



# Research-mid

**Doctor: Munir Abu-Helalah**

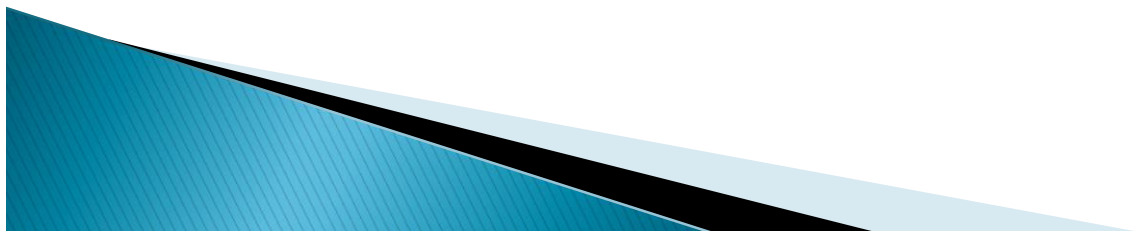
# Pretesting the Questionnaire

- ▶ Discover poor question wording or ordering
- ▶ Identify errors in questionnaire layout or instructions
- ▶ Determine if respondents were unable or unwilling to answer questions
- ▶ Suggest additional response categories
- ▶ Determine if questionnaire is appropriate length



# A pilot test

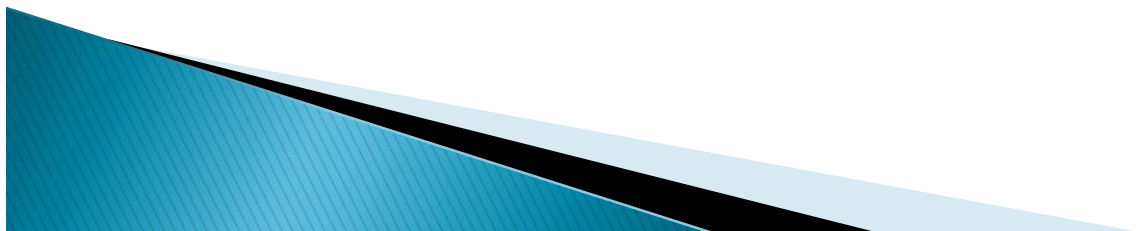
- ▶ Is an evaluation of the specific questions, format, question
- ▶ sequence and instructions prior to use in the main survey.
- ▶ Pilot testing is a crucial step in conducting a survey. Even modest pretesting can avoid costly errors.



# A pilot test

**Questions answered by the pilot test include:**

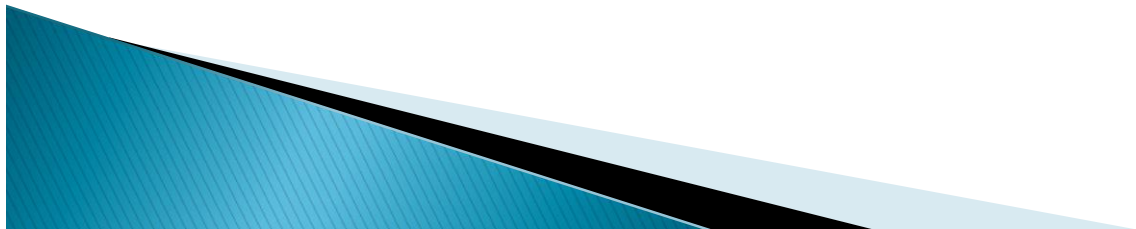
1. Is each of the questions measuring what it is intended to measure?
2. Are questions interpreted in a similar way by all respondents?
3. Do close-ended questions have a response which applies to all respondents?
4. Are the questions clear and understandable?



