



Principles for Presenting Patient-Reported Outcomes to Promote Understanding and Use

*Funded by the
Patient-Centered
Outcomes Research
Institute and Pfizer*

Claire Snyder, PhD
Professor of Medicine, Oncology, and Health Policy & Management
Johns Hopkins Schools of Medicine and Public Health



Patient-Reported Outcomes Tools: Engaging Users & Stakeholders

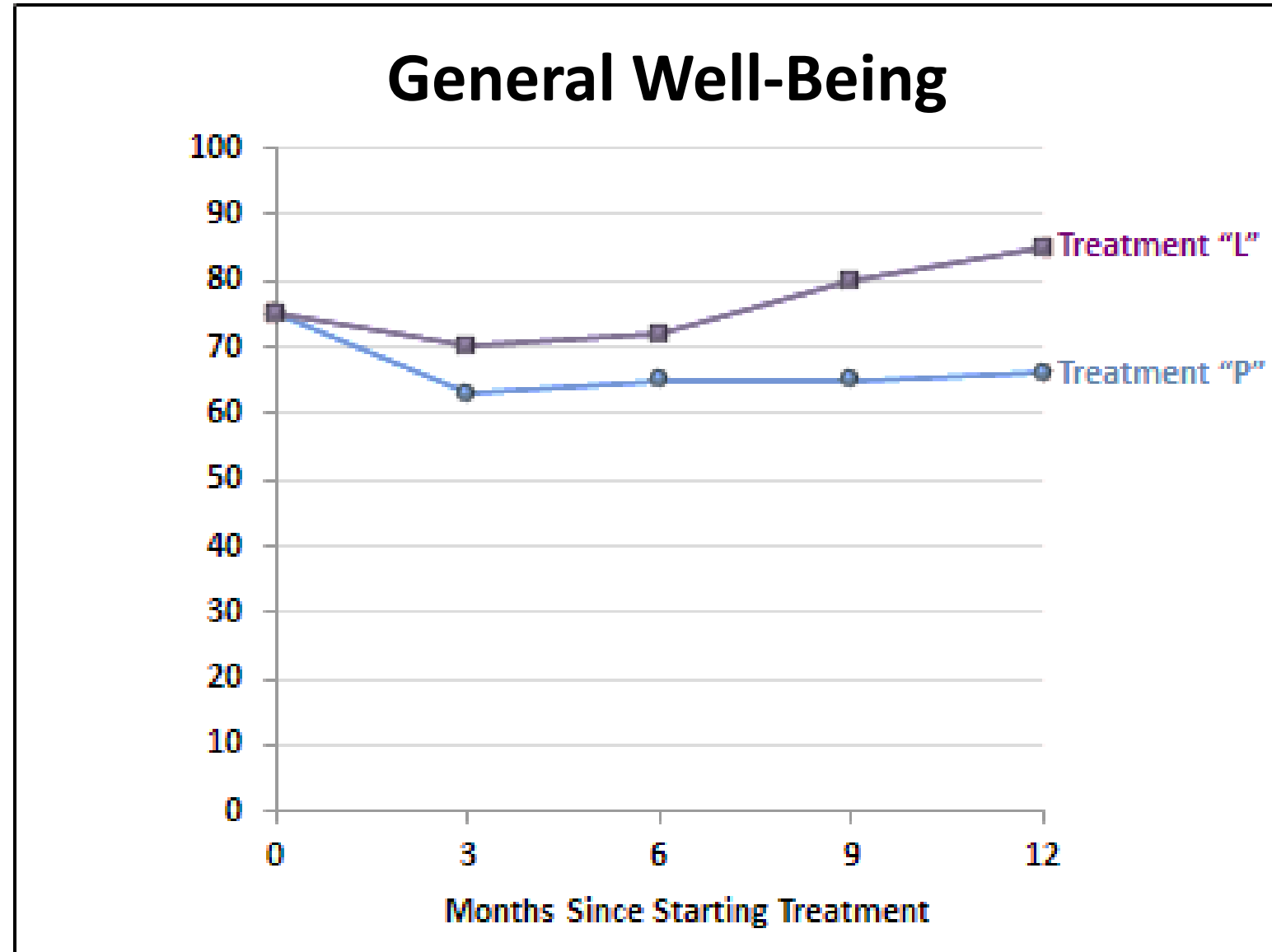
Helping you navigate the use of patient-reported outcomes (PROs)
in clinical trials and clinical practice

TheProteusConsortium.org

POLL QUESTION:

Which
treatment is
better?

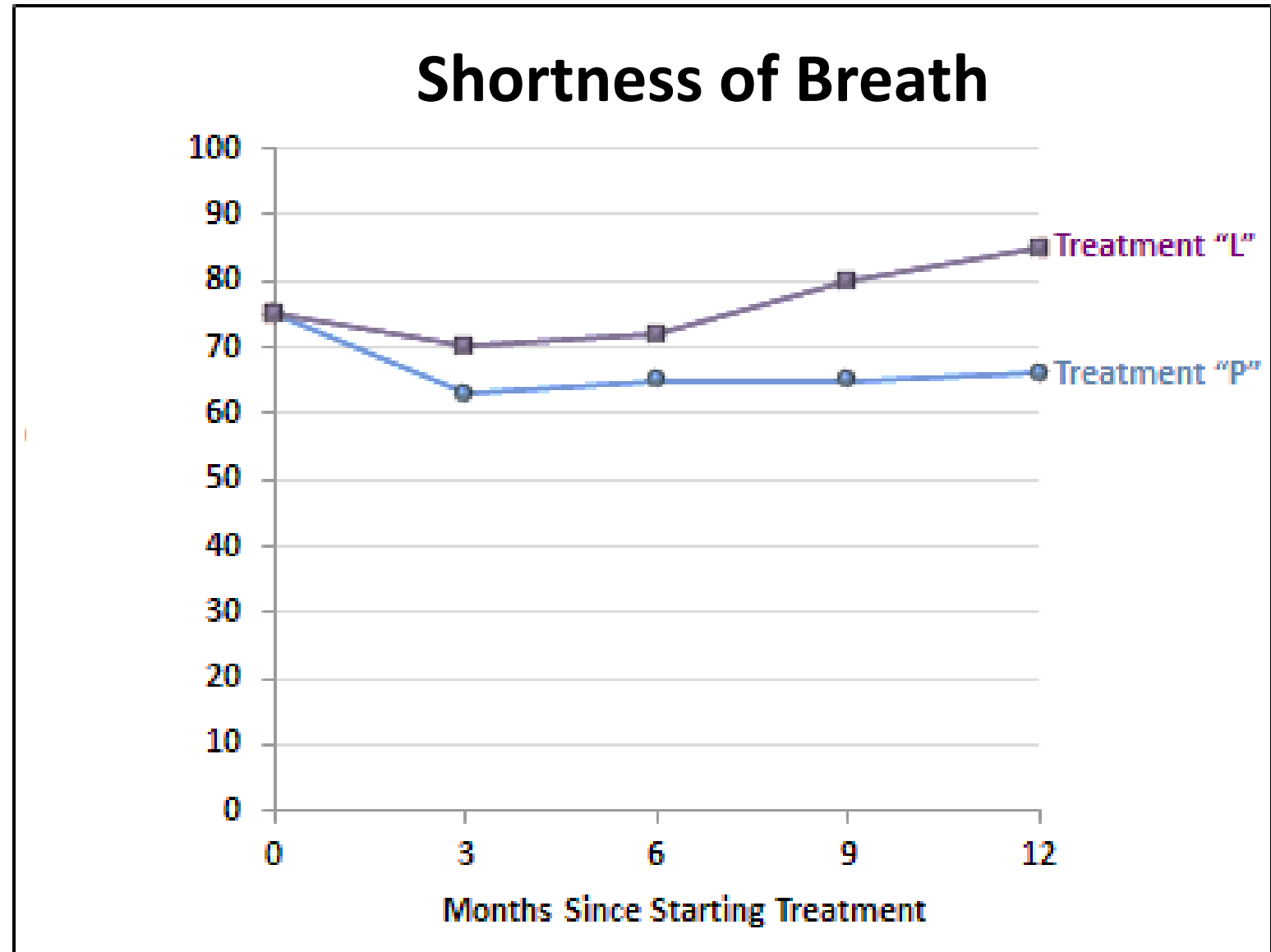
- Treatment L
- Treatment P



POLL QUESTION:

Which treatment is better?

- Treatment L
- Treatment P



- *“Of course I have no idea if this is a good score or a bad score,”*
- *“Until you address the scaling issues it isn't very useful...”*
- *“A score of say, 50, meant one thing on one graph and something different on another one, which I thought was strange.”*
- *“I don't know what the numbers mean,”*



3-Part Mixed Methods Study

1. To what extent do current practices of PRO reporting limit clinician and patient understanding and use? What are the most/least desirable attributes of current practices?
2. What are novel ways to present PRO results to clinicians and patients to improve their usefulness?
3. Are these novel ways of presenting PROs effective in improving understanding and use of the data?

Communicating patient-reported outcome scores using graphic formats: results from a mixed-methods evaluation

Michael D. Brundage¹ · Katherine C. Smith^{2,3} · Emily A. Little⁴ · Elissa T. Bantug² · Claire F. Snyder^{2,3,4} · The PRO Data Presentation Stakeholder Advisory Board

Engaging stakeholders to improve presentation of patient-reported outcomes data in clinical practice

Katherine C. Smith^{1,4} · Michael D. Brundage² · Elliott Tolbert³ · Emily A. Little³ · Elissa T. Bantug⁴ · Claire F. Snyder^{3,4} · PRO Data Presentation Stakeholder Advisory Board

Graphical displays of patient-reported outcomes (PRO) for use in clinical practice: What makes a pro picture worth a thousand words?

Elissa T. Bantug^{a,*}, Theresa Coles^b, Katherine C. Smith^{a,c}, Claire F. Snyder^{a,c,d}, Julie Rouette^e, Michael D. Brundage^e On the behalf of: the PRO Data Presentation Stakeholder Advisory Board

What Do These Scores Mean? Presenting Patient-Reported Outcomes Data to Patients and Clinicians to Improve Interpretability

Claire F. Snyder, PhD^{1,2,3}; Katherine C. Smith, PhD^{2,3}; Elissa T. Bantug, MHS³; Elliott E. Tolbert, PhD^{1,2}; Amanda L. Blackford, ScM³; and Michael D. Brundage, MD, MSc⁴; and the PRO Data Presentation Stakeholder Advisory Board

Original Article

MDM
Medical Decision Making

Medical Decision Making
1–12
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0272989X18791177
journals.sagepub.com/home/mdm
SAGE

Picture This: Presenting Longitudinal Patient-Reported Outcome Research Study Results to Patients

Elliott Tolbert, Michael Brundage, Elissa Bantug, Amanda L. Blackford, Katherine Smith, Claire Snyder, and PRO Data Presentation Stakeholder Advisory Board

Quality of Life Research
<https://doi.org/10.1007/s11136-018-2065-3>

In proportion: approaches for displaying patient-reported outcome research study results as percentages responding to treatment

Elliott Tolbert^{1,2} · Michael Brundage³ · Elissa Bantug⁴ · Amanda L. Blackford^{4,5} · Katherine Smith^{2,4,6} · Claire Snyder^{1,2,4,7} · PRO Data Presentation Stakeholder Advisory Board

Accepted: 22 November 2018
© Springer Nature Switzerland AG 2018

Qual Life Res
DOI 10.1007/s11136-017-1710-6



Presenting comparative study PRO results to clinicians and researchers: beyond the eye of the beholder

Michael Brundage^{1,8} · Amanda Blackford² · Elliott Tolbert^{3,4} · Katherine Smith^{5,7} · Elissa Bantug⁷ · Claire Snyder^{3,6,7} · PRO Data Presentation Stakeholder Advisory Board (various names and locations)

