานาน <input type="checkbox"/> วิเคราะห์ข้อมูลยาก เพราะใช้ คำถามปลายเปิด (เป็น semi-structure questionnaire)





ข้อดี-ข้อด้อยของวิธีการเก็บข้อมูลโดย การทดสอบแบบสอบถาม (Administering written)

ข้อดี

- ไม่แพง
- อาจได้คำตอบที่เป็นจริง
ซื่อสัตย์
- ไม่ต้องใช้พนักงานสัมภาษณ์

ข้อด้อย

- ใช้ได้กับกลุ่มผู้อ่านออกเขียนได้
- Response rate ต่ำ

มีวิธีเพิ่ม response rate

การออกแบบสอบถาม





ประเด็นที่ต้องคำนึงถึงในการออกแบบสอบถาม

- เนื้อหาและความยาวของแบบสอบถาม
- ลำดับคำถาม และการจัดเรียง
- การปฏิเสธการให้ข้อมูล
- อัตราการตอบ-การให้ข้อมูลที่สมบูรณ์
- รูปแบบและการจัดวางคำถาม



ขั้นตอนในการออกแบบสอบถาม 1)

□ พิจารณาเนื้อหา

- ทบทวนวรรณกรรม
- ความเห็นจากผู้เชี่ยวชาญ
- ความเห็นจากกลุ่มเป้าหมาย





ขั้นตอนในการออกแบบสอบถาม 2)

□ ตั้งคำถาม & เลือกประเภทคำถามที่ใช้

- ไม่ควรถาม 2 ประเด็นในคำถามเดียวกัน
- ไม่ควรใช้ข้อความปฏิเสธซ้อน
- ไม่ใช้ข้อความที่กำกวม
- ไม่ควรถามนำ
- หลีกเลี่ยงคำแสลง
- ประเภทของคำถามอาจเป็น ปลายเปิด หรือปลายปิด



ประเภทคำถาม—ข้อดี/ข้อด้อย

Open-ended questions

จงระบุสาเหตุที่ทำให้ท่านเกิดความเครียด

.....
.....

Closed-ended questions

สถานภาพสมรสในปัจจุบันของท่านคือ

- โสด
- คู่
- หม้าย
- หย่า
- แยกกันอยู่

- Completion
(Fill-in the- blank)

- ปัจจุบันท่านอายุ
.....ปี





ข้อควรพิจารณาเพิ่มเติมในการตั้งคำถาม

□ เรื่องที่มีความอ่อนไหว

- ควรถามโดยยกสถานการณ์ ไม่ให้ผู้ตอบรู้สึกว่าเป็นข้อมูลของตัวเอง

□ เรื่องที่เกี่ยวข้องกับความจำ

- มีกรอบเวลาที่ชัดเจน --- ควรทบทวนวรรณกรรม
- มีคำถามอื่นนำก่อน



ข้อควรพิจารณาเพิ่มเติมในการตั้งคำถาม

□ คำถามที่มีคำตอบได้มากกว่า 1 คำตอบ ควรมีรูปแบบอย่างไร

ท่านได้รับข้อมูล...จากสื่อใดบ้าง (ตอบได้มากกว่าหนึ่งข้อ)

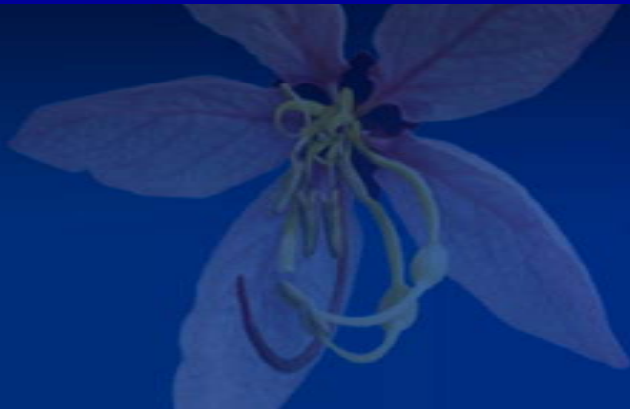
	ได้รับ	ไม่ได้รับ
โทรทัศน์	<input type="checkbox"/>	<input type="checkbox"/>
วิทยุ	<input type="checkbox"/>	<input type="checkbox"/>
หนังสือพิมพ์	<input type="checkbox"/>	<input type="checkbox"/>
เพื่อน	<input type="checkbox"/>	<input type="checkbox"/>

