Clinician-Based Neuropathy Outcome Measures

Peripheral Neuropathy: Public Workshop
Feb 11-12, 2013

Gil I. Wolfe, M.D.

Dept. of Neurology

Univ. at Buffalo/SUNY SMBS

Buffalo, NY

University at Buffalo

The State University of New York

Neuropathy outcome measures

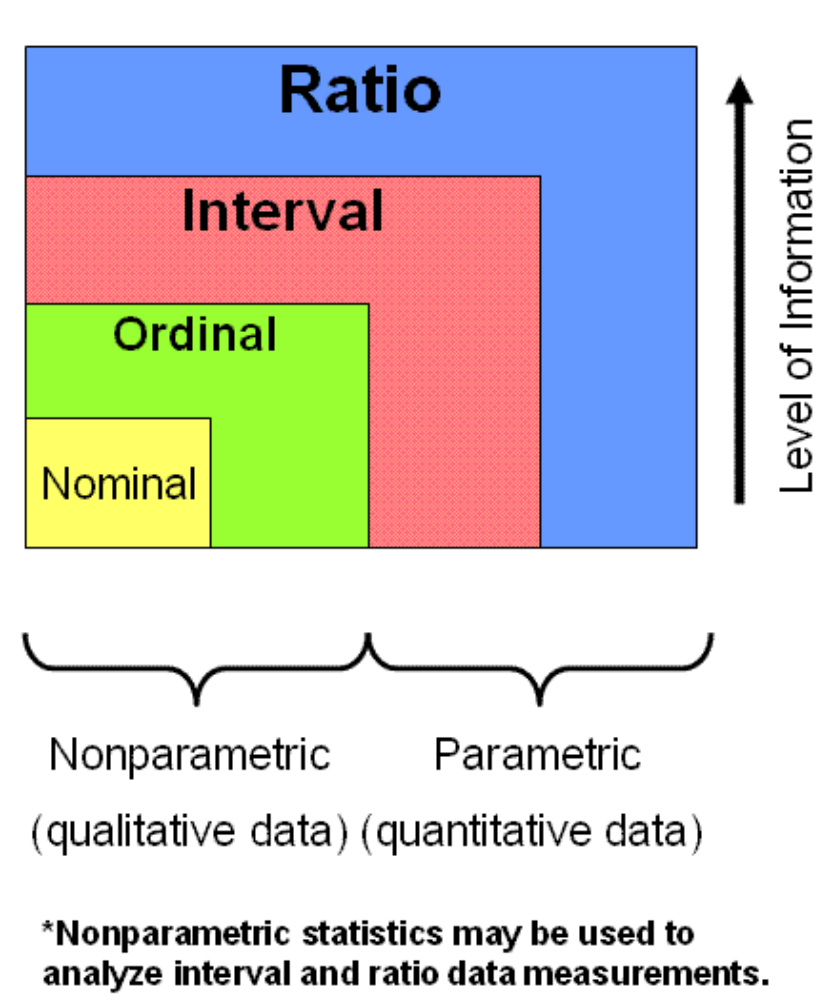
- Conventional clinimetrics
  1. Simple
  2. Communicable
  3. Construct validity
     - Are we measuring what we intend to measure
  4. Reliability
     - Consistency of measure
     - Local dependency of items will inflate this
  5. Responsiveness
     - How accurately outcome measure responds to clinical change

- Wide collection of measures used in "inflammatory" neuropathies
  - van Nes et al. *JPNS* 2008;13:136-147
- Efforts by INC, others to standardize measures
- Method hindrance in translating preclinical success
  - State-of-art trials without measures to match

Merkies et al. *Current Opin Neurol* 2012;25:556-563
Hobart et al. *Lancet Neurol* 2007;6:1094-1105

Types of data

- Nominal
  - Religion, ethnicity
  - Not amenable to numerical values outside of sorting
- Ordinal
  - 0,1,2,3,4,5 (like MRC scale)
  - Do not assume linearity!!
- Interval
  - Continuous value (temp, ht)
  - Attempt to move in this direction with item response theory