Part 4: Delphi Consensus Development

Quality of Life Research
<https://doi.org/10.1007/s11136-018-2020-3>



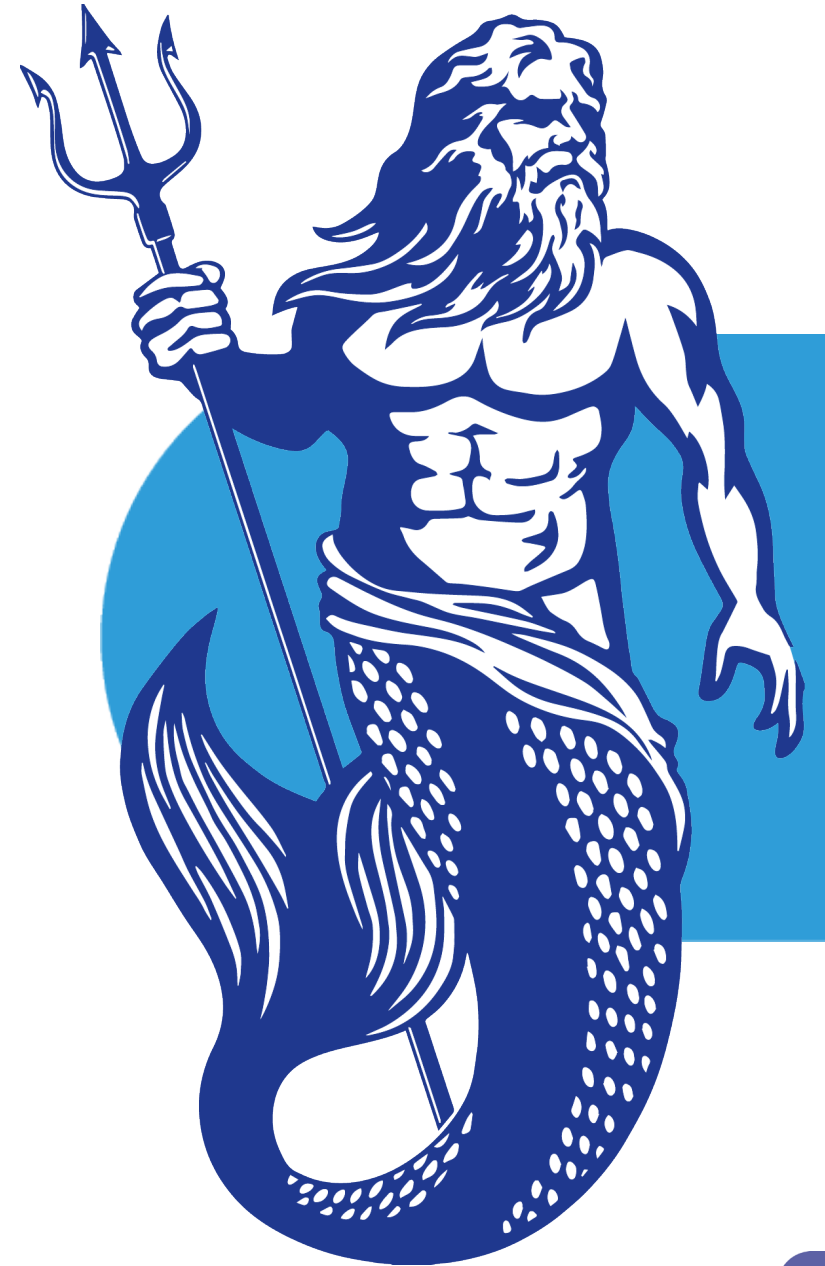
Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data

Claire Snyder^{1,2,3} · Katherine Smith^{2,3} · Bernhard Holzner⁴ · Yonaira M. Rivera² · Elissa Bantug³ · Michael Brundage⁵ · PRO Data Presentation Delphi Panel

Accepted: 29 September 2018
© Springer Nature Switzerland AG 2018

Four Principles

1. Know Your Audience
 - Clinicians Like It Simple
 - Patients Like It Even Simpler
2. There Is No Agreement on What “Up” Means
3. Don’t Over-Estimate Statistical Savvy
4. Research Questions Drive Data Display



Four Principles

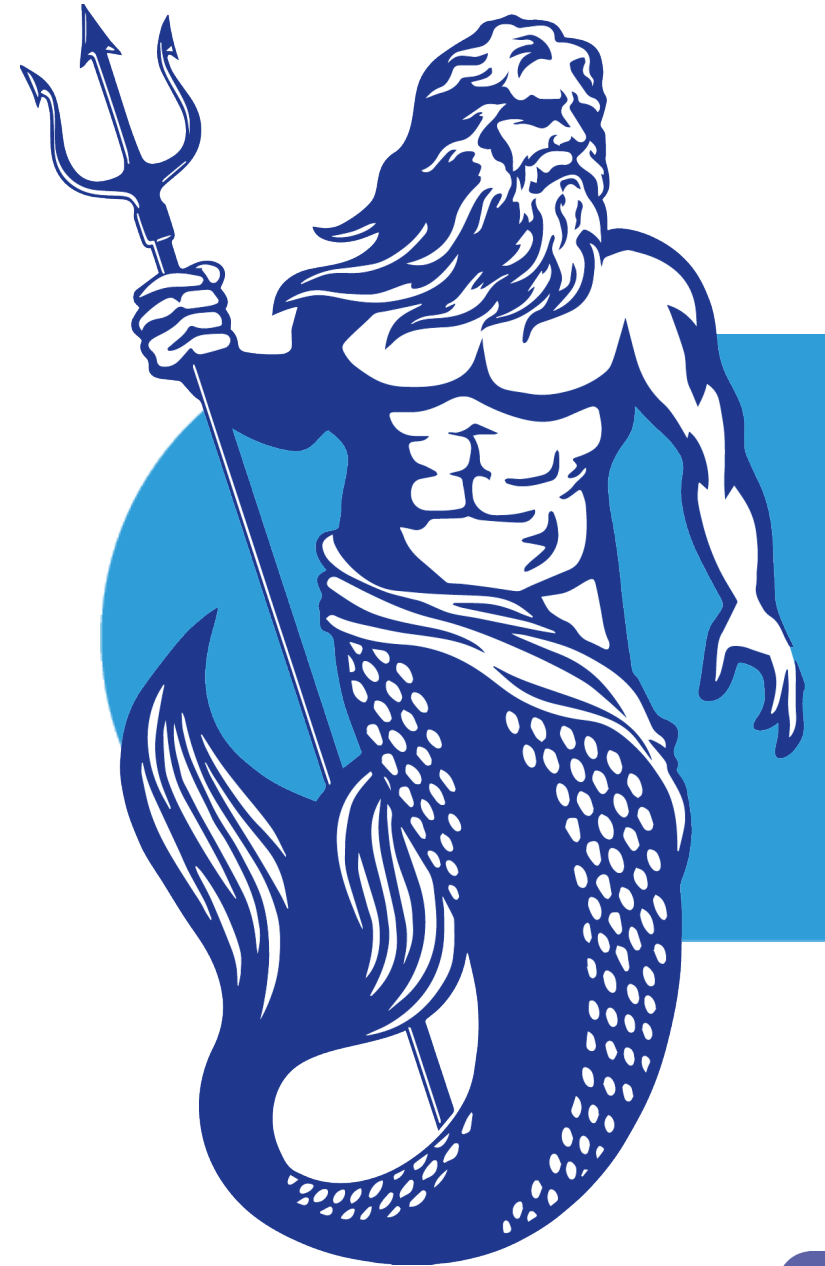
1. Know Your Audience

- Clinicians Like It Simple
- Patients Like It Even Simpler

2. There Is No Agreement on What “Up” Means

3. Don't Over-Estimate Statistical Savvy

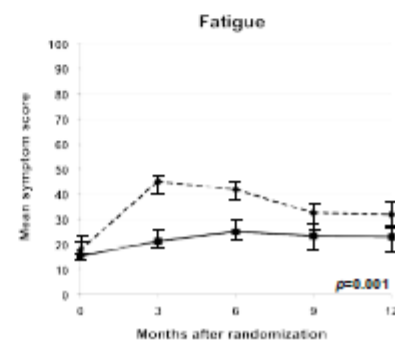
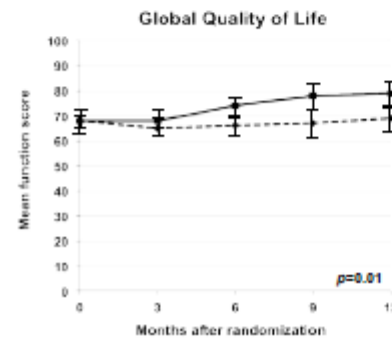
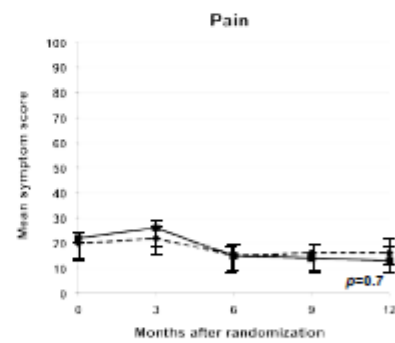
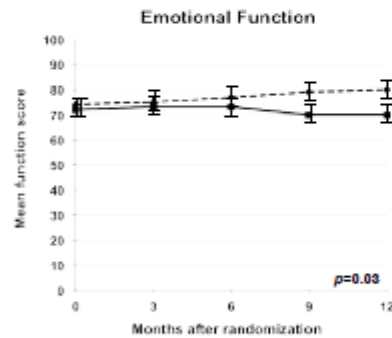
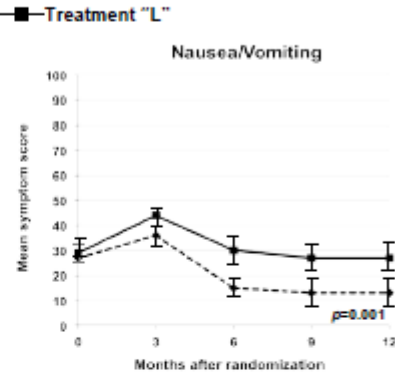
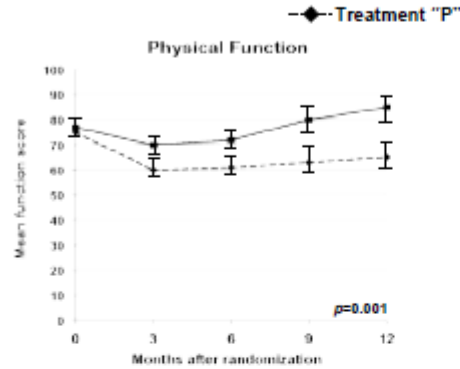
4. Research Questions Drive Data Display



The lines represent the average (mean) scores of patients on each treatment at baseline (0) and at each assessment point up to 12 months. The vertical bars represent variation in the scores at each point (95% confidence limits around the average scores). "p-values" less than 0.05 indicate significant differences between treatments.

Higher scores represent better function or global QOL

Higher scores represent a higher level of symptoms (greater symptom burden)

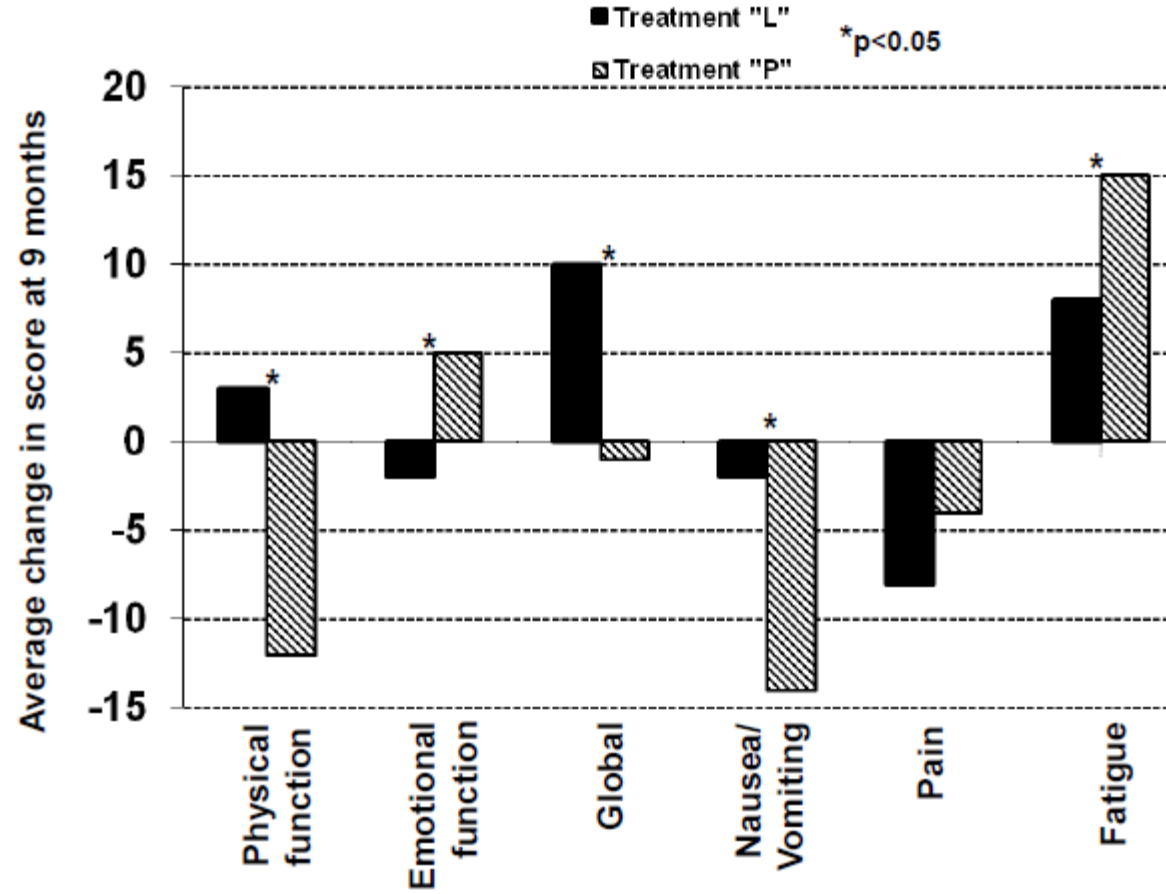


Also shown
-plain (without confidence limits)
-normed to population average

Each bar represents the average change in score at nine months.