□ คำถามที่ให้เรียงลำดับมากที่สุด/บอยที่สุด-น้อยที่สุด



ขั้นตอนในการออกแบบสอบถาม 3)

- เรียงลำดับคำถาม
- กำหนดรูปแบบของแบบสอบถาม
- ร่างแบบสอบถาม





ขั้นตอนในการออกแบบสอบถาม 4)

- ขัดเกลาภาษา ปรับปรุงแบบสอบถาม
- วิเคราะห์คำถามในแบบสอบถาม
- สร้างคู่มือแบบสอบถาม
- สร้างคู่มือลงรหัส
- ทดสอบแบบสอบถาม
- แก้ไขและจัดพิมพ์



การสร้างคู่มือแบบสอบถาม

□ อธิบายรายละเอียดของ

- วิธีการเก็บข้อมูล
- ลักษณะตัวอย่าง
- ความหมายของข้อมูล / ตัวแปร แต่ละข้อ

ตัวอย่าง-งานวิทยานิพนธ์ของคุณเปรมกมล ขวนขวาย(ได้รับอนุญาต)
เรื่อง "ปัจจัยที่เกี่ยวข้องกับการหกล้มของผู้สูงอายุที่อาศัยอยู่ในชุมชน"



การสร้างคู่มือรหัส

ลำดับข้อ/ หน้า	ชื่อตัวแปร/ คำอธิบาย	ประเภทข้อมูล	ความกว้าง	รหัส(ค่าที่เป็นไปได้)/ความหมายรหัส
หน้าปก	ID/ เลขที่ แบบสอบถาม	Number	3	001-600 แต่ละ ID เป็นข้อมูลของตัวอย่างแต่ละราย
1/1	Sex/เพศผู้ตอบ	Number	1	1=ชาย 2=หญิง 9=missing
2/1	Age/อายุผู้ตอบ	Number	2	15-60(ลงอายุเป็นปีเต็ม) 99=missing
:				
:				
10/2	ความพึงพอใจต่อ	Number	1	1=ไม่พึงพอใจเลย 2=ไม่พึงพอใจ 3=เฉยๆ 4=พึงพอใจมาก 5=พึงพอใจมากที่สุด 8=ไม่เข้าข่าย เช่น ไม่เคยใช้บริการ 9=missing



การทดสอบแบบสอบถาม

- ทดสอบกับตัวอย่าง ที่มีลักษณะคล้ายคลึงกับตัวอย่างที่ใช้ในการศึกษา
- ผลจะชี้ข้อบกพร่องต่าง ๆ ที่ควรแก้ไข
 - ความไม่ชัดเจนของคำถาม
 - การเรียงลำดับ
 - ความยาวของแบบสอบถาม



ประเด็นอื่นที่ควรคำนึงถึง: การเพิ่มอัตราตอบกลับ

- เป็นแบบสอบถามที่ดี
- มีจดหมายนำที่ดี
 - มีคำอธิบายถึงความสำคัญของข้อมูลที่ต้องการ
 - ลงนามโดยผู้มีอำนาจ
- กรณีสอบถามทางไปรษณีย์ ควรเตรียมซองและแสตมป์สำหรับการส่งกลับให้พร้อม
- มีจดหมายตาม (ไม่ใช่โทรศัพท์)



ประเด็นที่ควรทราบ: อคติในการเก็บรวบรวมข้อมูล

- เครื่องมือไม่ดี
- ผู้เก็บข้อมูลมีอคติ
- ผลของวิธีการที่ใช้ต่อผู้ให้ข้อมูล

ข้อพิจารณาด้านจริยธรรม

การสร้างสเกล





ขั้นตอนการสร้างสเกล

□ กำหนด(โครงสร้าง)เนื้อหาของสเกล

□ การนิยามตัวแปร

□ การรวบรวมรายการคำถาม

□ การออกแบบสเกล

□ การประเมินความตรงตามเนื้อหา

□ การประเมินความชัดเจนของภาษา

□ การลดรายการคำถาม

□ การประเมินความตรงเชิงโครงสร้างของสเกล

□ การประเมิน internal consistency

□ หาข้อสรุปในการนำไปใช้

- Adjust from previous study
- New construct based on theory/content
- Interview experts
- Focus group