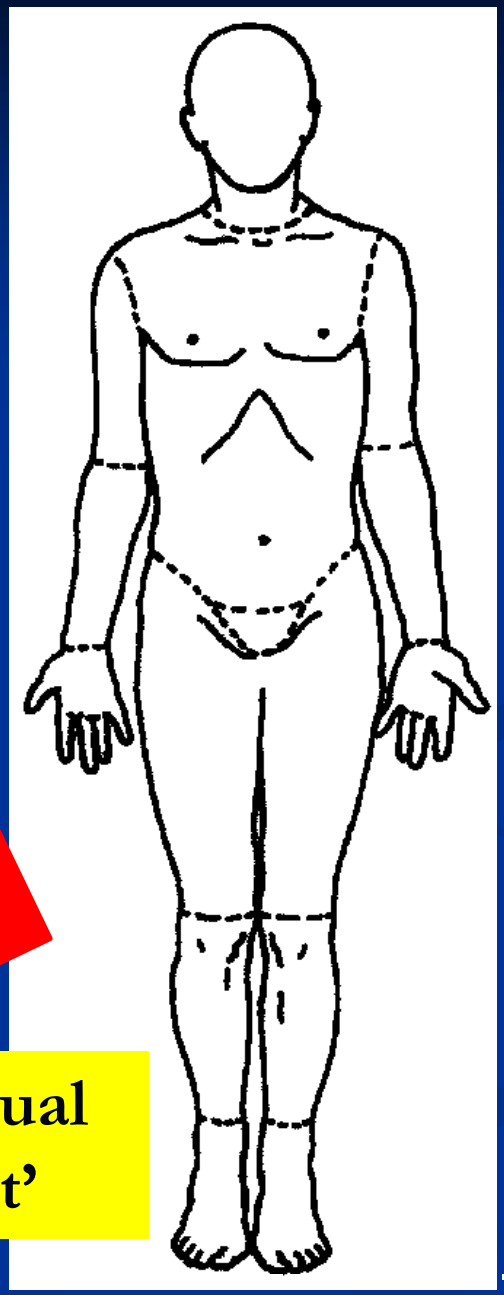


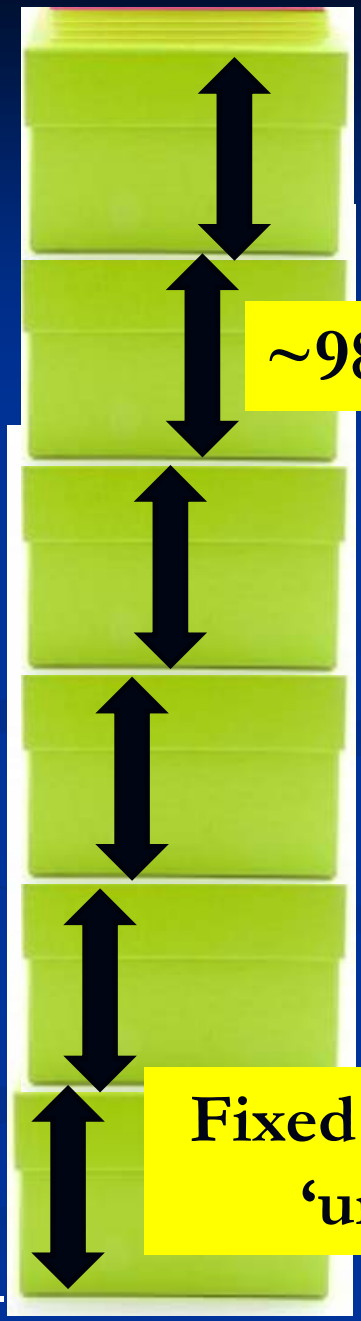
Measuring height (direct measures "easy")

I

REFLECTS REALITY

Unequal 'unit'





~98%

II

Fixed height 'unit'

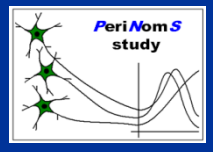
Traditional Scale

Measuring daily activities

Single item scales have poor reliability, validity, responsiveness

	Impossible 0	Very difficult 1	Difficult 2	Easy 3
Washing one's face			X	
Standing up from a chair		X		
Walking on level ground		X		
Walking upstairs	X			