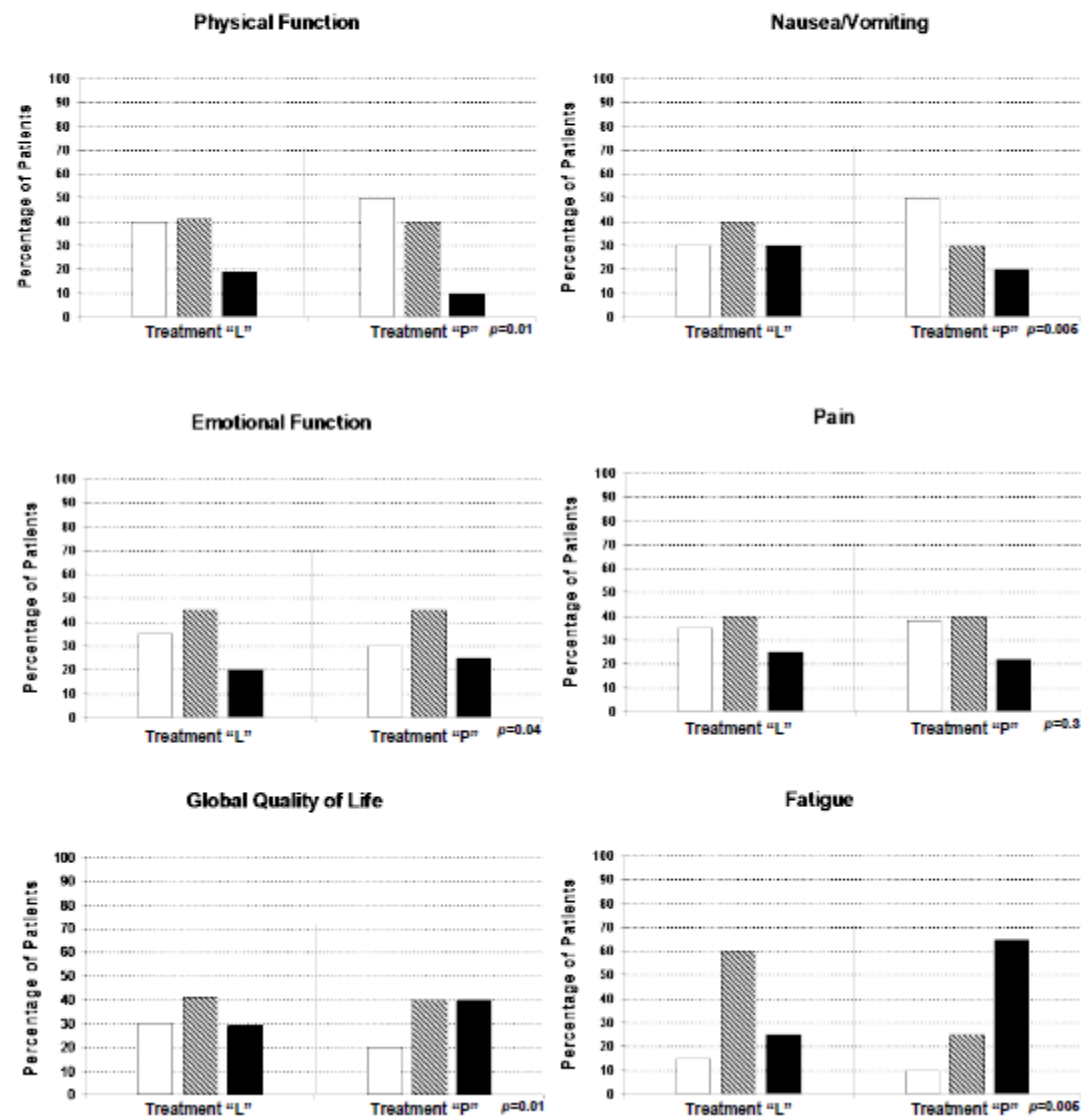
A bar below the "0" line represents improved scores (better functioning or improved symptoms), on average, at nine months compared with prior to starting treatment.

Statistically significant differences between treatments are shown with an *.



The bars represent the percentage of patients that show improved, stable, or worsened scores at nine months. Improved or worsened is defined as a 10 point change (scale range 0-100). "p-values" less than 0.05 are considered significant differences between treatments.

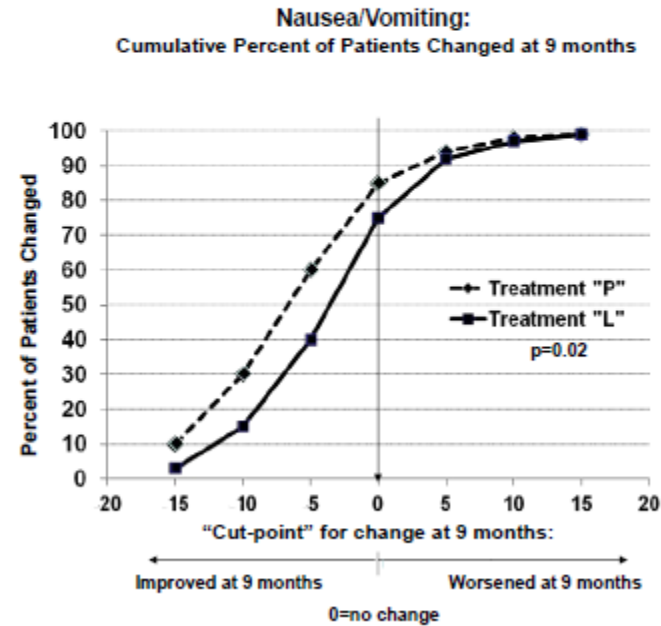
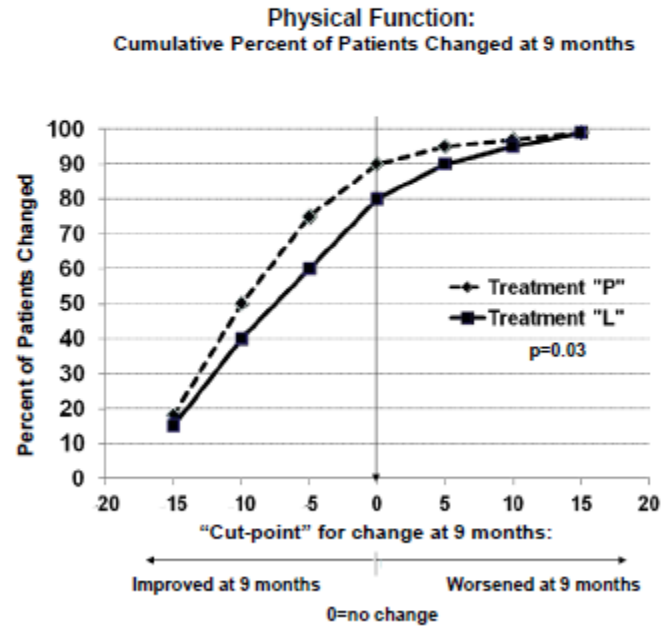
□ Improved at 9 months ▨ Stable at 9 months ■ Worsened at 9 months



The curves represent the cumulative percentage of patients that show a change in scores at nine months.

Various thresholds (up to 15 points improved) for calculating the percentage of patients changing from baseline are shown across the x-axis.

"p-values" less than 0.05 are considered significant differences between treatments.



Selected Qualitative Feedback

- Plain Line Graphs
 - Straightforward, clear
 - Some clinicians missed having confidence limits
 - Some patients didn't understand p-values

“This is pretty straightforward.”

“Well, it just seemed extremely clear.”
- Normed Line Graphs
 - Many clinicians like having something to compare scores to
 - Some patients felt normalized scores too “convoluted” and lacked confidence interpreting
- Line Graphs with Confidence Limits
 - Many clinicians appreciated inclusion of confidence limits
 - Patients did not know what the confidence limits were, were confused by them, and were not sure they added anything

Selected Qualitative Feedback

- Bar Chart of Average Changes
 - Seemed intuitive
 - Easier to compare treatments
 - But only a single point in time
 - Some patients did not understand statistical significance
- Bar Chart of Responders
 - Some liked concept of proportion stable/worsened/improved
 - Single cut-point not as helpful
- Cumulative Distribution Function
 - Unfamiliar, non-intuitive, difficult to understand
 - Inferior accuracy of interpretation
 - “makes a person feel dumb”
 - “it stinks and keeps getting worse”

Summary

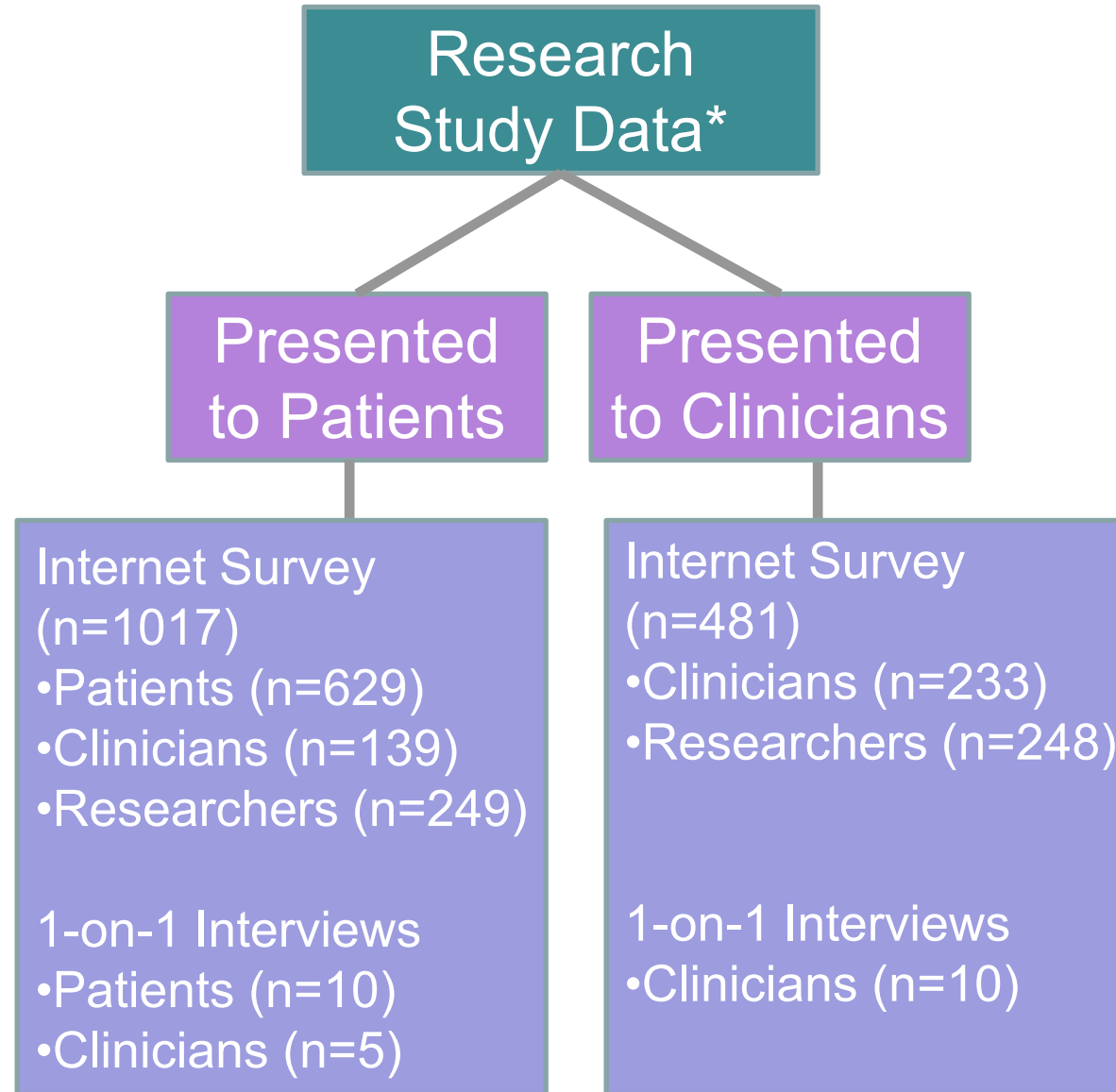
Quantitative Data: Patients

Patients	Under-standing Median	Usefulness Median	Preferred Easy %	Preferred Use %
Line Graphs	8.0	8.0	33	33
LG Normed	7.0	7.0	8	12
LG with C.I.	7.0	7.0	8	8
Bar Chart Averages	7.0	6.0	17	12
Bar Chart Responders	7.0	7.0	20	21
Cumulative Distribution Function	5.0	5.0	4	8