ขั้นตอนการสร้างสเกล (ต่อ)

□ การออกแบบสเกล

□ Process

- ระบุชุดของ item
- ตรวจสอบเพื่อยืนยันด้านเนื้อหา

□ Issues to be considered

- กลุ่มตัวอย่าง และ ขนาดตัวอย่าง
- การเลือกใช้ข้อความทางลบ
- จำนวน item
- Scaling of items

Scaling method, Scaling response



Scaling method, Scaling response

สเกลประเภทต่าง ๆ

- Likert Scale
- Thurstone Scale
- Guttman Scale
- Semantic Differential Scale

การกำหนดตัวเลือกให้กับคำถามแต่ละข้อ

- Scaling method
- Resources
- Scale administration
- Characters of target group



ขั้นตอนการสร้างสเกล (ต่อ)

- การประเมินความตรงตามเนื้อหา
- การประเมินความชัดเจนของภาษา
- การลดรายการคำถาม
- การประเมินความตรงเชิงโครงสร้างของสเกล
- การประเมิน internal consistency

Item analysis

- Distribution of response
- Factor analysis
- Inter-item correlation
- Item-total correlation
- Inter-item consistency (internal consistency)



ขั้นตอนการสร้างสเกล (ต่อ)

- Reliability assessment
 - Cronbach's Alpha coefficient
- Criterion-related validity—ถ้าทำได้

□ หาข้อสรุปในการนำไปใช้

- การแปลความหมาย
- วิธีการนำไปใช้





การตรวจสอบคุณสมบัติ ของเครื่องมือ





ประเด็นการนำเสนอ

- เครื่องมือประเภทมาตรวัด/สเกล
 - ทำไมจึงต้องมีการตรวจสอบคุณภาพของเครื่องมือ?
 - ถ้าจะตรวจสอบจะตรวจสอบอะไรบ้าง?
 - มีวิธีการตรวจสอบและเกณฑ์การประเมินอย่างไร?
 - ปัญหาที่มักพบในการตรวจสอบคุณภาพเครื่องมือ
-
- **การตรวจสอบคุณสมบัติของเครื่องมือ(เพิ่มเติม)**



เครื่องมือประเภทสเกล(มาตรวัด)

รายการข้อคำถาม	ระดับความเห็น				
	1	2	3	4	5
คำถามข้อ 1.....					
คำถามข้อ 2.....					
.....					
.....					
คำถามข้อ 10.....					



ทำไม?

จึงต้องมีการตรวจสอบเครื่องมือ

งานวิจัย--ต้องการข้อมูลที่มีคุณภาพ

คุณภาพข้อมูล
ขึ้นอยู่กับ
อะไรบ้าง?

คุณภาพของเครื่องมือ

คุณภาพของคนเก็บข้อมูล

คุณภาพของวิธีดำเนินการเก็บข้อมูล

คุณภาพของการจัดการข้อมูล



ถ้าจะตรวจสอบเครื่องมือ จะตรวจสอบอะไรบ้าง?

แบบวัดทัศนคติ (ตัวเลือก 1 2 3 4 5) ในการทดสอบ
เครื่องมือ ควรหาเฉพาะ Cronbach-Alpha เท่านั้นใช้หรือไม่?