Patient's traditional total score: 4 / 12



Disadvantages of traditional 'classical test theory' based scales

- Ordinal multi-item composite measures – construct items arbitrarily / each item has Likert-type response options
- All patients must complete all items, even though some may be inappropriate for their level of ability (weight or significance of each item not determined)
- Sum scores used, assuming all items have equal relevance and weight; also assume linearity
- Assume that obtained differences in sum score have equal meaning (e.g., scale range: 0 – 28; 2 point change from 0 to 2 is equivalent to 2 point change at the other end, i.e. 26 to 28)
  - Meaning of 1 pt change can vary up to 15-fold across scale

Davies. *Med Care* 2006
Hobart. *Lancet Neurol* 2007

Traditional versus Linear scales

Ordinal scales: lack linearity

Ordinal scales

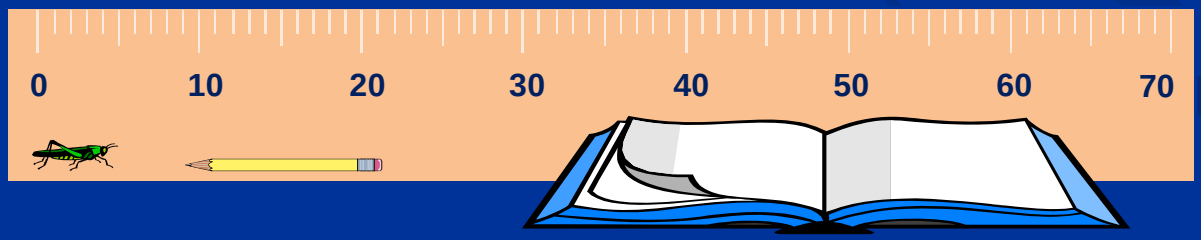
Impossible 0	Very difficult 1	Difficult 2	Easy 3

? ? ?

Points of transition

Linearity

5 cm 10 cm



Problems with traditional scales

an example

Original Fatigue Severity Scale

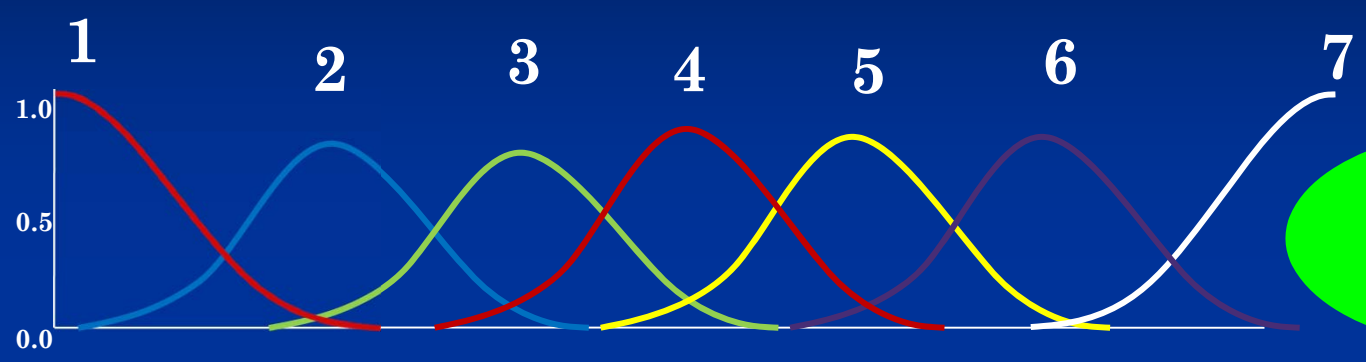
1=strongly disagree; 2=mainly disagree; 3=partially disagree; 4=do not agree/disagree; 5=partially agree; 6=mainly agree 7=strongly agree (circle one answer per question)

1	My motivation is lower when I am fatigued	1	2	3	4	5	6	7
2	Exercise brings on my fatigue	1	2	3	4	5	6	7
3	I am easily fatigued	1	2	3	4	5	6	7
4	Fatigue interferes with my physical functioning	1	2	3	4	5	6	7
5	Fatigue causes frequent problems for me	1	2	3	4	5	6	7
6	My fatigue prevents sustained physical functioning	1	2	3	4	5	6	7
7	Fatigue interferes with carrying out certain duties/responsib.	1	2	3	4	5	6	7
8	Fatigue is among my three most disabling symptoms	1	2	3	4	5	6	7
9	Fatigue interferes with my work, family, or social life	1	2	3	4	5	6	7

Are patients able to differentiate between these response options?

