



QUINTILES®



OUTCOME®

Real-World and Late Phase Research

W33: The Correlation between Patient Reported Outcomes and Clinician Reported Outcomes



Roche

Eric Gemmen, Quintiles Outcome

Katie Zarzar, Roche Products Ltd in the UK

Shital Kamble, Quintiles Outcome

Rebecca Dawsey, TransPerfect



TRANSPERFECT

November 6, 2013

Purpose

Explore the evidence of the degree of correlation between patient reported outcomes (PROs) and clinician reported outcomes (ClinROs), and how this varies by:

- disease/therapeutic area and measure:
 - disease/symptom presence
 - symptom frequency
 - symptom severity

Methods

- A review of the literature and analysis of existing patient registry data was conducted to qualitatively assess degree of correlation between PROs and ClinROs.
- Statistical measures of correlation and concordance are expressed in terms of Spearman's rho, Pearson's rho, weighted Kappa, and Kendall's Tau.
- A review of translation and linguistic validation projects involving PRO and ClinROs was also conducted to examine language-related differences and correlations between the scales.

Methods

- The results are organized as follows:
 1. Direct Comparison of PRO and ClinRO Responses;
 2. Relative Impacts of Language to the Responses of PROs and ClinRO Measures
- Types of outcome measures are specified:
 - diagnosis, symptom presence (yes/no),
 - symptom frequency
 - symptom severity

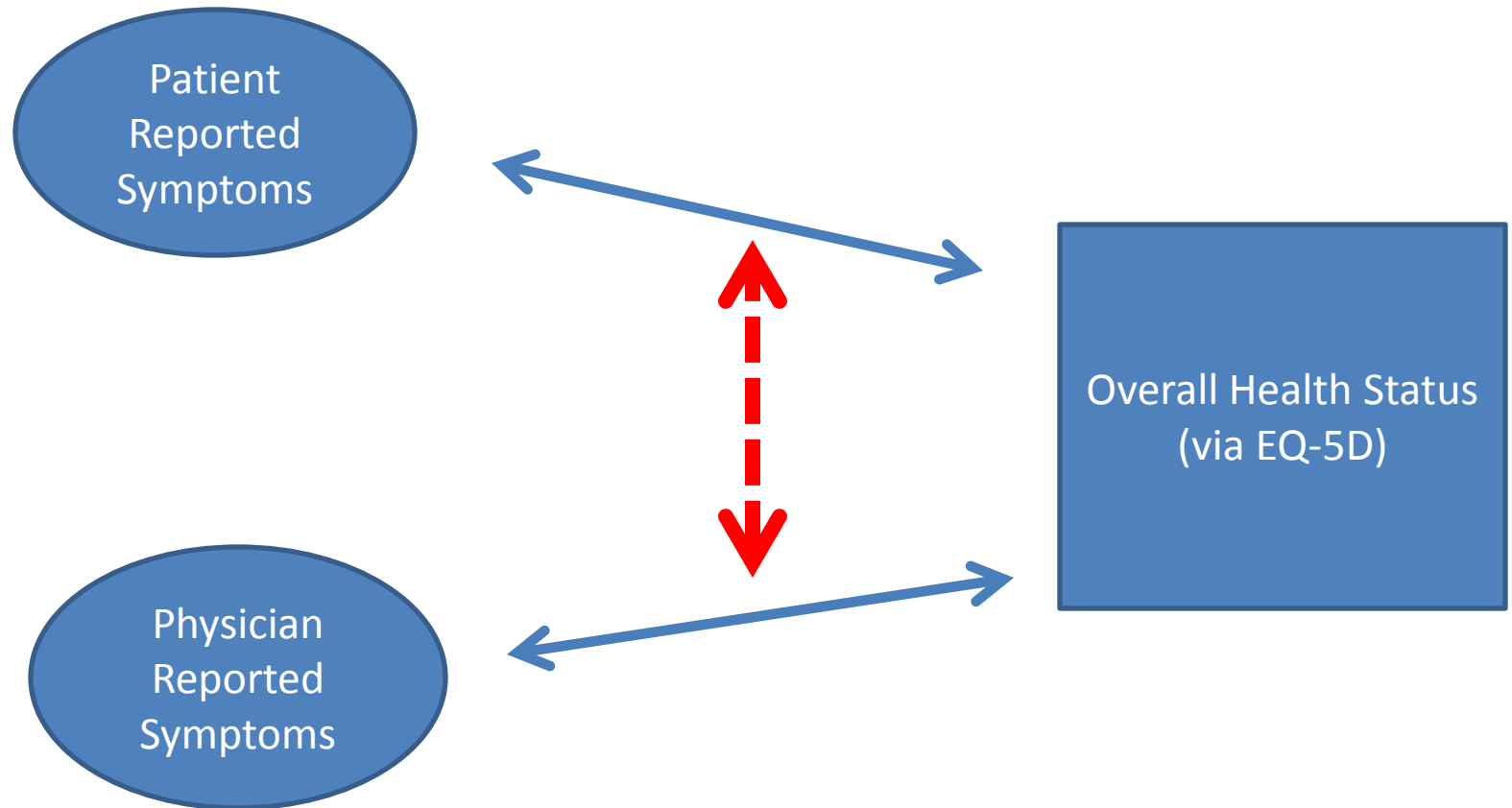
Methods (cont.)