Summary

Quantitative Data: Clinicians

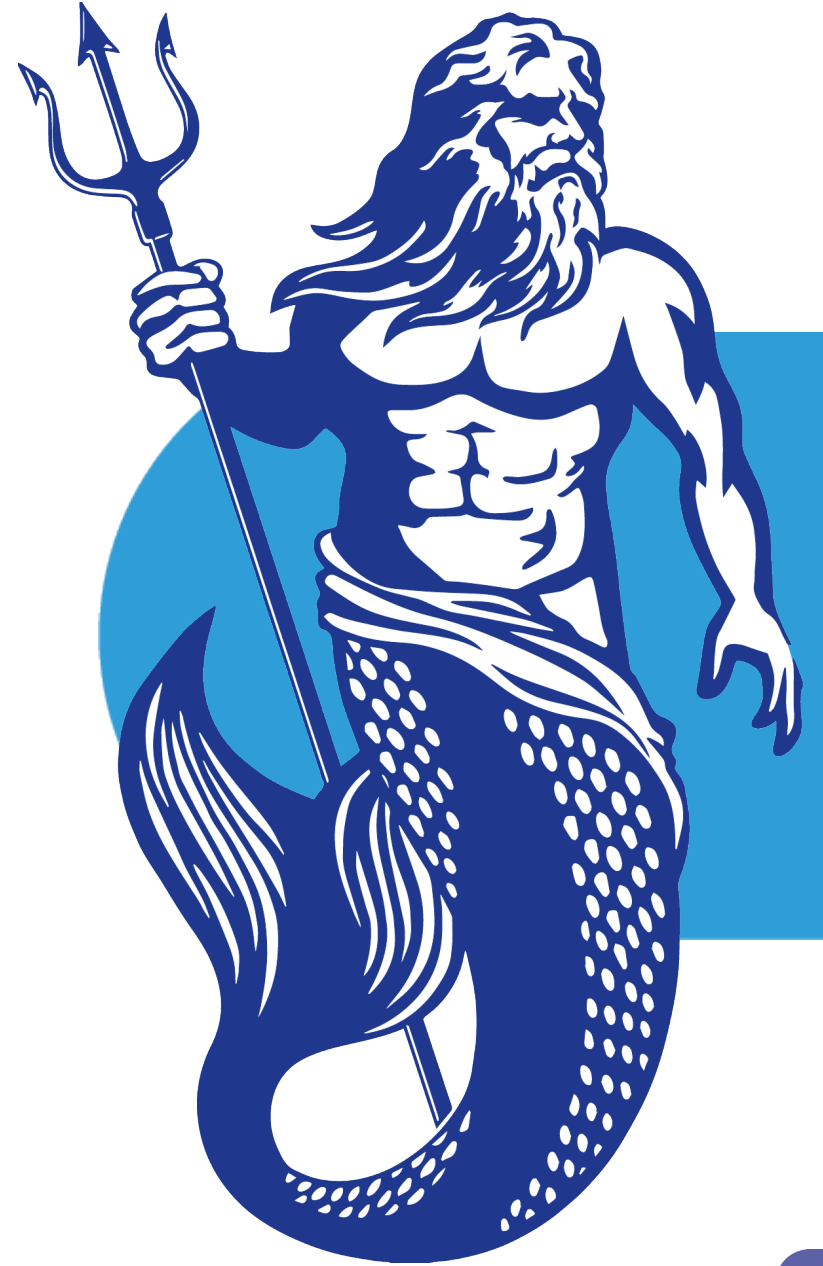
Clinicians	Under-standing Median	Usefulness Median	Preferred Easy %	Preferred Use %
Line Graphs	9.0	8.5	0	0
LG Normed	7.0	8.0	30	30
LG with CI	7.5	8.0	30	30
Bar Chart Averages	8.0	8.0	20	20
Bar Chart Responders	8.0	8.0	20	20
Cumulative Distribution Function	3.0	4.0	0	0



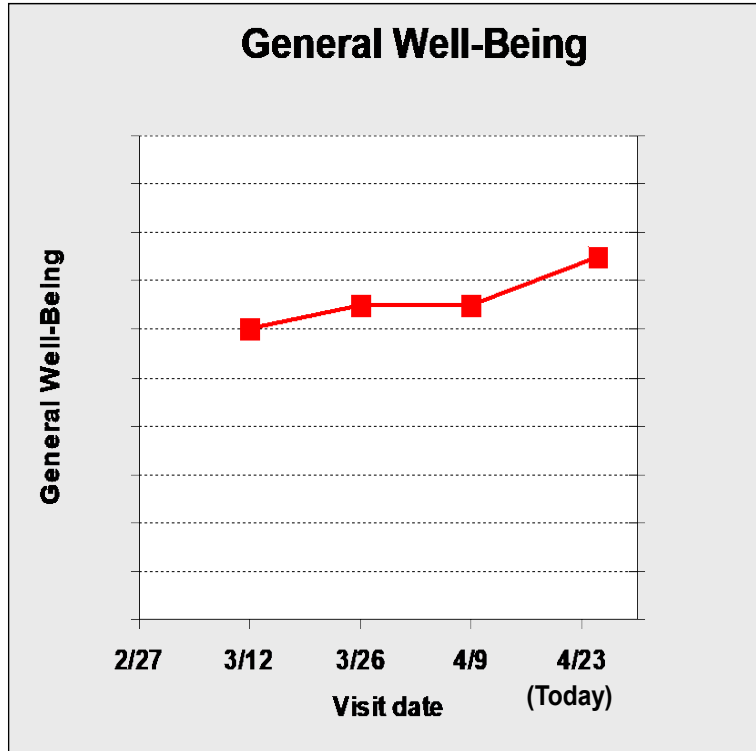
**The project also addressed presentation of individual patient data for monitoring and management (not covered today)*

Four Principles

1. Know Your Audience
 - Clinicians Like It Simple
 - Patients Like It Even Simpler
2. **There Is No Agreement on What “Up” Means**
3. Don't Over-Estimate Statistical Savvy
4. Research Questions Drive Data Display



General Well-Being

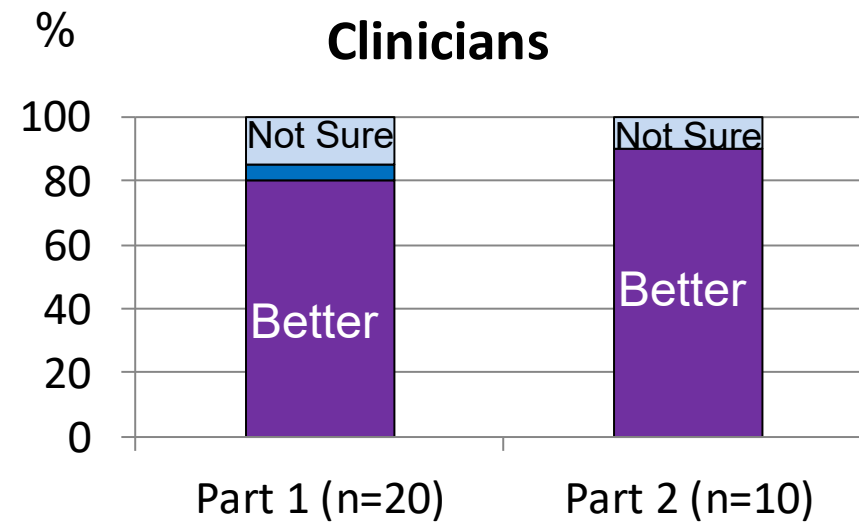
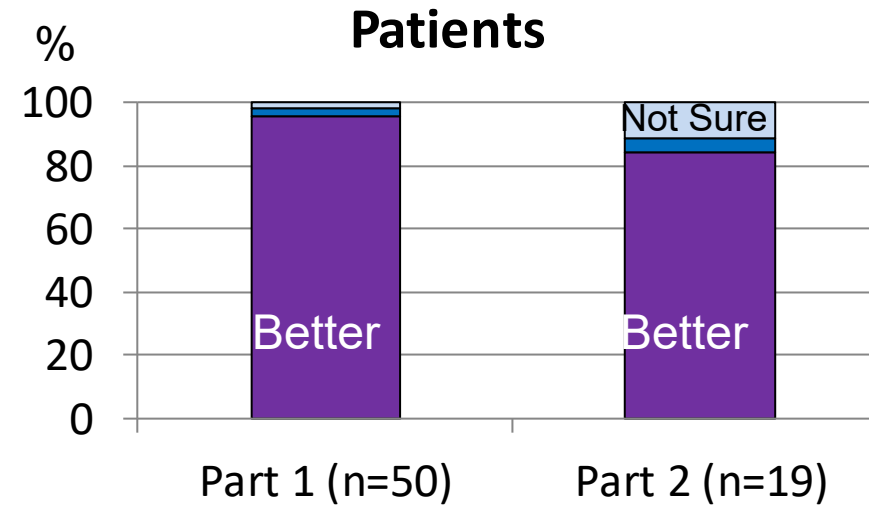


What do you think is happening to this patient's **well-being** over the past few visits?
(circle one)

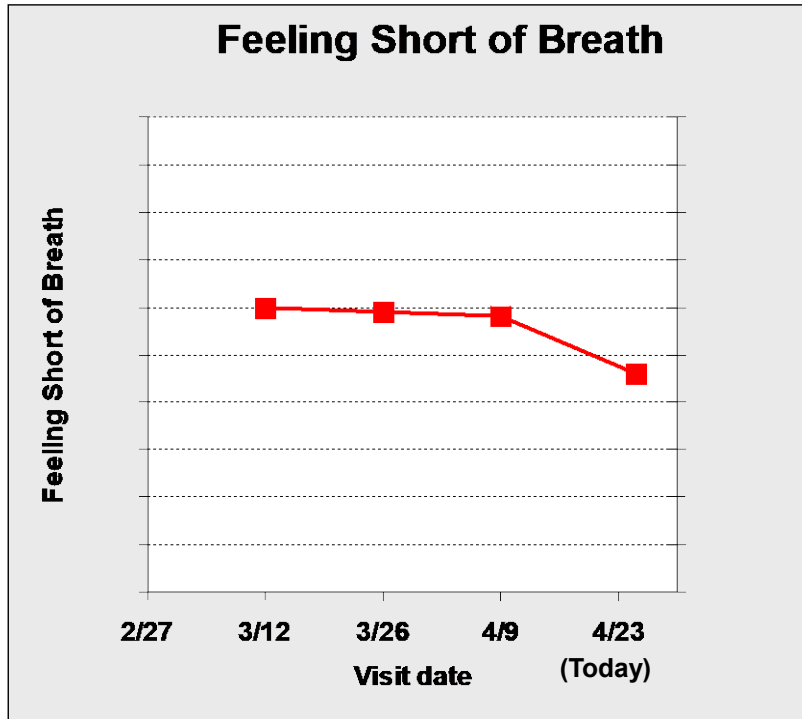
Getting better

Getting worse

Not sure



Shortness of Breath

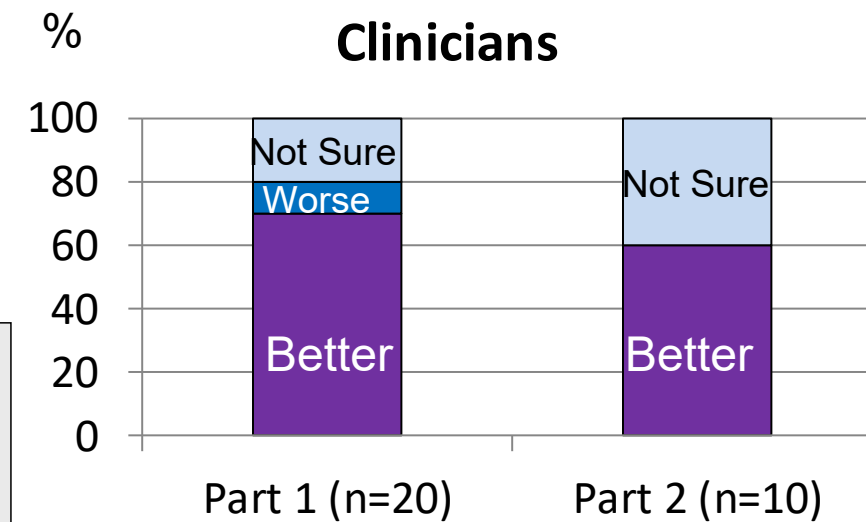
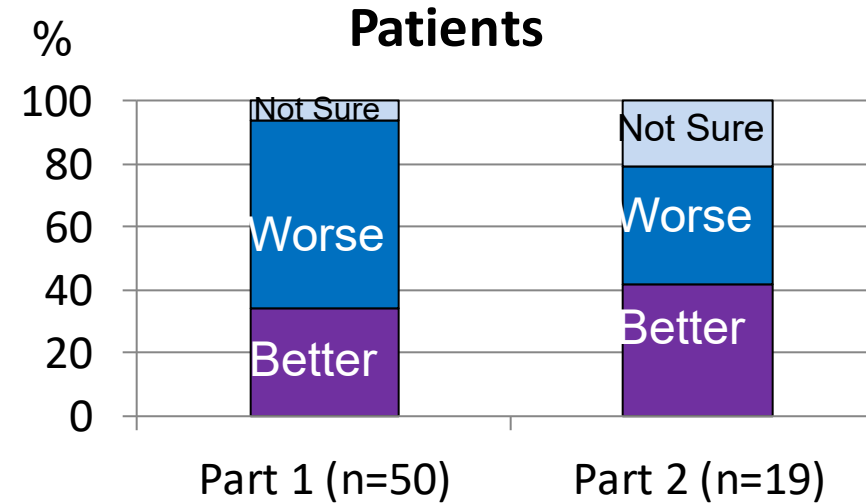


What do you think is happening to this patient's **shortness of breath** over the past few visits? (*circle one*)