การทดสอบเครื่องมือที่เขานิยมกันนั้น ไม่ว่าจะเป็
□ ความรู้ วัดออกมาเป็นตอบถูกได้ 1, ตอบผิดได้ 0
□ การวัดทัศนคติ วัดออกมาเป็น rating 1 2 3 4 5
จากเห็นด้วยน้อยที่สุด->มากที่สุด

ควรจะทดสอบอะไรบ้าง



ถ้าจะตรวจสอบเครื่องมือ จะตรวจสอบอะไรบ้าง

Validity

- ❑ Content validity
- ❑ Construct validity
- ❑ Criterion validity

Reliability

- ❑ Rater reliability
- ❑ Stability/Test-retest
- ❑ Alternate form reliability
- ❑ Internal consistency



การตรวจสอบ Validity

Content Validity

- Expert opinion
- CVI-Content validity index

Rates the content relevance on a four-point scale;
1-totally irrelevant content; 4-totally relevant content
CVI-the proportion of judges who rate the item 3 or 4

Number of raters and CVI:

5 rater ----CVI => 1

6-10 raters ----CVI=> 0.83-0.90 (no more than 1 rater – score an item less than 3)

Lynn 1986

Nursing research
35, 382-5



การตรวจสอบ Reliability

Internal consistency

Cronbach's Alpha Coefficient 0.80-0.90

ค่าสัมประสิทธิ์แอลฟา	ระดับการยอมรับ
> 0.90	ดีมาก แต่....
0.85-0.90	ดีมาก
0.80-0.84	ดี
0.75-0.79	ค่อนข้างดี
0.70-0.74	พอใช้
<0.70	จำเป็นต้องแก้ไข

Output of Reliability Analysis with SPSS : Cronbach's Alpha Coefficient

RELIABILITY ANALYSIS - SCALE (ALPHA)

N of Cases = 30.0
 Statistics for Mean Variance Std Dev N of Variables
 Scale 36.2000 33.5448 5.7918 10

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item- Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
X1	33.7000	34.2172	-.1432	.3656	.7999
X2	32.3667	27.5506	.6936	.7746	.7024
X3	32.6667	25.1264	.5638	.5400	.7038
X4	32.4667	26.8782	.5897	.6273	.7059
X5	32.9333	27.0989	.3585	.2714	.7408
X6	32.2667	30.1333	.3864	.6253	.7357
X7	32.2000	25.4069	.7241	.8013	.6849
X8	32.1667	28.2126	.4082	.6992	.7300
X9	32.6000	25.8345	.4140	.5514	.7330
X10	32.4333	28.8057	.3640	.5070	.7359

Reliability Coefficients 10 items

Alpha = .7499 Standardized item alpha = .7700

Output of Reliability Analysis with Stata: Cronbach's Alpha Coefficient

```
. alpha x1- x10
```

```
Test scale = mean(unstandardized items)
```

```
Reversed item: x1
```

```
Average interitem covariance: .2890932
```

```
Number of items in the scale: 10
```

```
Scale reliability coefficient: 0.7829
```

```
. alpha x1- x10, item
```

```
Test scale = mean(unstandardized items)
```

Item	Obs	Sign	item-test correlation	item-rest correlation	average inter-item covariance	alpha
x1	30	-	0.3038	0.1432	.3378991	0.7999
x2	30	+	0.7600	0.7001	.2810664	0.7444
x3	30	+	0.6295	0.4825	.2729246	0.7608
x4	30	+	0.7547	0.6744	.2658365	0.7386
x5	30	+	0.4832	0.2965	.3024266	0.7895
x6	30	+	0.5449	0.4573	.3110473	0.7676
x7	30	+	0.7970	0.7247	.2560026	0.7309
x8	30	+	0.6291	0.5131	.2829821	0.7569
x9	30	+	0.5076	0.3096	.2960409	0.7909
x10	30	+	0.6274	0.5147	.2847063	0.7571
Test scale					.2890932	0.7829



ตรวจสอบการกระจายของคำตอบ

Frequency of each category

- Delete Item ---- >0.95 or <0.05
- In practice ---- $0.2-0.8$

A Simple item discrimination index

$$d_i = \frac{U_i - L_i}{n_i}$$

U_i = no. of people above the median who score positive on item i ,

L_i = no. of people below the median who score positive on item i ,

n_i = no. of people below (or below) the median



ปัญหาที่มักพบในการ ตรวจสอบคุณภาพเครื่องมือ

ไม่ทราบว่า

- จะตรวจสอบอะไรบ้าง
- จะตรวจสอบอย่างไร
- จะใช้โปรแกรม/คำสั่งอะไร
- จะแปลผลอย่างไร
- และจะแก้ไขอย่างไร