- Specific examples of PRO-ClinRO pairs are provided for the following disease areas:
 - Oncology
 - Depression in Parkinson's Disease
 - Multiple sclerosis
 - Rheumatoid arthritis
 - Acne, Psoriasis, and Atopic Eczema
 - Dry eye
 - Crohn's Disease
 - Pediatrics

Limitations

- This workshop presents a survey of the evidence relating PROs to ClinROs primarily from the medical literature. As such, detail and results are limited to what is presented in the research article.

Concordance of Symptom Presence and Overall Health Status (cont.)



Concordance of Symptom Presence and Overall Health Status (cont.)

Strength of Concordance (Kendall's Tau) between patient/physician reported symptoms and overall health status as measured by EQ-5D

	Patient Reported		Physician Reported
Fatigue	0.36	>	0.24
Nausea	0.19	>	0.10
Vomiting	0.13	>	0.09
Diarrhea	0.14	>	0.05
Constipation	0.17	>	0.13
Dyspnea	0.27	>	0.15
Appetite Loss	0.28	>	0.22

Source: Basch (2010) from N=467 persons with breast, lung, genitourinary or gynecologic malignant conditions across a total of 4034 clinic visits at Memorial Sloan-Kettering Cancer Center, New York.

Diagnosis/Symptoms Agreement Example: Rheumatoid Arthritis

Agreement and Correlations between Rheumatoid Arthritis Values Findings by PROs and Physician

Patient-reported	Physician-reported					
	SJC	TJC	DAS28	MD-Global	CDAI	SDAI
SJC	0.772^b	0.499	0.525	0.531	0.563	0.541
TJC	0.429	0.75^b	0.552	0.493	0.611	0.598
RADAI	0.393	0.604	0.56	0.399 ^a	0.667	0.646
RAPID3	0.372	0.594	0.523	0.361 ^a	0.731	0.706
RAPID4	0.402	0.625	0.562	0.395 ^a	0.75	0.726
RAPID5	0.53	0.709	0.662	0.511 ^a	0.829	0.851
MDHAQ	0.246^d	0.491	0.442	0.304 ^a	0.531	0.531
VAS-Global	0.396	0.583	0.517	0.026^{c,e}	0.754	0.725
VAS-Pain	0.323	0.508	0.434	0.314 ^a	0.632	0.606

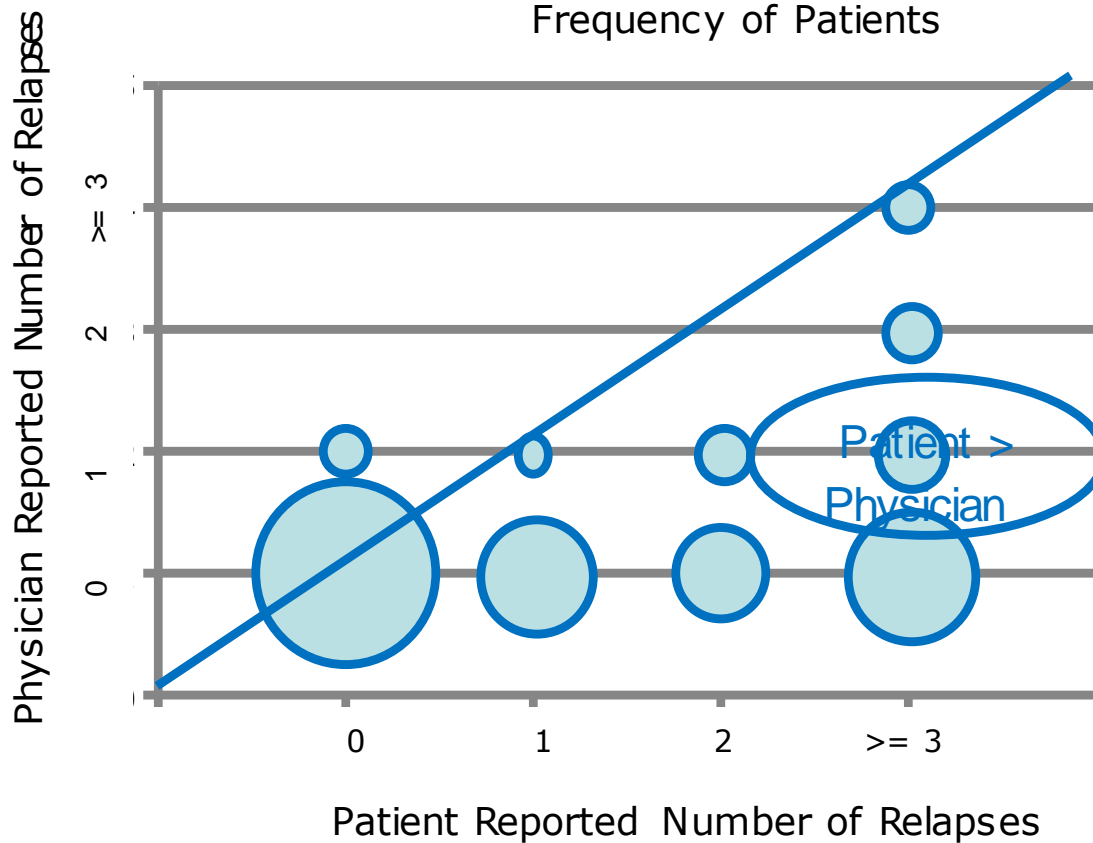
Source: Amaya-Amaya (2012). All correlations via Spearman's rho, except:

^a Correlation by Kendall's Tau; ^b Agreement by Kendall's W test; ^c Agreement by Weighted kappa

**All data P < 0.0001, except in ^dP = 0.004 and ^eP = 0.241.



Association of Patient and Physician Reported Number of Relapses During Months 1 to 6