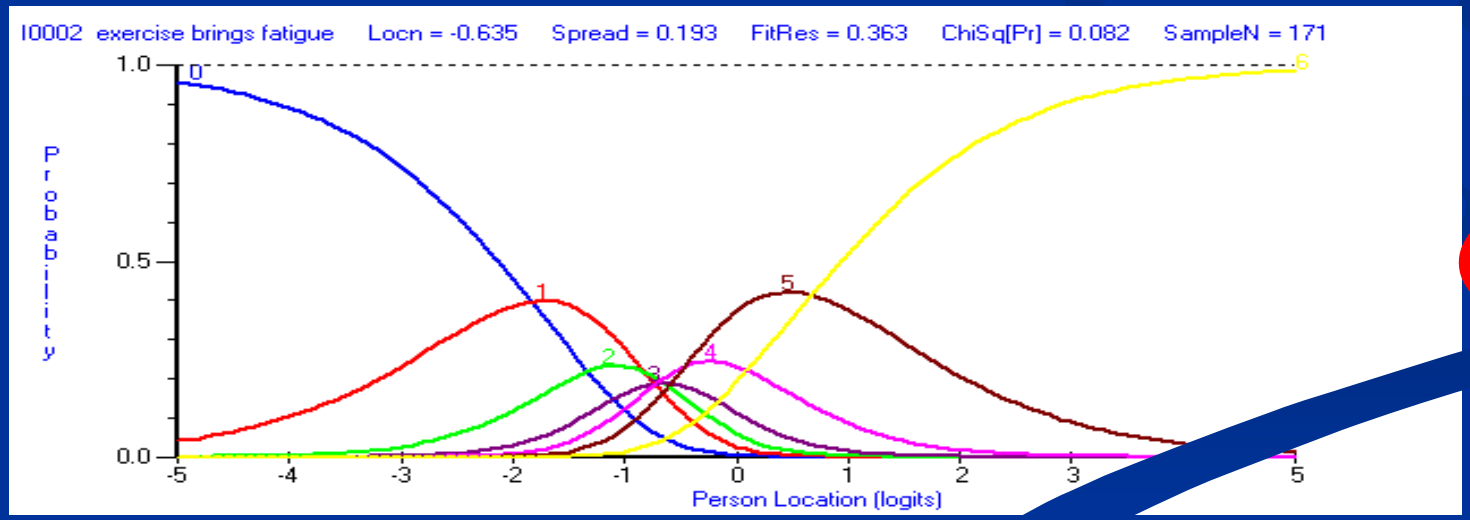
Discriminatory ability

Examined through Rasch analyses



Ordered thresholds

Item: Exercise brings on my fatigue



Disordered thresholds

May reflect pt or MD (MRC) difficulty with discrimination

Neuropathy outcome measures

Georg Rasch (1901-1980)

Latent Trait Theory

- Item response theory
  - Relate person's measurement to probability of response to item
  - Compare behavior between items
  - Estimate items' contribution to final score
  - Assess true change between ordinal values
- Person measurement

IRT vs. Rasch

For misfitting data:

- IRT seeks alternative mathematical model to best fit observed data
  - Prioritizes observed data at the cost of accepting poor “in”
- Rasch analysis explores why data did not fit the Rasch model
  - Prioritizes the math; models data to meet core requirements of an accurate outcome measure