ตัวอย่างปัญหา

1. แบบทดสอบที่เราวัดความรู้ ที่เป็นตัวเลือก 1 2 3 4 ในการทดสอบ เครื่องมือนั้น ควรทดสอบความเชื่อมั่นโดยใช้ Cronbach-Alpha ได้ (ตัวทดสอบวัดมากกว่า 2 ระดับ น่าจะใช้ได้) หาค่าความยากง่าย หาค่าอำนาจการจำแนก
2. แบบทดสอบที่เราวัดความรู้ (ใช่ /ไม่ใช่ ถูก/ไม่ถูก) ในการทดสอบ เครื่องมือสามารถทำได้เหมือนกับที่มีตัวเลือกเป็น 1 2 3 4 ได้หรือไม่ครับ
3. แบบวัดทัศนคติ (ตัวเลือก 1 2 3 4 5) ในการทดสอบเครื่องมือ ควรหาเฉพาะ Cronbach-Alpha เท่านั้นหรือไม่
4. การทดสอบเครื่องมือที่เขานิยามกันนั้น ไม่ว่าจะเป็ ความรู้ วัดออกมา 1 2 3 4 หรือ ใช่ /ไม่ใช่ การวัดทัศนคติ นักวิจัยควรจะทดสอบเครื่องมืออะไรบ้างครับที่เขานิยามกัน



เครื่องมือประเภทแบบทดสอบ

- ❑ Content validity

- ❑ IOC

- ❖ Index of consistency
 - ❖ Index of Item-Objective congruence

- ❑ Difficulty index

- ❑ Discrimination index

- ❑ Internal consistency

p ---- 0.2-0.8

$p_{27\%บน} - p_{27\%ล่าง}$

n ของกลุ่มที่มากที่สุด

0.2 - 0.8

KR20 หรือ KR21



ประเด็นเพื่อ discuss

แบบสอบถามที่เราสร้างขึ้น ถ้าไม่มีส่วนที่เป็น สเกล
จำเป็นต้องทดสอบคุณสมบัติของเครื่องมือหรือไม่

- ถ้าตอบว่า ไม่ ขอทราบเหตุผล
- ถ้าตอบว่าจำเป็น ถามว่าต้องทดสอบคุณสมบัติใดบ้าง



การตรวจสอบคุณสมบัติ ของเครื่องมือ (เพิ่มเติม)





Validity Test

Types

- Content validity →
- Construct validity →
- Criterion-based validity
 - Concurrent validity →
 - Predictive validity →

Methods of Assessment

- Expert, CVI
- Factor analysis,
- Test of Hypothesis
- Correlation, Agreement
- Regression



Validity of Diagnostic Test

Properties of Diagnostic Test

- Sensitivity
- Specificity
- Predictive validity
- Accuracy
- Likelihood Ratio

□ Criterion-Based Validity ----- Kappa



Reliability Test

Types

- Rater reliability
 - Inter-rater
 - Intra-rater
- Internal consistency
- Stability or Test-retest reliability
- Alternate forms reliability or Parallel form reliability

Reliability coefficient

- ICC, Kappa
- Cronbach's Alpha Coefficient
- Item-to-total correlation
- ICC, CCC, Bland-Altman Plot
- Kappa
- Correlation coefficients
- CCC, Kappa



Responsiveness

- ❑ The ability of a measure to reflect underlying change
- ❑ 4 statistics used to index responsiveness
 - Effect size
 - T-test comparisons
 - The standardized response mean
 - The responsiveness statistics

Staquet MJ, Hays RD,
Fayers PM. 1998. P.178-9



Formula of each statistics

- Effect size = D/SD
- Paired t-test = D/SE
- Standardized response mean = D/SD^*
- Responsiveness statistic = $D/SD\#$

D = raw score change on measure

SE = standard error of the difference

SD = standard deviation at time 1

SD* = standard deviation of D

SD# = standard deviation of D among stable subject



ขอบคุณค่ะ