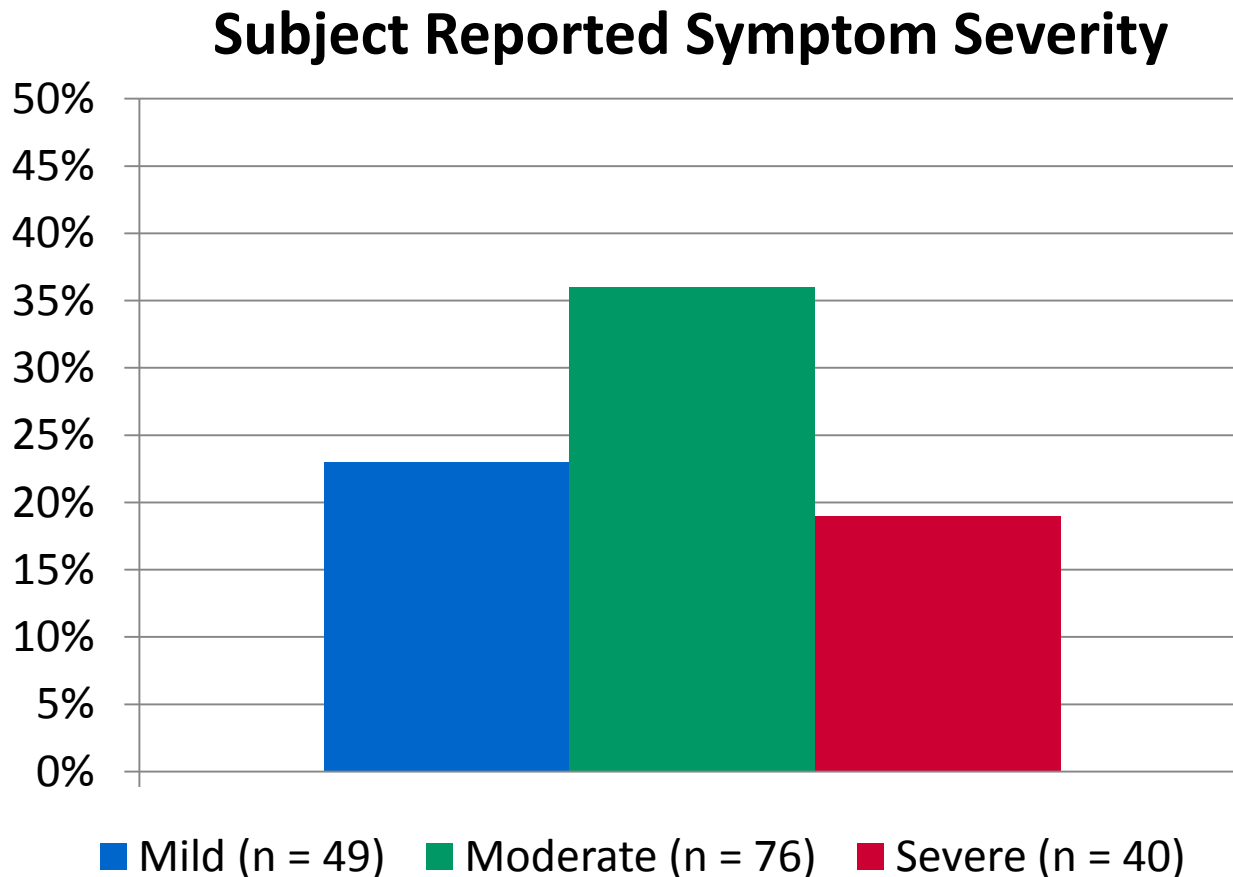


Symptom Severity Example: Dry Eyes

- 162 dry eye subjects, and 48 controls
- Self-assessment of severity of dry eye completed first
- Clinicians first completed a clinical assessment, and then a clinician assessment of severity of the subject's dry eye symptoms