Getting better

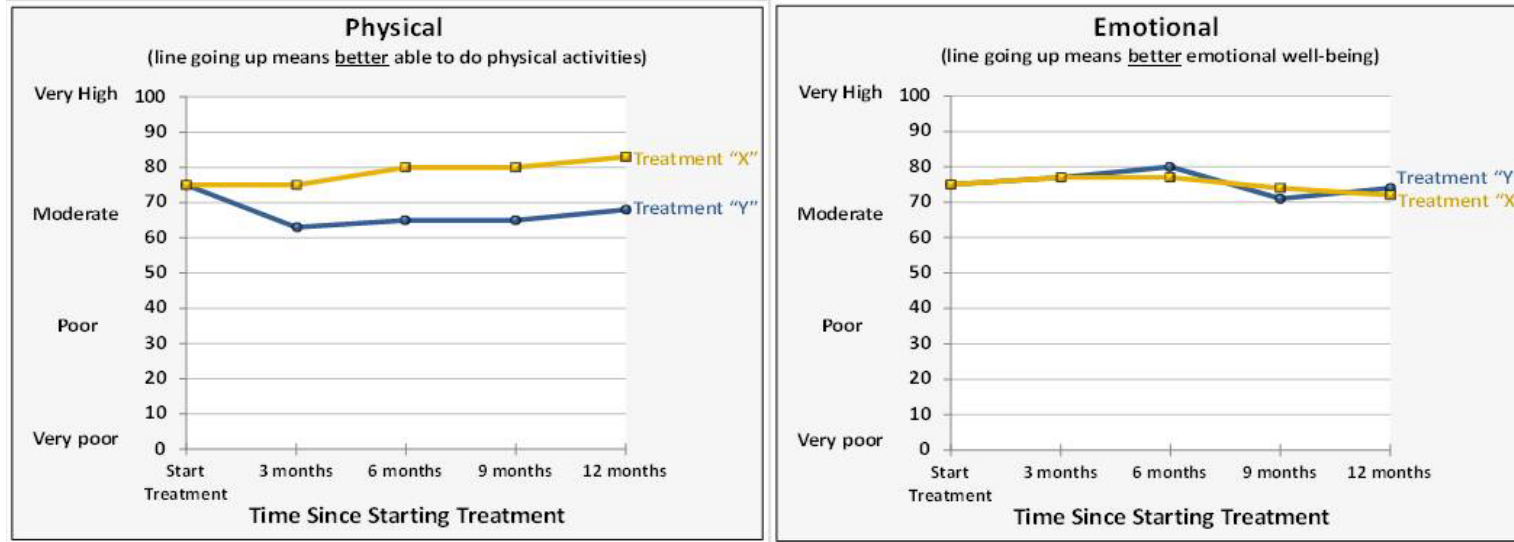
Getting worse

Not sure

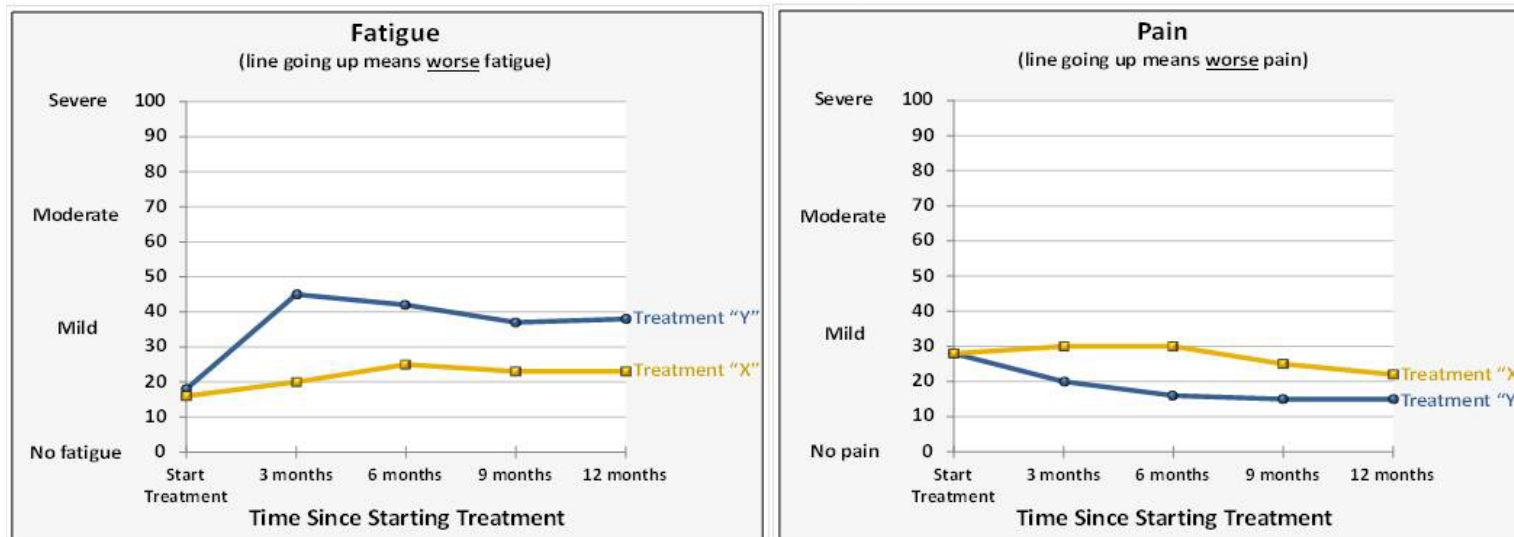


“More” Line Graphs

Patient's Functioning

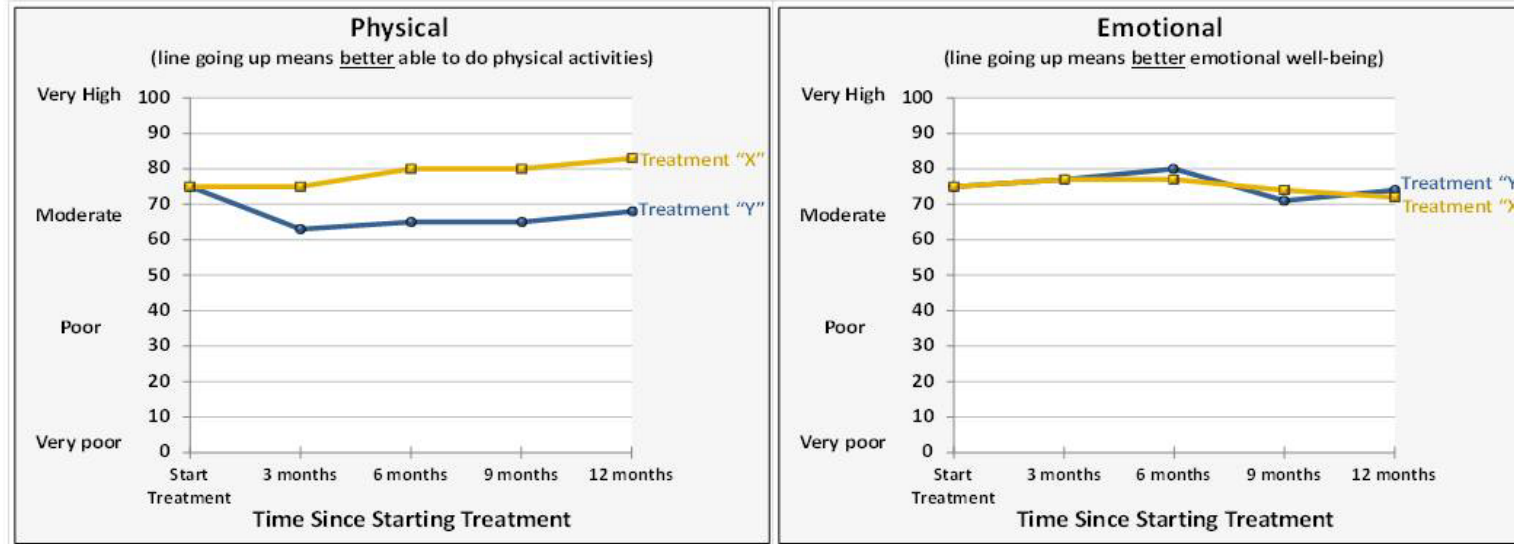


Patient's Symptoms

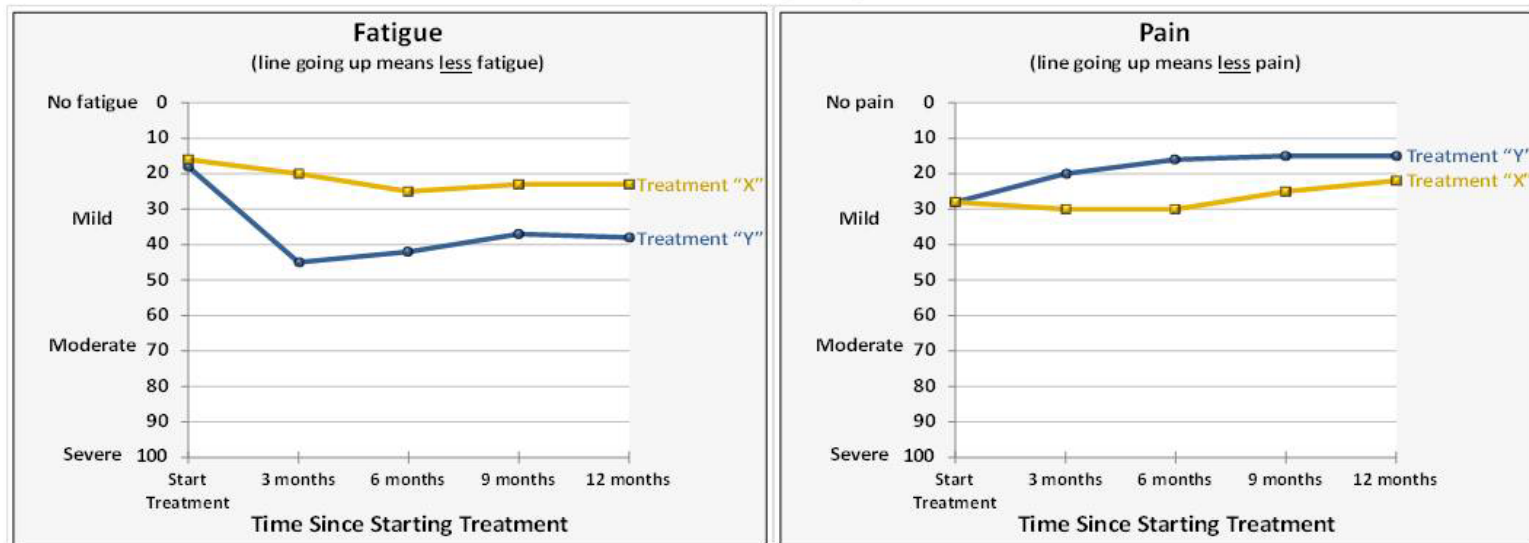


“Better” Line Graphs

Patient's Functioning

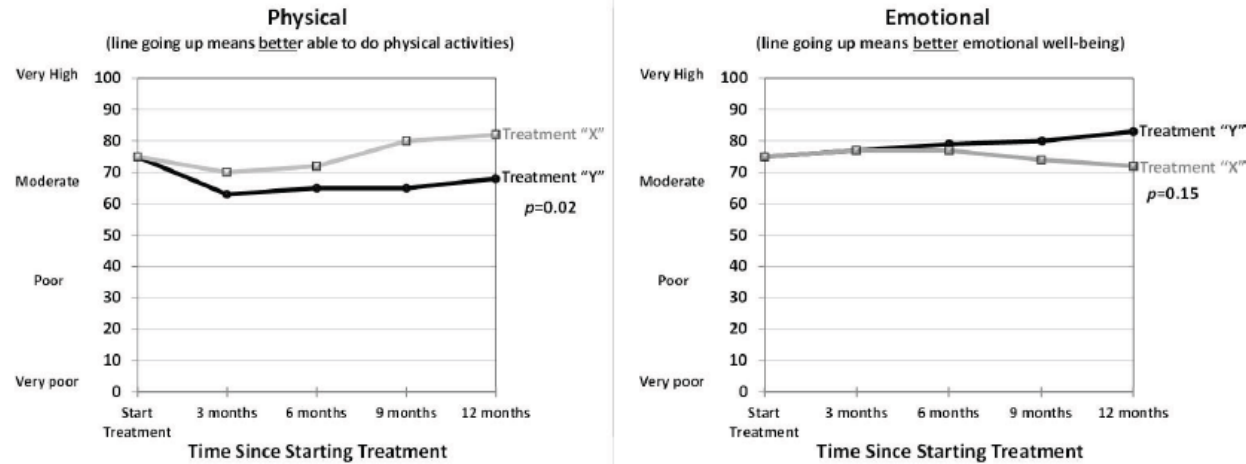


Patient's Symptoms

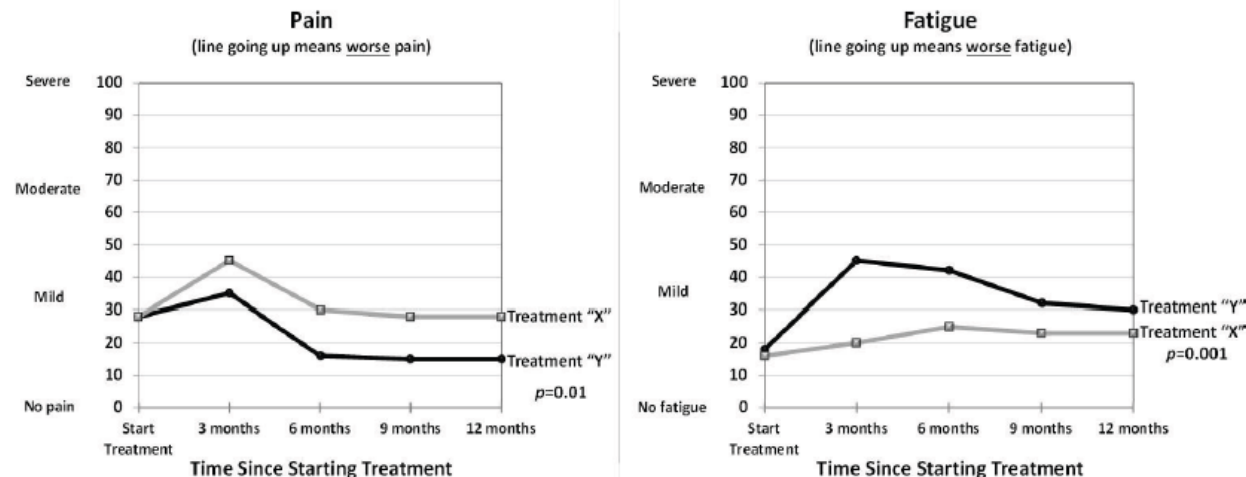


“More” - Clinician Version

Patient's Functioning



Patient's Symptoms



Legend: For all graphs, p -values are for between-treatment differences over time.