Neuropathy outcome measures

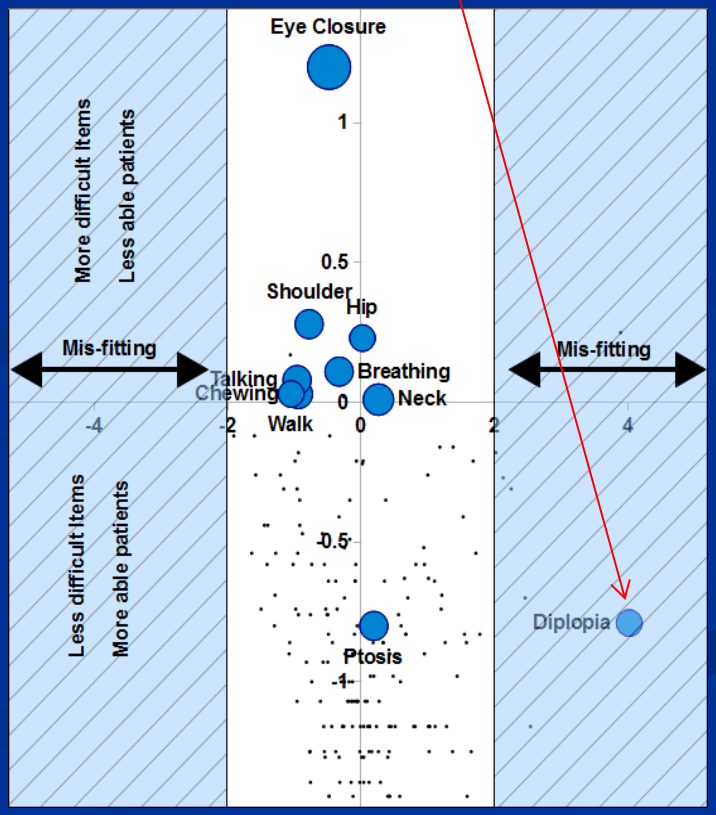
Rasch analysis:

Pt probability of completing a certain task/item dependent on

1. Person/patient ability
2. Item difficulty
   - Item addressed based on location on metric
   - Item's ability to differentiate performance levels

- Advantages
  - Item sensitivity in relation to ability levels
  - Scale's accuracy in estimating differences
    - Is a score of 6 twice as bad as a 3?

Impact on unidimensionality of MG Composite



Sadjadi et al. *Muscle Nerve* 2012;45:820-825

RASCH - based on Guttman scaling

0 = Unable
1 = Able

	Item 1	Item 2	Item 3	Item 4	*sumscore*
Person A	0	0	0	0	0
Person B	1	0	0	0	1
Person C	1	1	0	0	2
Person D	1	1	1	0	3
Person E	1	1	1	1	4
sumscore	4	3	2	1	

Easiest item

Most difficult item

RASCH - based on Guttman scaling

0 = Unable
1 = Able

	Item 1	Item 2	Item 3	Item 4	*sumscore*
Person A	0	0	0	0	0
Person B	1	0	0	0	1
Person C	1	1	0	0	2
Person D	1	1	1	0	3
Person E	1	1	1	1	4
sumscore	4	3	2	1	

Least able
Lowest ability
Most ill pt

Most able
Highest ability
Least ill pt

Neuropathy outcome measures

- Symptom measures
  - Sensory, motor, autonomic
  - Ordinal value assignments (NSS)
- Quantitative exams/functional scales
  - Motor (MRC)
  - Sensory (ISS)
  - Autonomic
- NCS
- Computerized sensory exams
- Nerve biopsy

NIS; ordinal

TNS; ordinal

Neuropathy symptom score (NSS)

1 point for symptom presence (DMPN studies)

- Muscle weakness
  - Bulbar
    - Extraocular
    - Facial
    - Tongue
    - Throat
  - Limbs
    - Shoulder girdle/upper arms
    - Hand
    - Glutei and thigh
    - Legs
- Sensory
  - Negative symptoms
    - Difficulty ID objects in mouth
    - Difficulty ID objects in hands
    - Unsteadiness walking
  - Positive symptoms
    - Numbness, falling asleep, like Novocain, prickling, any site
    - Pain, any location
- Autonomic
    - Postural fainting
    - Impotence (male)
    - Loss of urinary control
    - Night diarrhea

Dyck et al. *Ann Neurol* 1980;8:590-596

Neuropathy Impairment Score (NDS)

Motor	Motor	Reflexes & Sensory
(Papilledema)	Wrist extension	Biceps
CN III	Wrist flexion	Triceps
CN VI	Finger extension	Brachioradialis
Face	Finger flexion	Quadriceps
Palate	Hand intrinsics	Triceps surae
Tongue	Iliopsoas	Touch pressure (index)
Respiratory	Glutei	Pricking [illegible] (index)
Shoulder abd	Quadriceps	Vibration (index)
Biceps brachii	Hamstrings	Joint position (index)
Brachioradialis	Ankle dorsiflexors	Touch pressure (great toe)
Elbow extension	Ankle plantar flexors	Pricking pain (great toe)
Elbow flexion		Vibration (great toe)
		Touch pressure (great toe)