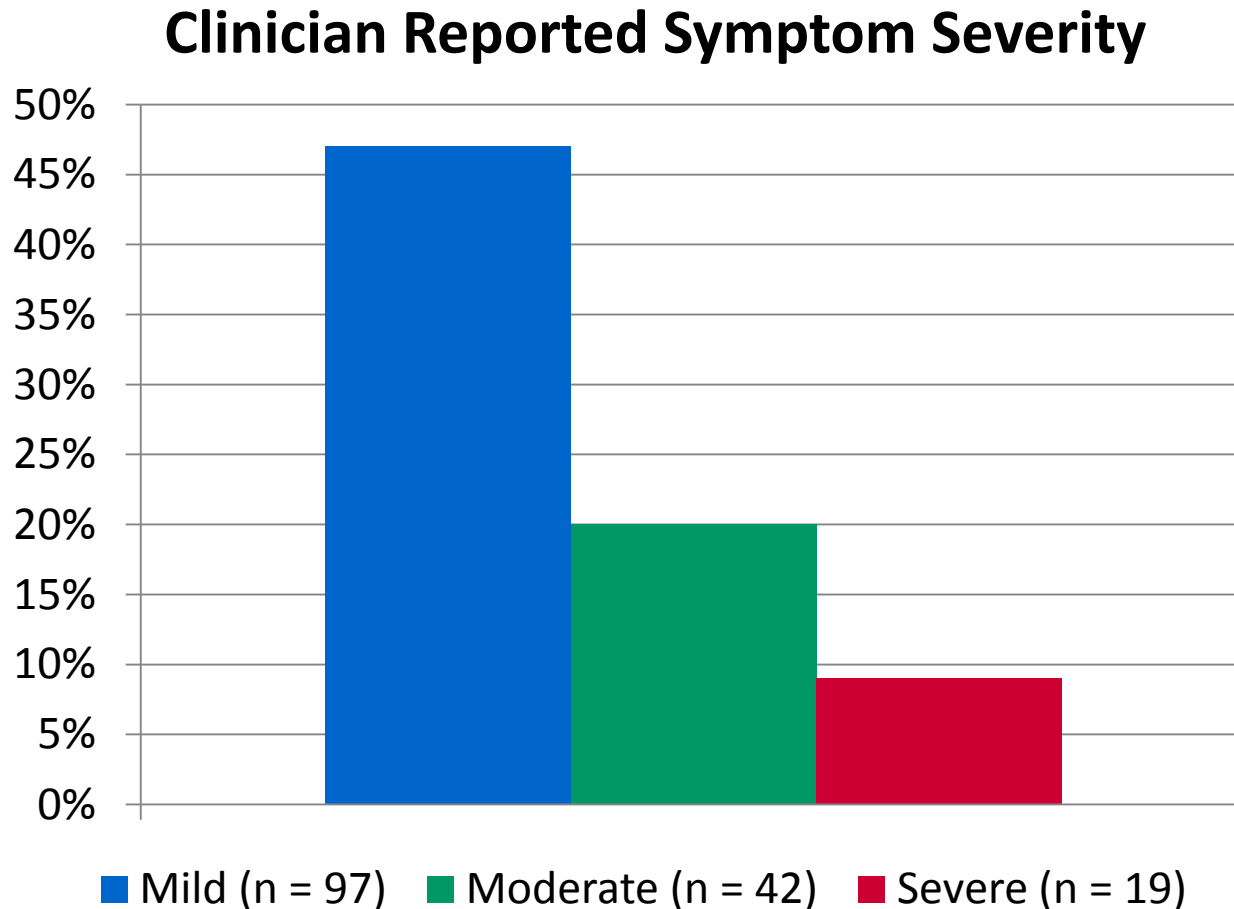
(Patients did not discuss their self-assessment with the clinician)

Patient Reported Symptom Severity



Source: Chalmers et al. (2005)

Clinician Reported Symptom Severity



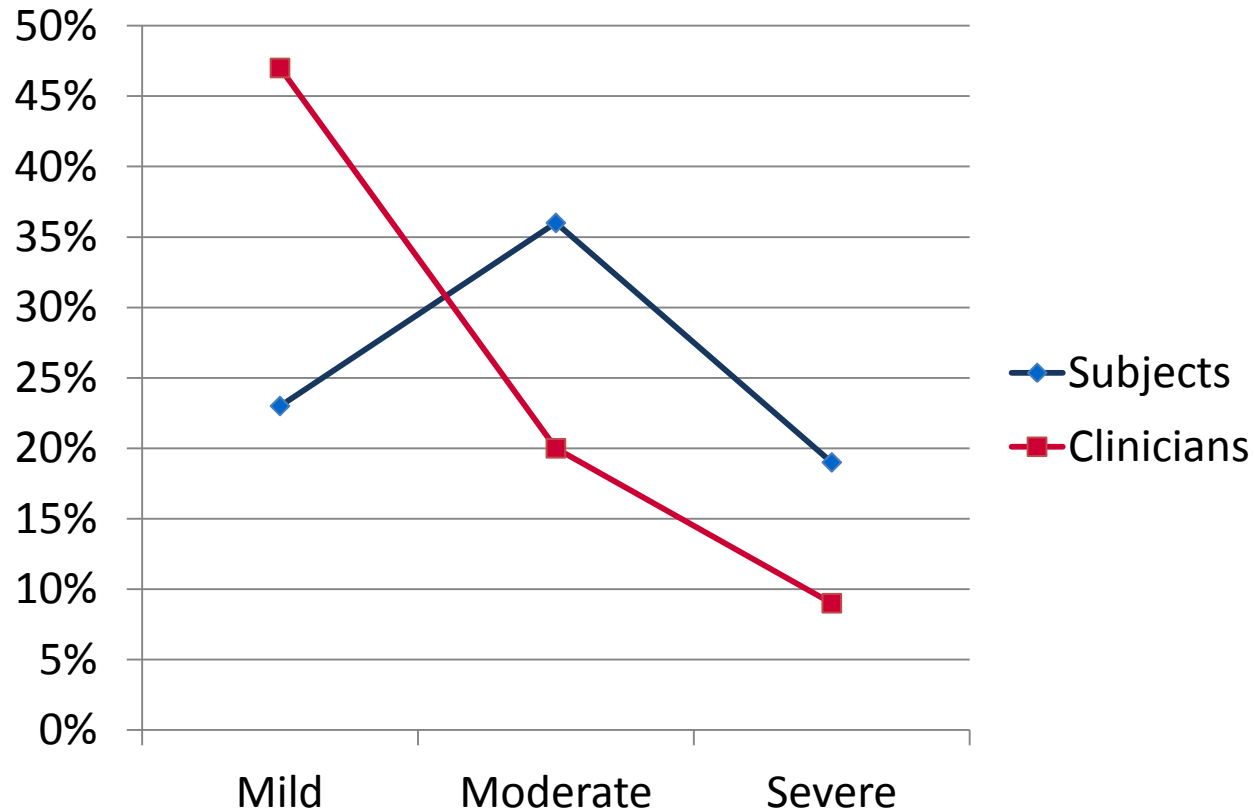
Source: Chalmers et al. (2005)