Key Findings: Line Graphs

- **FOR PATIENTS:** Line graphs with higher always indicating “**better**” outcomes
 - More accurately interpreted
 - More likely to be rated clear than “more” line graphs
- **FOR CLINICIANS:** No difference in accuracy of interpretation or clarity between “**more**” vs. “**better**” formats



Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data

Claire Snyder^{1,2,3} · Katherine Smith^{2,3} · Bernhard Holzner⁴ · Yonaira M. Rivera² · Elissa Bantug³ · Michael Brundage⁵ · PRO Data Presentation Delphi Panel

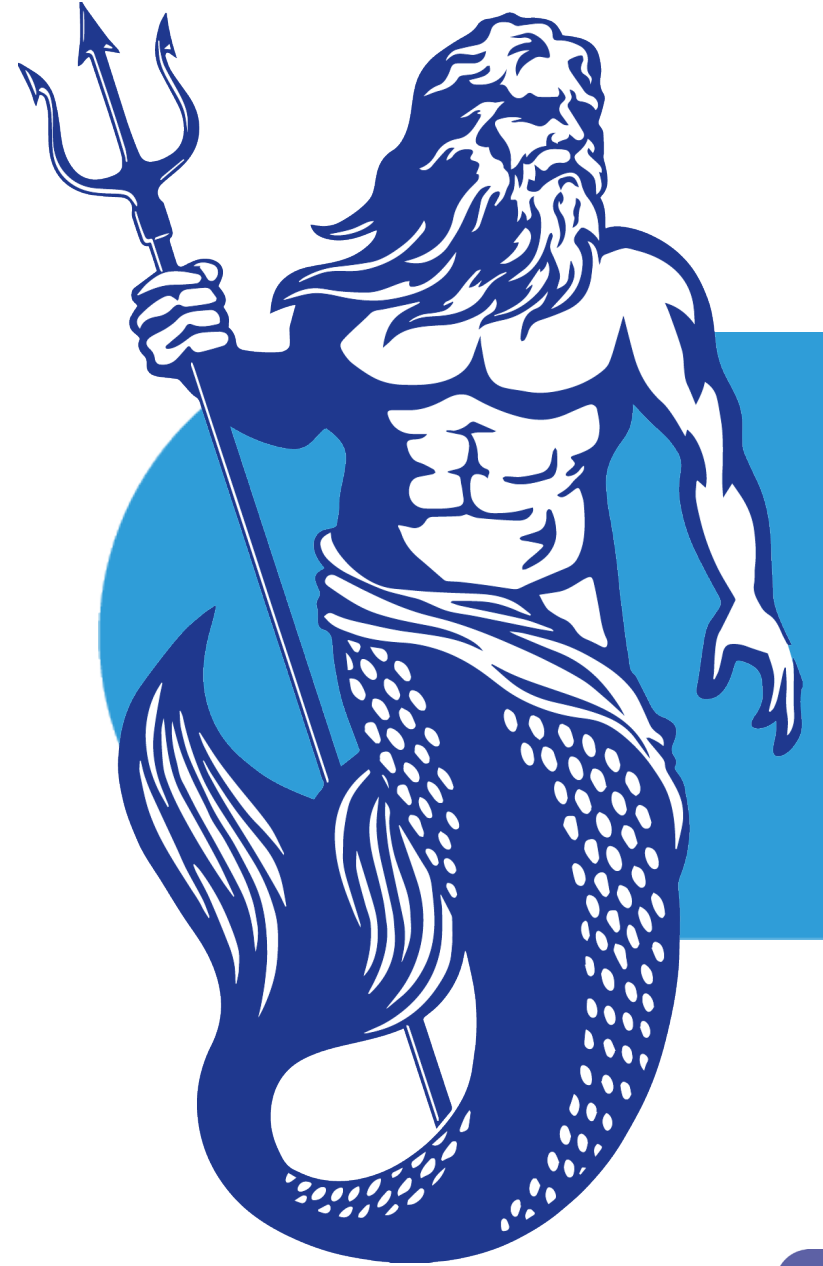
Accepted: 29 September 2018
© Springer Nature Switzerland AG 2018

Directionality/Conveying Score Meaning:

- Use exceptionally clear labeling, titling, and annotations to indicate whether higher scores are better or worse
- If possible, include descriptive labels (e.g., none/mild/moderate/severe) along the y-axis
- Avoid mixing score direction in a single display

Four Principles

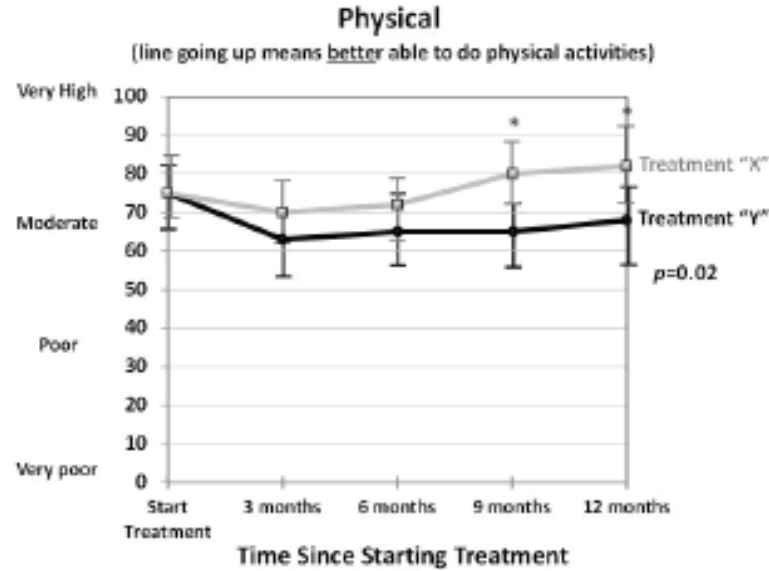
1. Know Your Audience
 - Clinicians Like It Simple
 - Patients Like It Even Simpler
2. There Is No Agreement on What “Up” Means
3. **Don't Over-Estimate Statistical Savvy**
4. Research Questions Drive Data Display



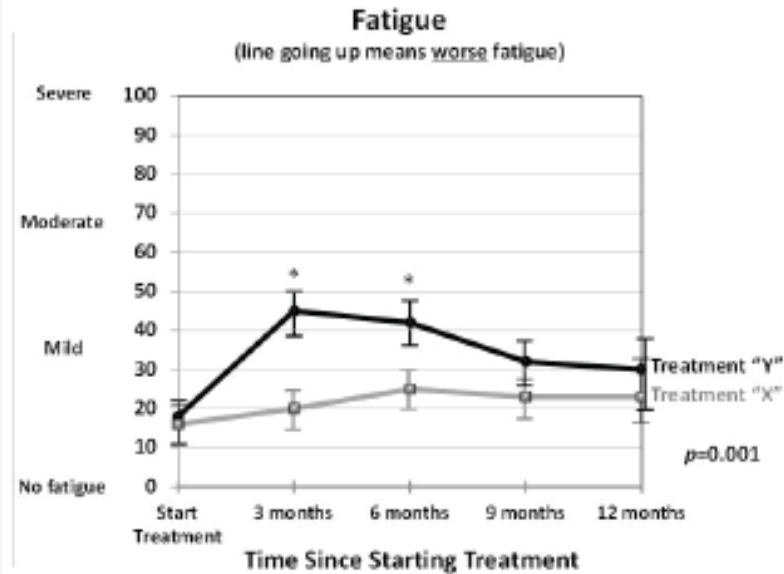


Statistically Significant Differences

Patient's Functioning



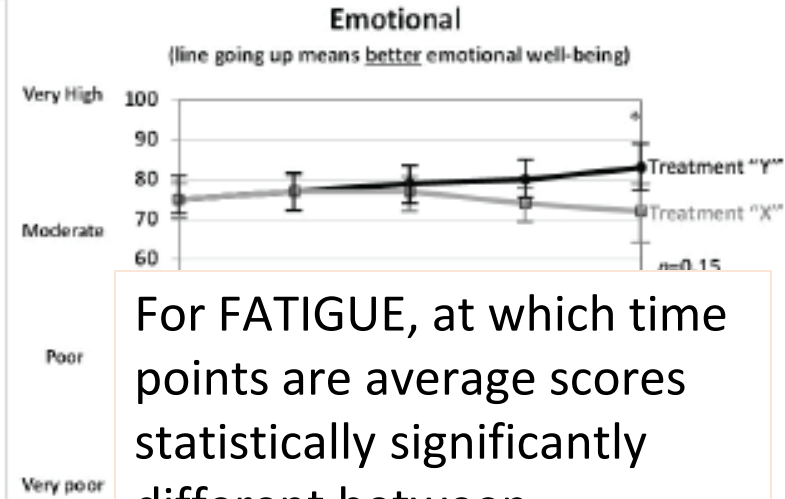
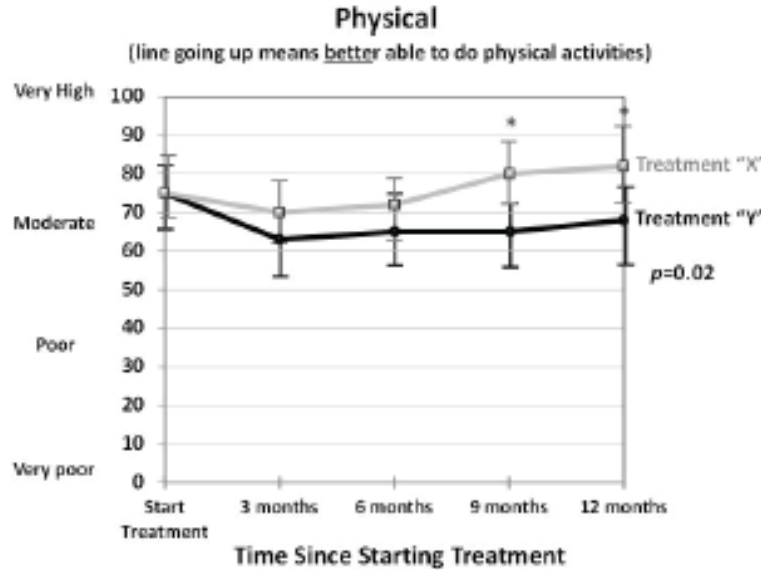
Patient's Symptoms



Legend: For all graphs, p -values are for between-treatment differences over time.

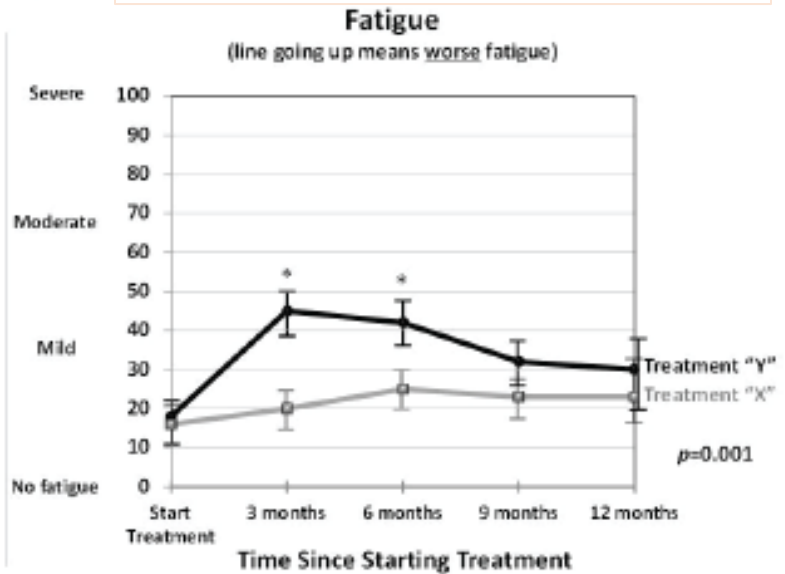
* = differences between treatments that were determined to be clinically important. Vertical lines indicate 95% confidence limits

Patient's Functioning



For FATIGUE, at which time points are average scores statistically significantly different between treatments? (select any that apply)