# A pilot test

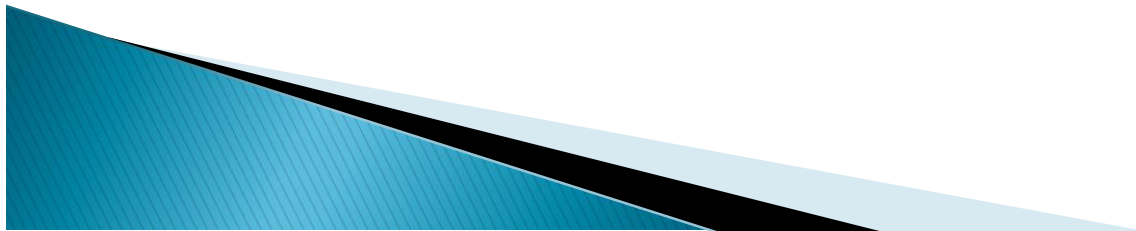
Questions answered by the pilot test include:

5. Is the questionnaire too long?
6. How long does the questionnaire take to complete?
7. Are the questions obtaining responses for all the different response categories or does everyone respond the same?



# A pilot test

**Always remember:  
Even modest edits can avoid costly errors!**



# Pilot your questionnaire!

- ▶ Pilot with a group of people
  - > similar to your target subjects

Usually 30 subjects

If the whole population is 100 or smaller

You can do it on 10 potential participants

For multi sites study: One site is enough

- Highlight problems before starting
- Misunderstandings
- Look for alternative wording
- Evaluate for missing data, consistency, reasonableness of answers
- Ask pre-test participants for direct, feedback
- Use duplicate administrations to assess reproducibility



Final polishing

# Bias

جدا مهم

Systematic difference in the response measurement

▶ Recall bias

- Cases more likely to remember than controls

▶ Observer bias

- Different interviewers – different interpretations
- Different interpretation of similar questions

▶ Non-response bias

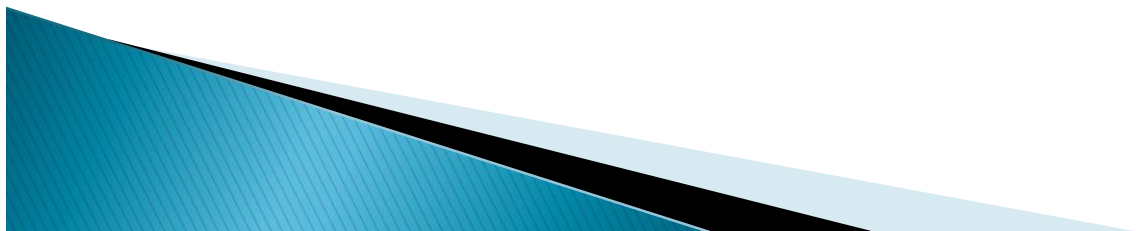
- telephone interviews





# How to reduce bias

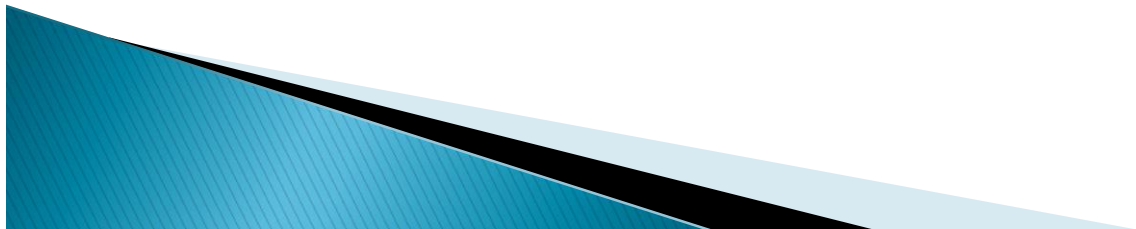
- ▶ Structured questionnaire
- ▶ Ensure high response rate
- ▶ Pretesting and piloting
- ▶ Training of interviewers



# Maximising the response rate

If you were sending out a questionnaire, what would you do to maximise the response rate?