Summary of PRO Versus ClinRO Severity Assessments for All Subjects

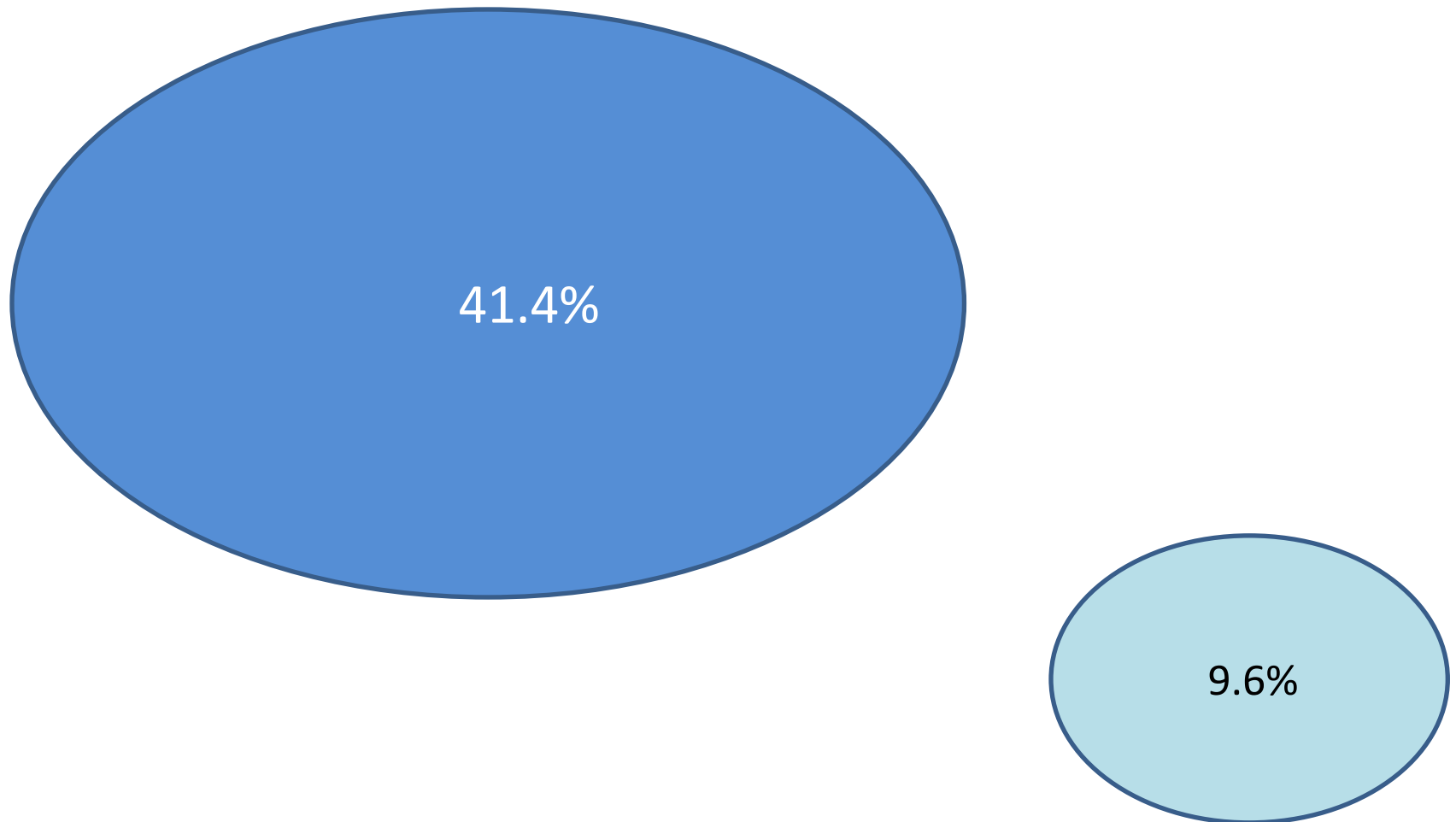
n = 209		Self-Assessment			
		None (n = 46)	Very Mild/Mild (n= 49)	Moderate (n= 74)	Severe/Extremely Severe (n = 40)
Clinician Assessment	None (n = 51)	39 (76.5%)	7 (13.7%)	5 (9.8%)	
	Mild (n= 97)	7 (7.2%)	37 (38.1%)	45 (46.4%)	8 (8.2%)
	Moderate (n =42)		5 (11.9%)	16 (38.1%)	21 (50%)
	Severe (n= 19)			8 (42.1%)	11 (57.9%)

Comparison of Subject and Clinician Responses

Comparison of Subject and Clinician Responses



Discordance of Symptom Severity Rating by Subjects and Clinicians



Symptom Severity Example: PD

- 50 patients diagnosed by a movement disorder specialist with idiopathic PD

Scales utilized:

- PRO
 - Beck Depression Inventory (BDI)
 - Geriatric Depression Scale (GDS)
- ClinRO
 - Hamilton Depression Rating scale (HAM-D)
 - Montgomery-Asberg Rating Scale (MADRS)
- 25 respondents received the PROs first, and 25 respondents received the ClinROs first

Correlation of PRO and ClinRO Measures

Correlation Matrix of Depression Measures – all Participants
(n = 50)

	BDI	GDS	HAM-D	MADRS
MADRS	0.61^a	0.5	0.76^{a,b}	1
HAM-D	0.74^a	0.57^a	1	
GDS	0.5^{a,b}	1		
BDI	1			

Abbreviations: BDI: Beck Depression Inventory, GDS: Geriatric Depression scale, HAM-D: Hamilton Depression Rating scale, MADRS: Montgomery-Asberg Rating Scale