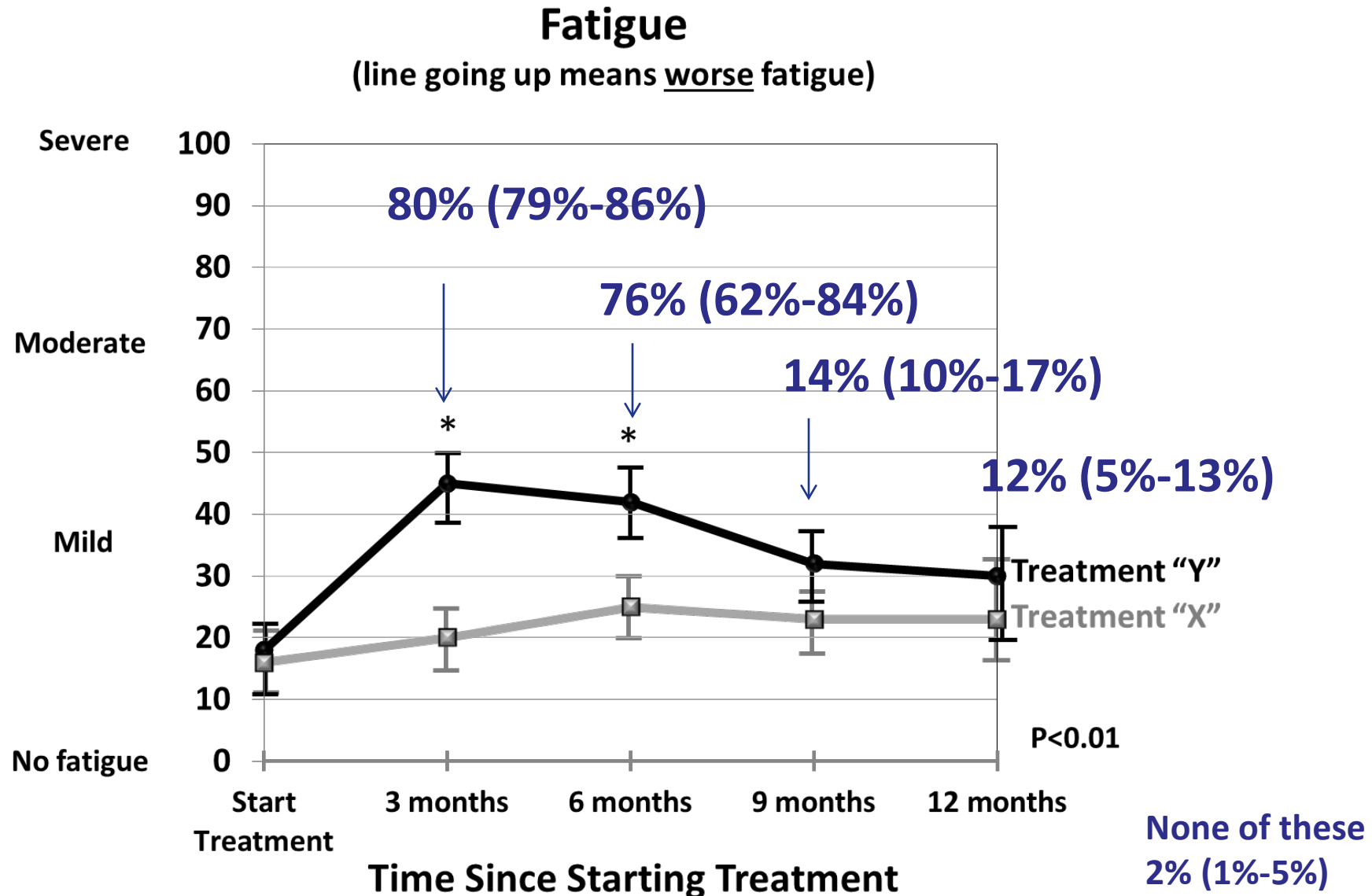
Patient's Symptom



Legend: For all graphs, p -values are for between-treatment differences over time.

* = differences between treatments that were determined to be clinically important. Vertical lines indicate 95% confidence limits

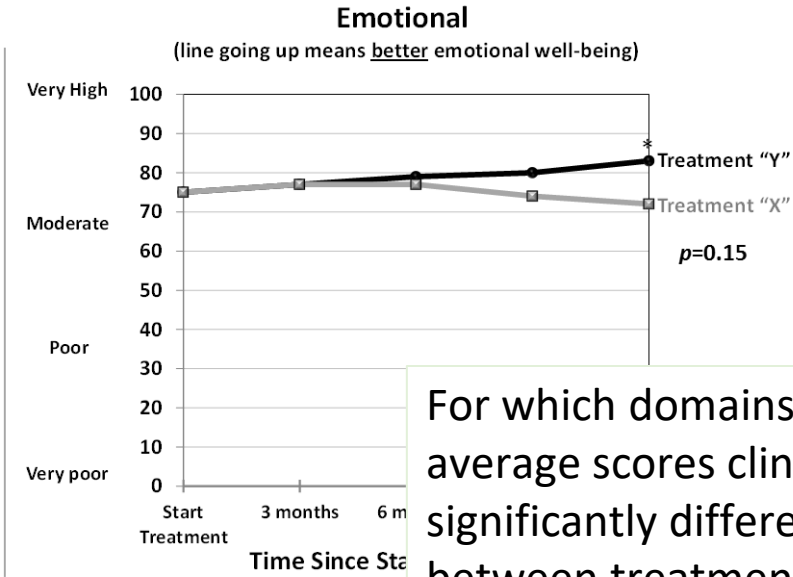
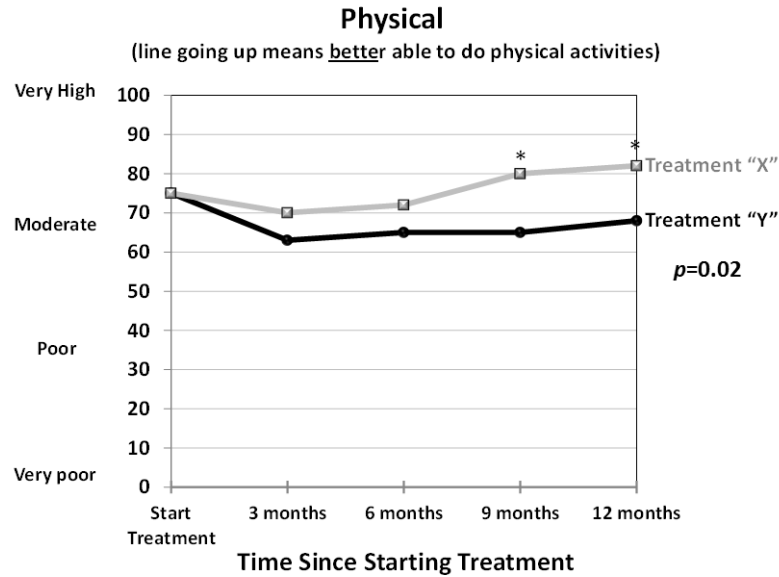
Correct Responses: %Clinicians (range across respondent groups and formats)





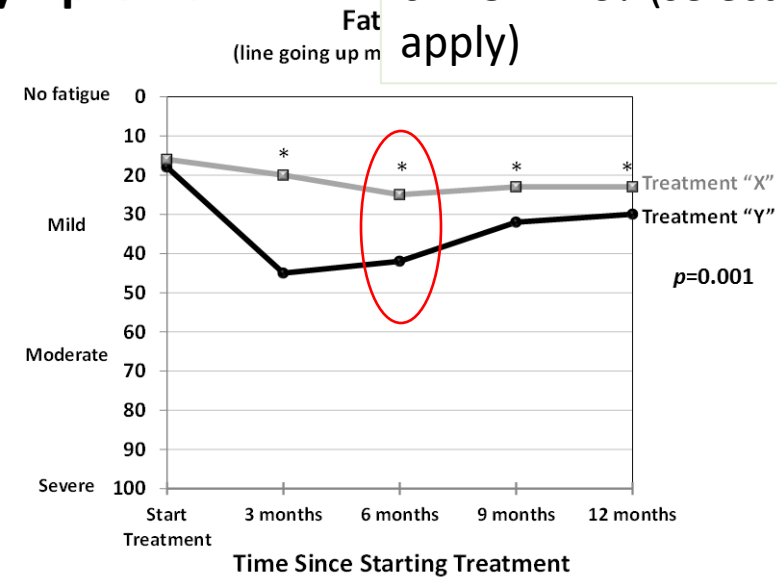
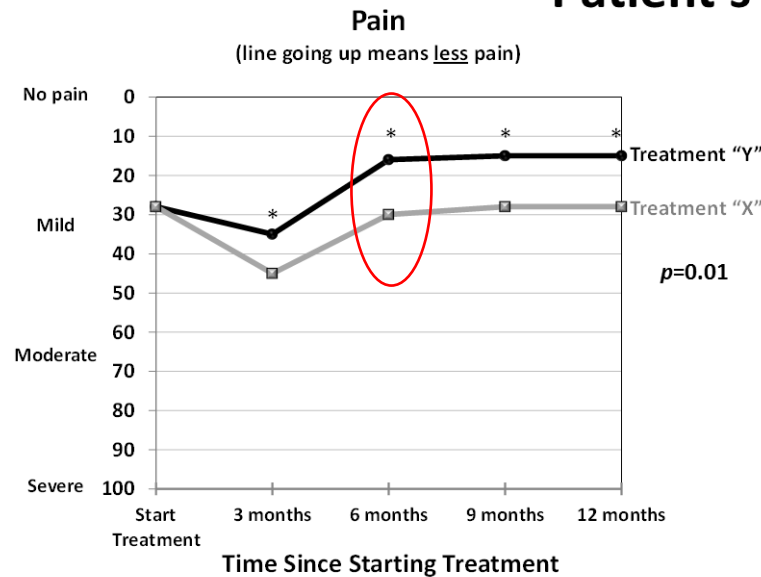
Clinically Important Differences

Patient's Functioning



For which domains are average scores clinically significantly different between treatments at 6 MONTHS? (select all that apply)

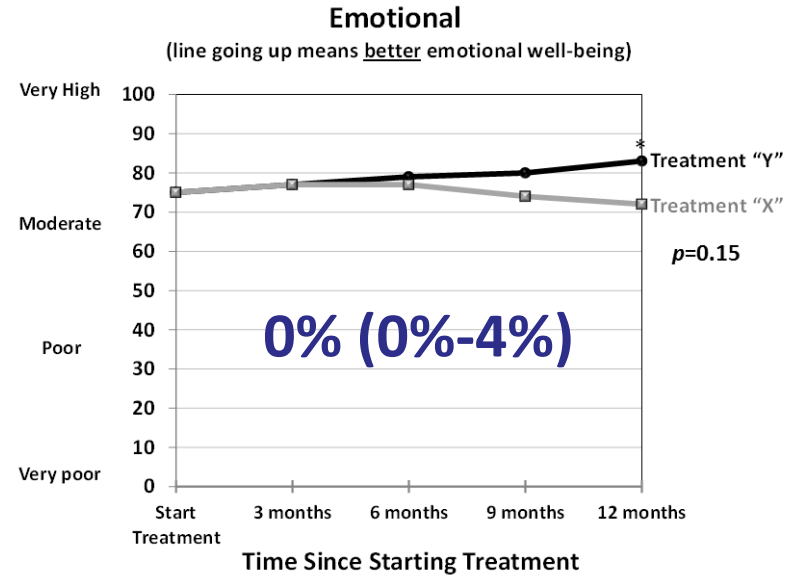
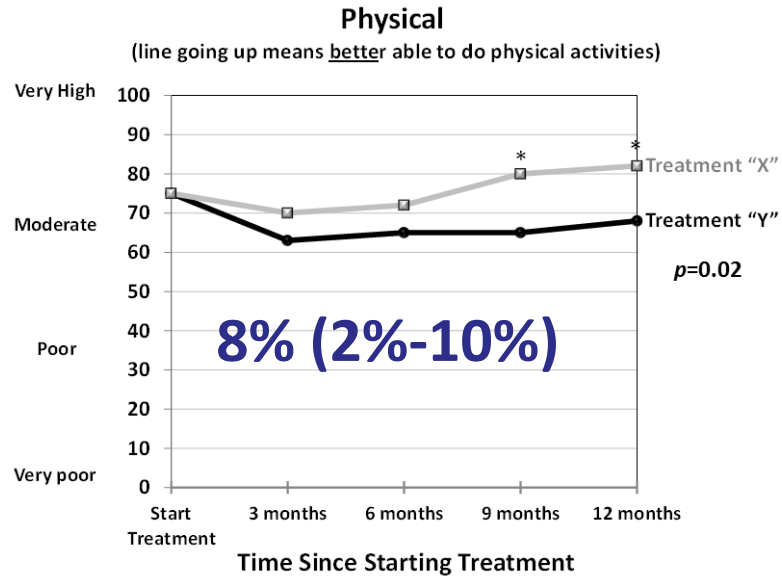
Patient's Symptoms