Dyck et al. *Ann Neurol* 1980;8:590-596

Tafamidis for TTR-FAP

Coelho T et al. *Neurology* 2012;79:785-792

- Tafamidis 20 mg qd vs placebo x18 mos
  - Occupies thyroxine binding sites to stabilize tetramers
- Primary outcomes taken from DMPN
  - NIS-LL (single neurologist, mean of 2 assessments)
  - Norfolk QOL-DN
- Results
  - NIS-LL: trend (p=0.068) toward more responders (<2 pt worsening) in ITT (n=128) population
  - Norfolk QOL: p=0.116 in ITT population
  - In efficacy evaluable pts (n=102), p values just below 0.05
  - Some 2º outcomes favored tafamidis
- Conundrum: Type 2 error??

Muscle strength scales

Description	MRC Scale	MRC Modified Scale
Normal	**5**	10
Barely detectable weakness	4+	9
Holds test position; moderate weakness	**4**	
Transient resistance with abrupt collapse		6
Anti-gravity but no resistance		5
Anti-gravity but not full range	3-	4
Full range with gravity removed	2+	
Partial range with gravity removed	**2**	
Flicker of contraction, no joint mvmt	**1**	1
No contraction	**0**	0

MRC Rasch Analysis

- Based on 1065 patients with neuropathies, myopathies
- Disordered thresholds in ¾ of muscles examined
  - Too many options
  - Option labelling is confusing
- Physician experience or disease type had no influence
- Proposed MRC revision (ordered thresholds except for masseter)

3 Normal strength

2 Slight weakness (<50% loss)

1 Severe weakness (>50% loss)

0 Paralysis

Vanhoutte et al. *Brain* 2012;135:1639-1649

INCAT Sensory Sumscore (ISS)

Scored 0 to 20

Pinprick Arms	Pinprick Legs	Vibration Arms (Rydel-Seiffert)	Vibration Legs (Rydel-Seiffert)	Two-point discrimination at index finger
Normal 0 = index finger	Normal 0 = hallux	Normal 0 = index finger	Normal 0 = hallux	Normal
Abnormal	Abnormal	Abnormal	Abnormal	Abnormal
1 = index finger	1 = hallux	1 = index finger	1 = hallux	1 = 5-9 mm
2 = wrist	2 = ankle	2 = wrist	2 = ankle	2 = 10-14 mm
3 = elbow	3 = knee	3 = elbow	3 = knee	3 = 15-19 mm
4 = shoulder	4 = groin	4 = shoulder	4 = groin	4 = ≥ 20 mm

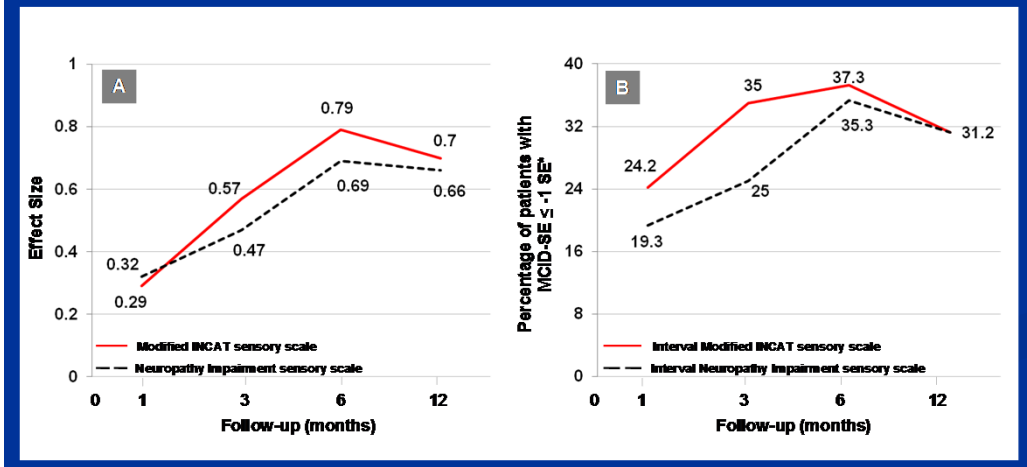
Validity, reliability, responsiveness assessed