^a Significant correlation $P < 0.01$

^b Significant difference between bolded correlations ($0.76 > 0.50$), Fisher $z P < 0.05$

Agreement of PROs and ClinROs

Rate of Agreement of Self-Report and Clinician-Based
Depression Measures
(n = 50)

		Self-Report BDI		
		Asymptomatic	Symptomatic	Fully Symptomatic
Clinician-based HAM-D	Asymptomatic	24	4	1
	Symptomatic	5	7	1
	Fully Symptomatic	2	1	5

Abbreviation: BDI, Beck Depression Inventory; HAM-D, Hamilton Depression Rating Scale

Overall Results

Dry Eyes:

- Under reporting of severity by clinicians

Parkinson's Disease:

- Strong correlation (72%) between patient and clinician ratings of depressive symptom

Agreement of Self-assessed and Clinician assessed Severity Example: Skin Disease- Acne, Psoriasis and Atopic Eczema

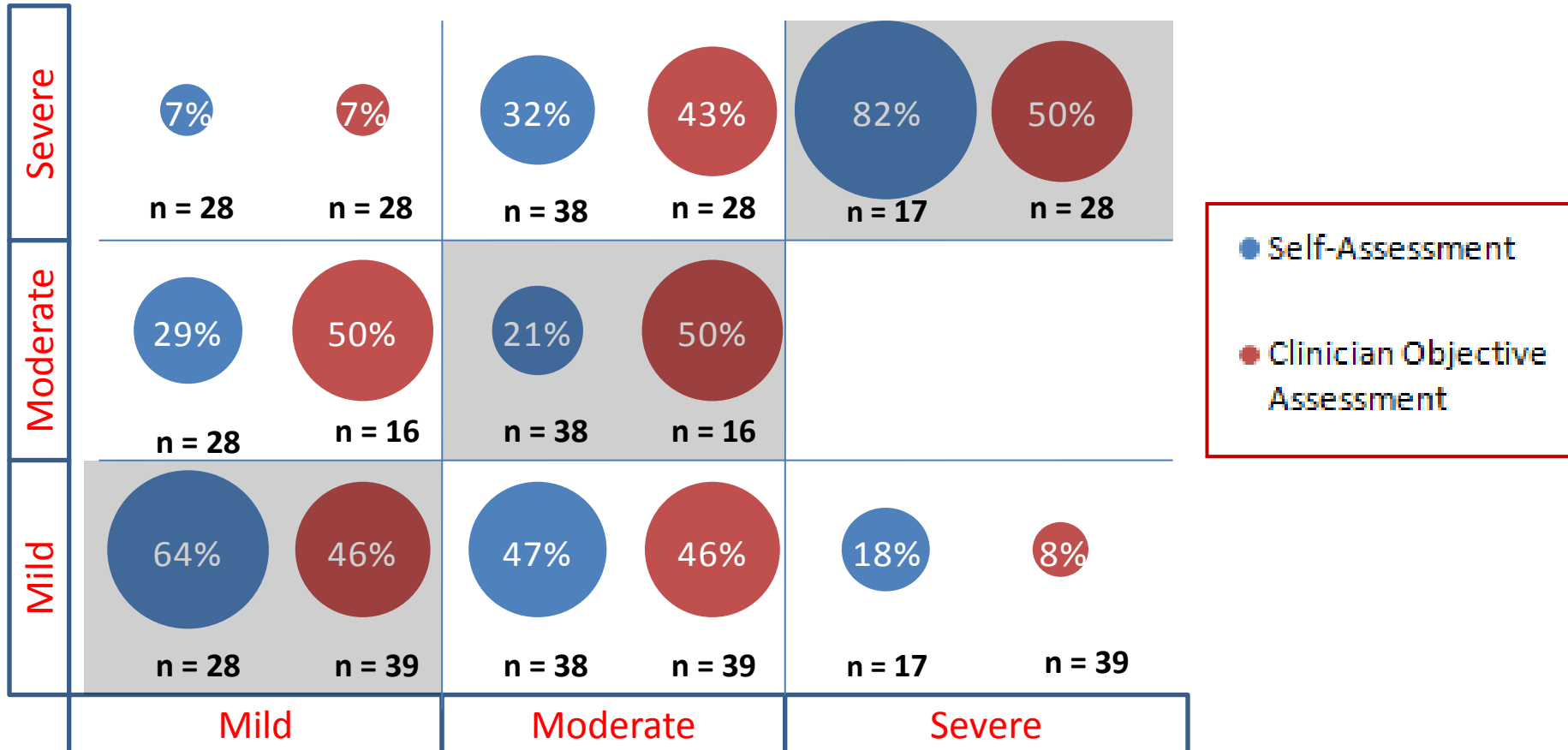
- **Objective and Study Population**
 - A cross-sectional study examined psychological associations of acne, psoriasis, or atopic eczema
 - Comparison- Self-assessed versus clinician objective responses regarding skin disease severity
 - 108 patients from general and specialist dermatology practices:
 - Acne (n=41),
 - psoriasis (n=47), and
 - Atopic eczema (n=20)

Skin Disease: Acne, Psoriasis and Atopic Eczema

- **Objective severity assessment:**
 - Leeds technique (Acne),
 - Psoriasis Area and Severity Index (PASI),
 - Six Area Six Sign Atopic Dermatitis (SASSAD) instruments.
 - Continuous scores on these instruments converted to accepted cut-points: “mild,” “moderate,” and “severe”
- **Patients disease severity self assessment:**
 - “mild,” “moderate,” or “severe”

Skin Disease: Acne, Psoriasis and Atopic Eczema

Agreement: Self-Assessed and Clinical-Assessed Severity for Acne, Psoriasis or Atopic Eczema (n=108)



Weighted kappa (κ) = 0.35 (95% CI: 0.1981, 0.5084)

Skin Disease: Acne, Psoriasis and Atopic Eczema

- **Study Considerations and Limitations**
 - Patients recruited from both general practice and specialist dermatology practice. Findings do not reflect the perceptions only of patients who have been “self-selected” to some extent by referral to specialist or secondary care.
 - Small sample size- agreement across three different diseases pooled approach may fail to detect variation in agreement between particular skin diseases.