In groups, 5 minutes

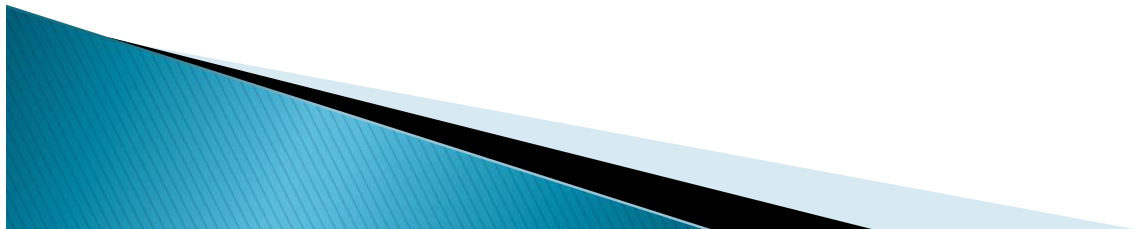


# Non-responders

- ▶ Understanding the characteristics of those who did not respond to the survey is important to quantify what, if any, bias exists in the results.

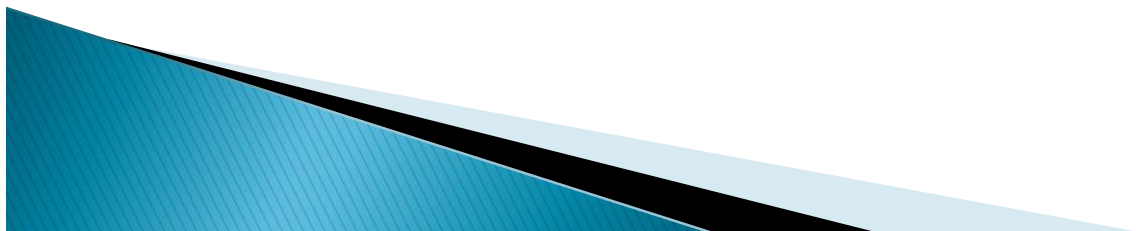
# Ways to Improve Response Rates

- ▶ Keep questionnaire short
- ▶ Ensure confidentiality
- ▶ Target the appropriate population
- ▶ Make it convenient for respondents
- ▶ Clearly communicate research purpose
- ▶ Give advance notice (advertising)
- ▶ Reward for completing the questionnaire



# Techniques for minimising non-response

- ▶ Good design
  - Thoughtful layout, easy to follow, simple questions, appearance, length, degree of interest and importance, thank people for taking part
- ▶ Pre-notification
- ▶ Explanation of selection
- ▶ Sponsorship, e.g. letter of introduction / recommendation
- ▶ Cover letter



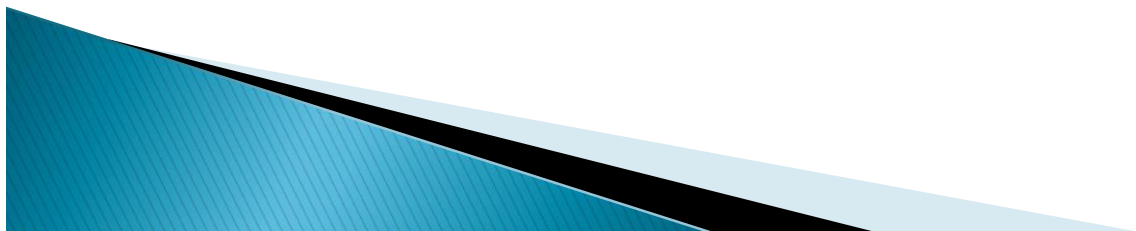
# Techniques for minimising non-response

- ▶ Incentives
  - Small future incentives, e.g. prize draw
  - Understanding why their input is important
- ▶ Reminders
- ▶ Confidentiality
- ▶ Anonymity
- ▶ Pre-paid return envelopes




# Ready to use questionnaires

- ▶ In general, for a tool to be validated for use in assessment, it should be:
- ▶ Valid
- ✓ Assess clinical important difference: smallest improvement considered worthwhile by a patient
- ✓ Tool sensitive for changes
- Reliable
- Precise
- Easy to administer
- Acceptable by the study population.