Merkies et al. *Neurology* 2000;54:943-949

Modifed ISS (mISS)

Touch pressure and joint position incorporated

- Rasch analysis on 162 GBS/CIDP/MGUS PN pts
  - Validity and reliability in stable cohort
  - Responsiveness in newly diagnosed patients
  - Conclusions
    - NIS sensory subset and mISS could be converted from ordinal to interval scoring
    - mISS superior to NIS with greater responsiveness



van Nes et al. *PNS* 2011 poster

Conclusions

- Latent trait theory a powerful tool to transform ordinal data into interval/linear measurements
- If model's expectations are not met 'CLEANING UP' the scale may meet the requirements
  - Possible with Rasch; not a focus of IRT
- LTT generates outcome measures that satisfy scientific criteria and allow assessment of statistical vs. clinical significance

NINDS Common Data Elements

Harmonizing information. Streamlining research.

No widely used data standards in NINDS-funded clinical research

Researchers create data collection instruments for each new project

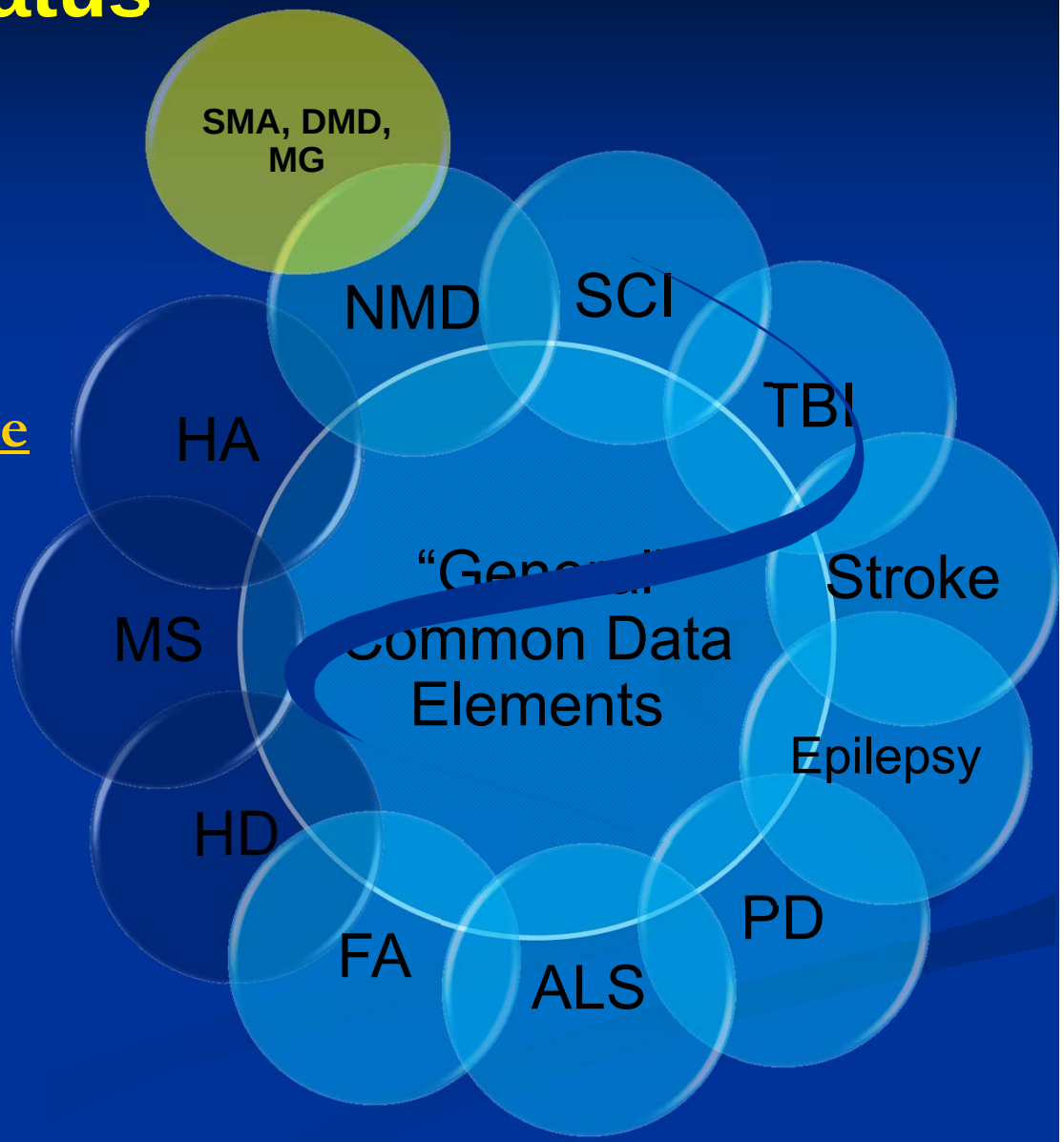
Meta-analyses across studies require extensive data re-formatting