Agreement of Self-reported and Clinician-reported Symptoms Example: Cancer Patients

- **Objective, Design, Patient Population**
 - To examine the extent to which patient and clinician symptom scoring and their agreement could contribute to the estimation of overall survival among cancer patients.
 - Retrospective pooled analysis (n=2279) conducted using secondary data from 14 Phase III European Organization for Research and Treatment of Cancer (EORTC) randomized clinical trials (1990-2002).

Example: Cancer- Baseline Symptom Assessment in 14 selected trials

Patient Symptom Burden Assessment: EORTC Quality of Life Core questionnaire (QLQ-C30)

The patient rated his or her symptoms on a 4-point ordinal scale:

- Score of 1 = “not at all,”
- Score of 2 = “a little,”
- Score of 3 = “quite a bit,” and
- Score of 4 = “very much.”

Clinician Assessments of Patient Symptoms: National Cancer Institute’s Common Terminology Criteria for Adverse Events (NCI- CTCAE)

The clinician rated the patient’s symptoms on a 5-point scale:

- Score 0 = “none or normal,”
- Score 1 = “mild,”
- Score 2 = “moderate,”
- Score 3 = “severe,” and
- Score 4 = “life threatening or disabling.”

For purposes of comparison, each of the following pairs were considered to be identical responses:

- EORTC QLQ-C30 score 1 vs NCI-CTCAE score 0;
- EORTC QLQ-C30 score 2 vs NCI-CTCAE score 1;
- EORTC QLQ-C30 score 3 vs NCI-CTCAE score 2;
- EORTC QLQ-C30 score 4 vs NCI-CTCAE scores 3 and 4 combined.

Example: Cancer- The Mean Scores and 95% Confidence Intervals (CIs) for the Symptoms Pain, Fatigue, Vomiting, Nausea, Diarrhea, and Constipation

Clinical Symptom	Patient Score (EORTC QLQ-C30), Mean (95% CI)	Clinician Score (NCI-CTCAE), Mean (95% CI)
Pain	2.31 (2.26 to 2.36)	2.13 (2.07 to 2.18)
Fatigue	2.10 (2.05 to 2.15)	1.36 (1.33 to 1.40)
Vomiting	1.11 (1.08 to 1.14)	1.18 (1.15 to 1.21)
Nausea	1.38 (1.35 to 1.41)	1.20 (1.16 to 1.24)
Diarrhea	1.27 (1.23 to 1.31)	1.10 (1.08 to 1.12)
Constipation	1.50 (1.44 to 1.56)	1.11 (1.09 to 1.14)

 High Variability

Example: Cancer- Comparisons between Clinician and Patient Assessments of Pain, Fatigue, Vomiting, Nausea, Diarrhea, and Constipation

Clinician (NCI-CTCAE)	Patient (EORTC QLQ-C30)	r	k (95% Confidence Interval)
Pain	Have you had pain?	0.58	0.29 (0.26 to 0.33)
	Did pain interfere with your daily activities?	0.50	0.27 (0.23 to 0.30)
Fatigue	Did you need to rest?	0.30	0.07 (0.03 to 0.10)
	Have you felt weak?	0.28	0.07 (0.03 to 0.10)
	Were you tired?	0.30	0.08 (0.04 to 0.11)
Vomiting	Have you vomited?	0.32	0.22 (0.13 to 0.30)
Nausea	Have you felt nauseated?	0.32	0.14 (0.10 to 0.18)
Diarrhea	Have you had diarrhea?	0.20	0.14 (0.07 to 0.20)
Constipation	Have you been constipated?	0.38	0.16 (0.11 to 0.21)

Agreement of Self-reported and Clinician-reported Symptoms Example: Cancer Patients

- **Limitations**

- No evidence-based consensus regarding how to compare scoring from patient-reported vs. clinician-reported measurements
 - Different purpose of assessment for EORTC QLQ-C30 vs. NCI-CTCAE may explain the rationale for low levels of agreement reported between patients and clinicians at baseline.
- Generalizability- limited to relative asymptomatic population

Summary

- Although study examples show modest agreements between self-reports and clinician-reports, results suggest that clinical studies would benefit from assessment of both self-reported and clinician-reported disease severity and symptom burden.
- Compared to clinician objective severity, self-assessed severity is associated with patients' psychological well-being (Skin Disease Example).
- Further patients provide subjective measure of symptom severity that complements clinician scoring in predicting overall survival (Cancer Example).

Impact of Language



VS



- Level of Education
- Context
- Personal Experience
- Age

Examples...