## **Definitions of terminology used in ready to use questionnaires selection**

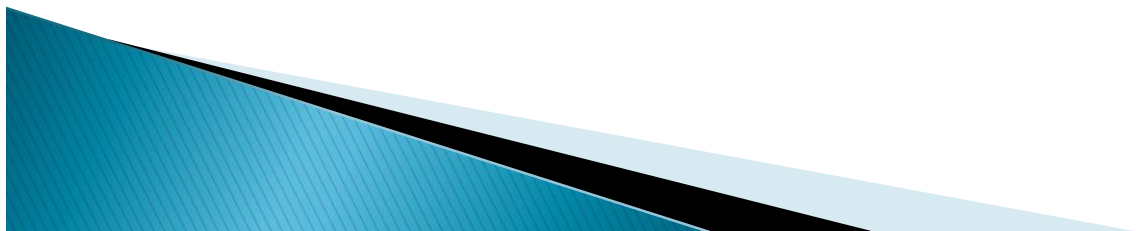
<b>Instrument</b>	<b>A questionnaire or interview or simple test (or some combination of these), used to measure and quantify health or disease status</b>
<b>Domain</b>	<b>An area or realm, one particular aspect within a broad assessment</b>
<b>Measure</b>	<b>A score, generally from a series of items designed to quantify some particular domain</b>
<b>Item or indicator</b>	<b>A single item, eg one question in a questionnaire</b>
<b>Scale</b>	<b>A simple test to quantify broad or single aspect of health using a numerical estimate from visual or numerical range</b>





# Questionnaire needs to be adapted to study population

- ▶ Know the respondents
  - Language
  - Education
  - occupation
  - ethnic group
  - sensitive issues



# Classification of ready to use questionnaires / scales

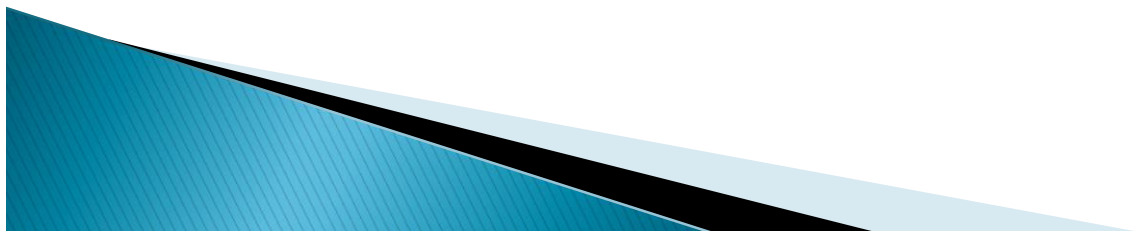
## A. Generic questionnaires:

- Developed to be applied for a large range of populations and health care problems
- They permit comparisons between populations or other groups of people, and also in the same group before and after an intervention.



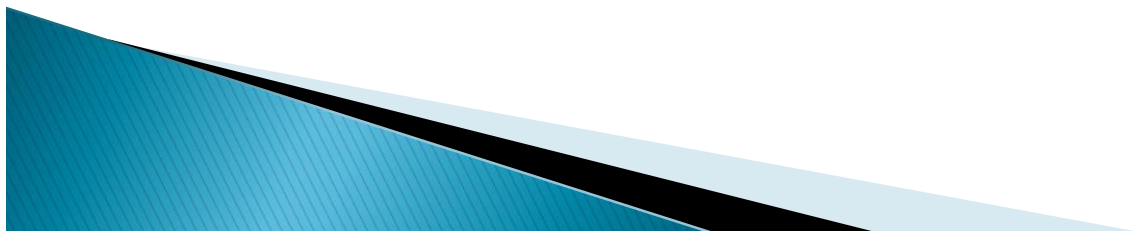
## A. Generic questionnaires:

- ▶ In order to apply any instrument for generic use, it should be validated across different groups and should be acceptable by these groups.