Multitude of data formats creates barriers to data sharing

Current Project Status

CDEs available for use are found at:
http://www.commondataelements.ninds.nih.gov/

- Available for use
- In development



NINDS-CDE

DMPN small fiber assessments

Instrument	Classification [Core, Supplemental, Exploratory]	Description/ Scoring info	Time	Copyright	Reference
IENF density	Supplemental	Morphological Number/mm	Lab test	Public domain	Hermann et al. 1999; McArthur et al. 1998
QSART	Supplemental	Commercial equipment Latency (sec) Volume (µL)	45-60 minutes	WR Medi[illegible] Patent #6269265	Low et al. 2006
Utah Early Neuropathy Score	Supplemental	PE based for preclinical and early detection	15 minutes	Protected rob.singleton@hsc.utah.edu for permission	Singleton et al. 2008

NINDS-CDE

DMPN large fiber assessments

Instrument	Classification	Description/ Scoring info	Time	Copyright	Reference
Michigan Neuropathy Screening Instrument	Supplemental	History questionnaire plus PE assessment	30 minutes	Public domain	Feldman et al. 1994
QST	Supplemental	Commercial equipment Psychophysical test	45 minutes	Public domain	Gruener et al. 1994
NIS	Supplemental	Two versions; most commonly used assigns scores to motor, DTR, sensory function	5-10 minutes	Common version in public domain	Dyck et al. 1987

NINDS-CDE

Inflammatory neuropathy assessments

Instrument	Classification	Description/ Scoring info	Time	Copyright	Reference
Rasch-built ODS (RODS-CIDP, MMN, MGUS)	Supplemental	Linearly weighted scale; activity and social scores	5 minutes by patient	Public domain	van Nes et al. 2011
INCAT-ODSS	Supplemental	UE/LE assessment	15 minutes	Public domain	*JNNP* 2006 [illegible].970
Overall Neuropathy Limitations Scale (modification of ODSS to address ceiling effect)	Supplemental	UE/LE scales	<5 minutes	Public do[illegible]	Graham et al. 2006
Pinch grip/handgrip dynamometry	Supplemental	Kg force 3 tests to develop mean	5-30 minutes	Public domain	Goonetilleke et al. 1994; Maurissen et al. 2003; *Arch PMR* 1997;12:1364

NINDS-CDE

Inherited neuropathy assessments

Instrument	Classification	Description/ Scoring info	Time	Copyright	Reference
CMT Disease Neuropathy Score (Versions 1 & 2)	Supplemental	Composite of symptoms, PE, NCS SNAPs and CMAPs	< 30 minutes	None	Shy et al. 2005; Murphy et al. 2011
NIS	Supplemental	Two versions; most commonly used assigns scores to motor, DTR, sensory function	5-10 minutes	Common version in public domain	Dyck et al. 1987
SF-36	Supplemental	General health survey	5-20 minutes	Copyrighted http://www.qualitymetric.com/DefaultPermissions/RequestInformation/tabid/233/Default.aspx	Ware et al. 1992, 1994, 2001; McHorney et al. 1993
Pinch grip/handgrip dynamometry	Supplemental	Kg force 3 tests to develop mean	5-30 minutes	Public domain	Goonetilleke et al. 1994; Maurissen et al. 2003; *Arch PMR* 1997;12:1364