- Pediatric Populations
 - ***“Average”***
- ***“ Did freezing of gait contribute to your falling in the past 24 hours?”***
- ***“Regurgitation”***
 - *“liquid or food coming up into your throat or mouth”*

Linguistic Validation of a Subject Diary Card for Patients Diagnosed with Crohn's Disease - Zulu

Source	Forward Translation	Interview Analysis	Linguist Feedback	Updated Forward Translation	Updated Back Translation
(Also mark "Yes" if any narcotics were used.)	(Phinda umake "Yebo" uma kunezidakam izwa ezisetshenzis iwe.)	R1 showed concern for the word "izidakamizwa".	The Zulu word in this context may be associated with street/illegal drugs, instead of drugs for medical purposes. The FT and BT are updated to the safer term that merely means "opioids" or "medical drugs".	(Phinda umake u- "Yebo" uma kuneminye imithi yokwelapha esetshenzisiwe.)	(Also mark "Yes" if any pain medication was used.)

Similar feedback in Afrikaans, Xhosa, and English (South Africa)

Linguistic Validation of TAPQoL for Parents of Children under 5 – Italian & Japanese

Clinician Review

Forward Trans.	Back Trans.	Clinician Feedback	Linguist Feedback	Updated Forward Translation	Updated Back Translation
Coliche?	Colics?	Coliche addominali is more precise.	Both the terms "coliche" and "coliche addominali" are correct. "Coliche addominali" is more precise and is understandable as the term "coliche". The FT and BT are revised.	Coliche addominali?	Abdominal colics?

Forward Trans.	Back Trans.	Clinician Feedback	Linguist Feedback	Updated Forward Translation	Updated Back Translation
疝痛	Colic	激しい反復性腹痛 (severe recurrent abdominal pain) 「疝痛」 is a right English term for Japanese physicians, but it seems to be a technical term which is unfamiliar to general people.	The clinician suggests that the medical term "colic" is not widely known among laymen. However, it's important to include the medical term as well. The FT and BT are revised to include the medical term, with an explanation in parentheses.	疝痛(反復する激しい腹痛)	Colic (Repetitive severe abdominal pain)

Linguistic Validation of TAPQoL for Parents of Children under 5 - Italian

Cognitive Interviewing

Source	Forward Translation	Interview Analysis	Linguist Feedback	Updated Forward Translation	Updated Back Translation
Colic	Coliche addominali?	R2 still shows confusion between stomach ache, abdominal pain and colic. He thinks this question is the same as the previous one. R3 says that colics are quite similar to the symptoms outlined in the previous question, so the two questions should be together. R5 believes that colics are diarrhea episodes. The other respondents have no difficulty.	The FT and BT are revised to add a very clear explanation so that colic cannot be confused with another condition.	Coliche (dolori intensi o crampi nella regione addominale)?	Colics (acute pains or cramps in the abdominal area)?

Linguistic Validation of TAPQoL for Parents of Children under 5 - Japanese

Cognitive Interviewing

Source	Forward Translation	Interview Analysis	Linguist Feedback	Updated Forward Translation	Updated Back Translation
Colic	疝痛(反復する激しい腹痛)	All the respondents felt odd about the technical term "疝痛."	The term "colic" is removed from the FT and BT as it is too technical and causes confusion amongst all respondents.	繰り返す激しい腹痛	Repetitive severe abdominal cramps

When to consider...

Adaptations

- ClinRO adapted to PRO
- PRO adapted to ClinRO
- Clinician-administered PRO adapted to PRO
- PRO adapted to Clinician-administered PRO

References

- Amaya-Amaya J. et al. *Arthritis* 2012; doi:10.1155/2012/935187.
- Basch E. The Missing Voice of Patients in Drug Safety Reporting. *N Engl J Med* 2010; 362; 10.
- Chalmers RL, OD, Begley CG, Edrington T, Caffery B, Nelson D, Snyder C, Simpson T. The Agreement Between Self-Assessment and Clinician Assessment of Dry Eye Severity. *Cornea* Volume 24, Number 7, October 2005.
- Chen RC et al. Patient-reported Acute Gastrointestinal Symptoms During Concurrent Chemoradiation Treatment for Rectal Cancer *Cancer* 2010; 116:1879-86.
- Cimino CR, Siders CA, Zesiewicz TA. Depressive Symptoms in Parkinson Disease: Degree of Association and Rate of Agreement of Clinician-Based and Self-Report Measures. *J Geriatr Psychiatry Neurol* 2011 24:199. DOI: 10.1177/0891988711422525
- Eddy L. Item Selection in Self-Report Measures for Children and Adolescents with Disabilities: Lessons from Cognitive Interviews. *Journal of Pediatric Nursing*. 2011 December; 26(6): 559-565.
- Magin PJ, Pond CD, Smith WT, Watson AB, Goode SM. Correlation and agreement of self-assessed and objective skin disease severity in a cross-sectional study of patients with acne, psoriasis, and atopic eczema. *Int J Dermatol*.2011;50(12):1486-90.
- Quinten C, Maringwa J, Gotay CC, Martinelli F, Coens C, Reeve BB, Flechtner H, Greimel E, King M, Osoba D, Cleeland C, Ringash J, Schmucker-Von Koch J, Taphoorn MJ, Weis J, Bottomley A. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J Natl Cancer Inst*. 2011; 103:1851–1858.
- Spiegel B, MD. Clinical Trial Design and Outcome Measurement. IFFGD 10th International Symposium on Functional GI Disorders. April 14, 2013.

The authors would like to thank TransPerfect and Quintiles Outcome for their support.

No funding or grant support was received for this project.

Conclusions / Discussion

- Correlation is generally stronger for discrete measures (symptom/disease presence=yes/no) than for continuous measures (severity, scales)
- Effects of language/culture
- Further research: Correlations over time