# A. Generic questionnaires:

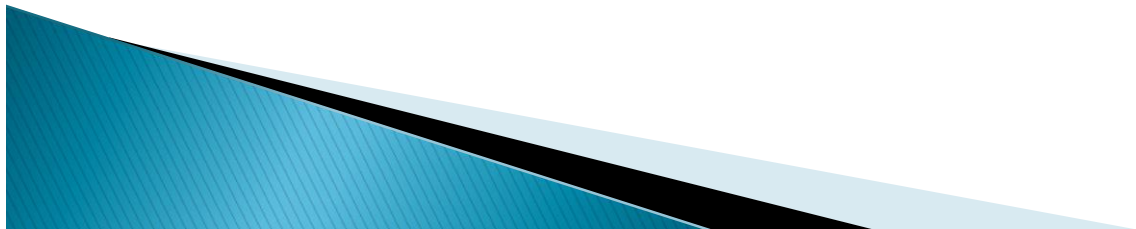
- ▶ It should be always considered that these measures are less responsive changes in health when compared with disease specific questionnaires.
- ▶ Therefore, if these are not used along with a disease specific questionnaire, it is advisable to choose a clinical outcome of direct relevance to the disease/health care problem under investigations.



# A. Generic questionnaires:

## Limitations:

- They may be insensitive to subtle but important changes in status with respect to a specific disease.
- They should be validated across a spectrum of different groups of people.



## B. Disease or population specific questionnaires:

- ▶ They are designed to target particular population or patients group.
- ▶ Examples of disease specific questionnaires.  
Asthma quality of life questionnaire  
Arthritis Impact Measurement scales  
Rand Diabetes Mellitus Battery



# الجداول بشكل عام كثير مهمة

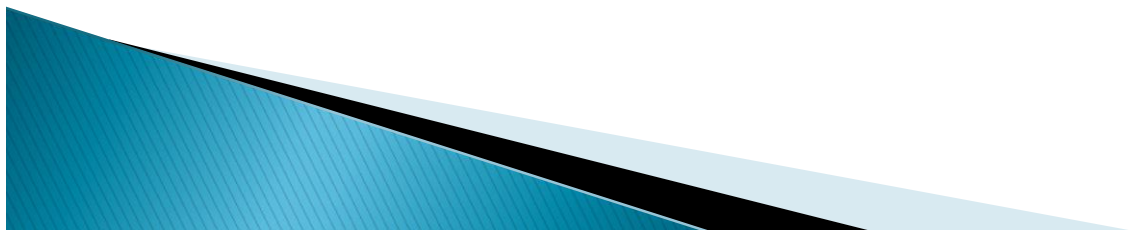
Concept	Comment
<b>1. Validity</b>	<b>Ability to measure what it supposed to measure.</b>
<b>a.Face validity</b>	<b>Refers to the investigators' subjective assessment of the questionnaire: a reasonable measure and items appears to be measuring what they intend to measure</b>
<b>b. Content validity</b>	<b>More systematic and comprehensive assessment than the face validity . It examines that extent to which items on a questionnaire covers all aspects that they intend to measure.</b>
<b>C.Construct validity</b>	<b>Construct: hypotheses are generated, then the questionnaire is tested to determine if it reflect these hypothesis. There two types of construct validity:</b> <b>1. Criterion validity: the extent that the results match with the pre-existing tools..<sup>3</sup></b> <b>2. Concurrent: when the new measure is administered at the same time with the pre-existing one</b>
<b>D. Convergent validity</b>	<b>The measure is correlated positively with other methods that measure the same concept.</b>
<b>E. Sensitivity (detection rate)</b>	<b>Proportion of actual cases. For example patients with clinical depression who score positive on measurement tool for depression</b>
<b>F. Specificity</b>	<b>It is the discriminative ability of a measure. Ie the proportion of people who are not cases and test negative on the measure</b>

<b>Concept</b>	<b>Comment</b>
<b>2.Responsiveness</b>	<b>Ability of an instrument to be responsive to actual changes that occurs over period of time.</b>
<b>3. Administration</b>	<b>Easy</b>
<b>4. Length</b>	<b>Not too long or too short.</b>
<b>5. Cost</b>	<b>Not expensive to obtain or to administer</b>
<b>6. Precision:</b>	<b>Ability to detect small changes</b>
<b>7. Reliability:</b>	<b>The extent to which a measure yields the same number or score each time it is administered.</b>
<b>a.Internal consistency</b>	<b>A test for the homogeneity and extent to which items are correlated within the same scale or domains in the scale. Cronbach's alpha gives an estimate of reliability based on all possible correlations between all items in the scale. Researchers have regarded that 0.7 is the minimum acceptable level for internal consistency. <sup>1,2</sup></b>
<b>Test-retest reliability</b>	<b>Relationship between scores obtained by the same person on two or more separate occasions. Kappa coefficient is used to test nominal data (ranging from -1 to 1,(0) if the agreement is not better than chance, negative if worse than chance and (1)if there is perfect agreement.</b>

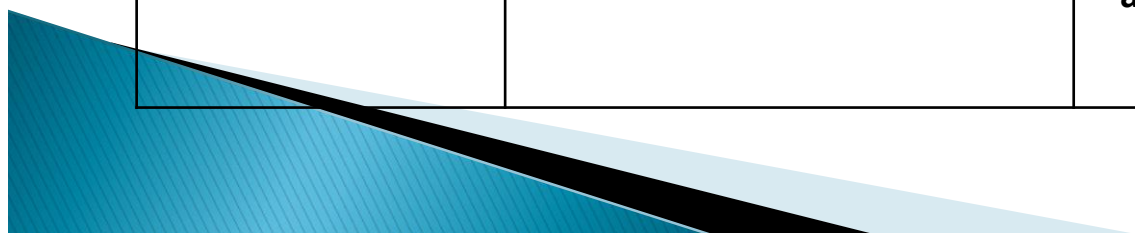


## Measures of *validity* of a new instrument

Measure	Concept measured	How measured
Face validity	The investigators' subjective assessment of the instrument; whether it appears to be measuring what it is intended to measure and whether each indicator is a reasonable one	Judgement (superficial)
Content validity	The extent to which the items in an instrument covers all aspects of the attribute to be measured. More systematic and comprehensive assessment than face validity	Judgement
Criterion validity	Validating an instrument by comparing it with a currently accepted reference measure <sup>6</sup>	Correlation coefficient, correlating the measure with some other accepted "criterion", ideally a gold standard <sup>6</sup>
Concurrent validity	Term for criterion validity when the two scales are administered at the same time; used when attempting to replace an existing scale with a new one that has some advantage (eg simplicity)	



Measures of <i>reliability</i> of a new instrument		
Measure	Concept measured	How measured
Internal consistency	A test for the homogeneity, the extent to which the items within a domain (which broadly should measure the same thing) are correlated.	<p><u>Cronbach's alpha</u>, an average of the correlation coefficients between all items. Takes values between 0 and 1. A low value (&lt;0.50) indicates that an item does not come from the same conceptual domain<sup>5</sup>, a value of 0.7 has been judged the minimum acceptable level for internal consistency<sup>6</sup>.</p> <p><u>Split half reliability</u>: correlation of two summary scores (for example from odd- and even-numbered questions in a questionnaire)</p>
Test-retest reliability	Relationship between scores obtained by the same person on two or more separate occasions.	<p><u>Kappa correlation coefficient</u>: Takes values between -1 and 1. A score of 1 indicates perfect agreement, 0 is the extent of agreement expected from chance, a negative score indicates worse agreement than would occur by chance</p>



## Measures of *validity* of a new instrument

فهم

<b>Construct validity</b>	Validating a new instrument by developing a hypothetical prediction of its performance, relevant where the variable of interest is abstract and cannot be directly observed <sup>1</sup>	For example a questionnaire for use in jaundice, measuring the extent of itching and excoriation, should show improvement when serum bilirubin decreases <sup>1</sup>
<b>Two subtypes:</b>		
<b>Convergent validity</b>	The measure is correlated positively with other methods accepted as measuring the same concept	Correlation coefficient
<b>Divergent or discriminant validity</b>	Lack of correlation with variables that measure a different unrelated topic	Correlation coefficient