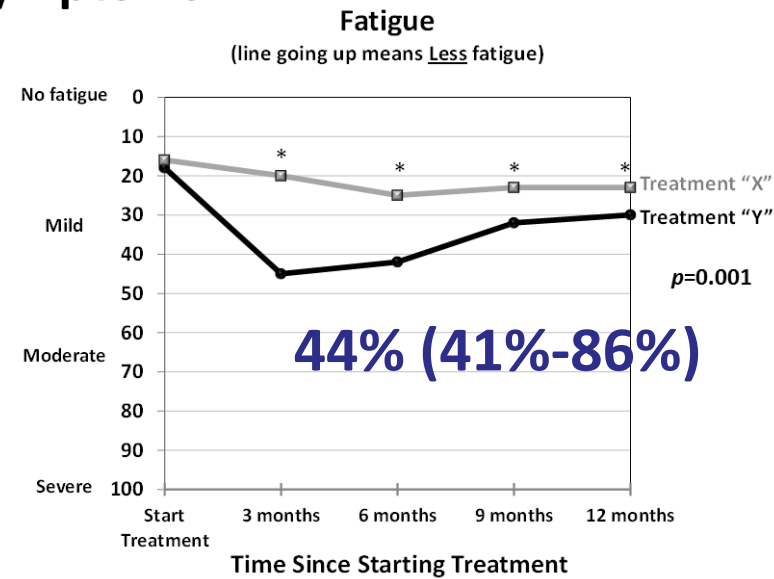
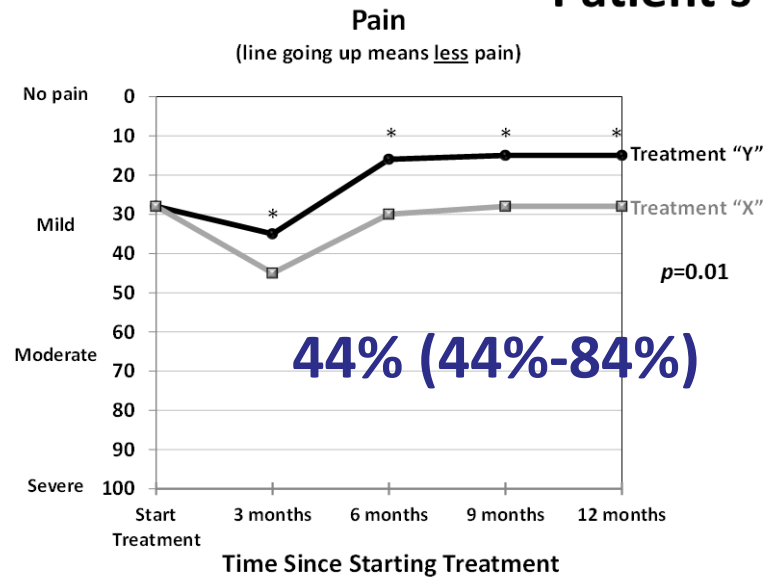
Legend: For all graphs, p -values are for between-treatment differences over time.

*indicates differences between treatments that are clinically important

Patient's Functioning

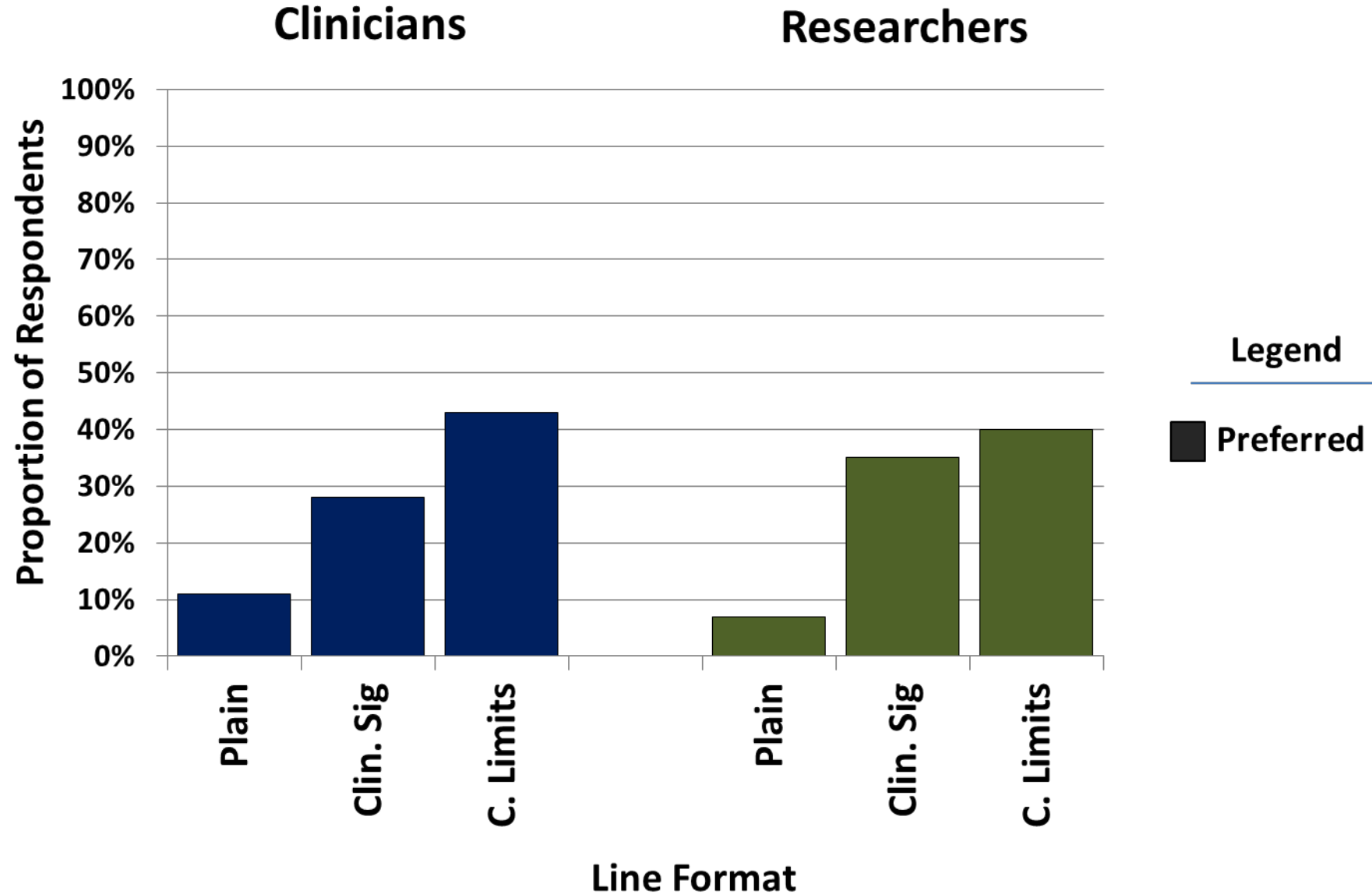


Patient's Symptoms



Legend: For all graphs, p -values are for between-treatment differences over time.
*indicates differences between treatments that are clinically important

Selected Most Useful: Lines





Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data

Claire Snyder^{1,2,3} · Katherine Smith^{2,3} · Bernhard Holzner⁴ · Yonaira M. Rivera² · Elissa Bantug³ · Michael Brundage⁵ · PRO Data Presentation Delphi Panel

Accepted: 29 September 2018
© Springer Nature Switzerland AG 2018

Conveying statistical significance (for clinicians and researchers only):

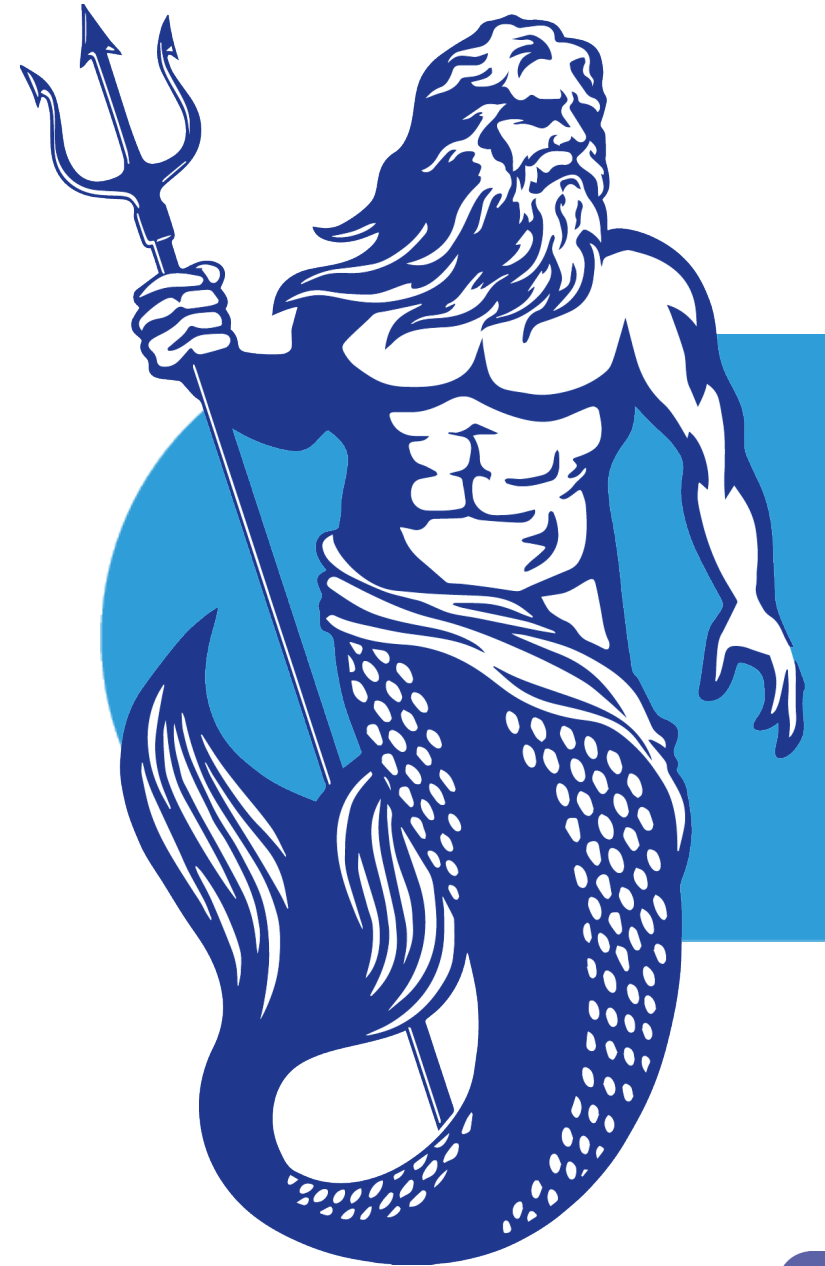
- Confidence intervals should be displayed
- P-values may also be appreciated

Clinically important differences:

- Include an indication of what differences between scores matter

Four Principles

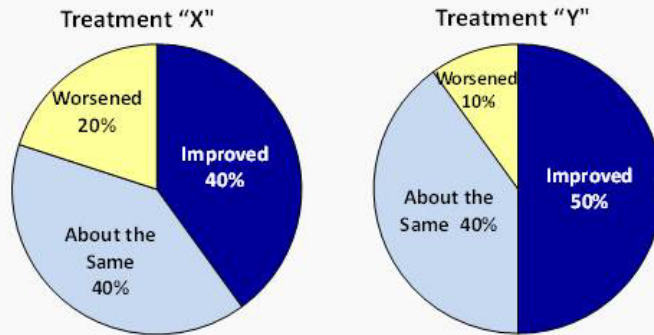
1. Know Your Audience
 - Clinicians Like It Simple
 - Patients Like It Even Simpler
2. There Is No Agreement on What “Up” Means
3. Don’t Over-Estimate Statistical Savvy
4. **Research Questions Drive Data Display**



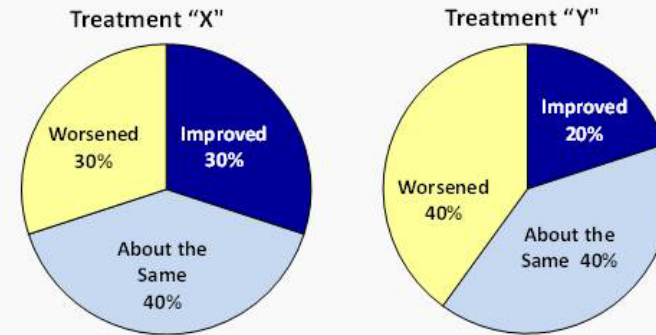
Pies

Status of 100 patients 9 months after starting treatment

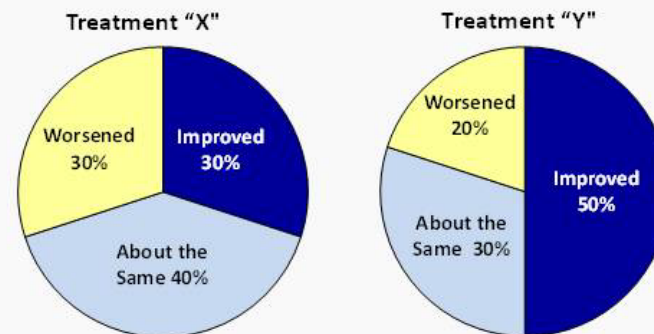
Ability to Do Physical Activities



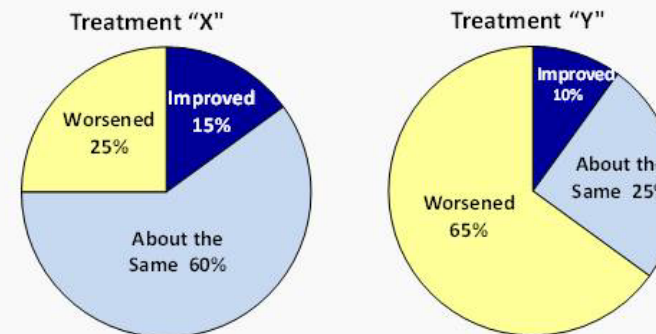
Emotional Well-Being



Pain

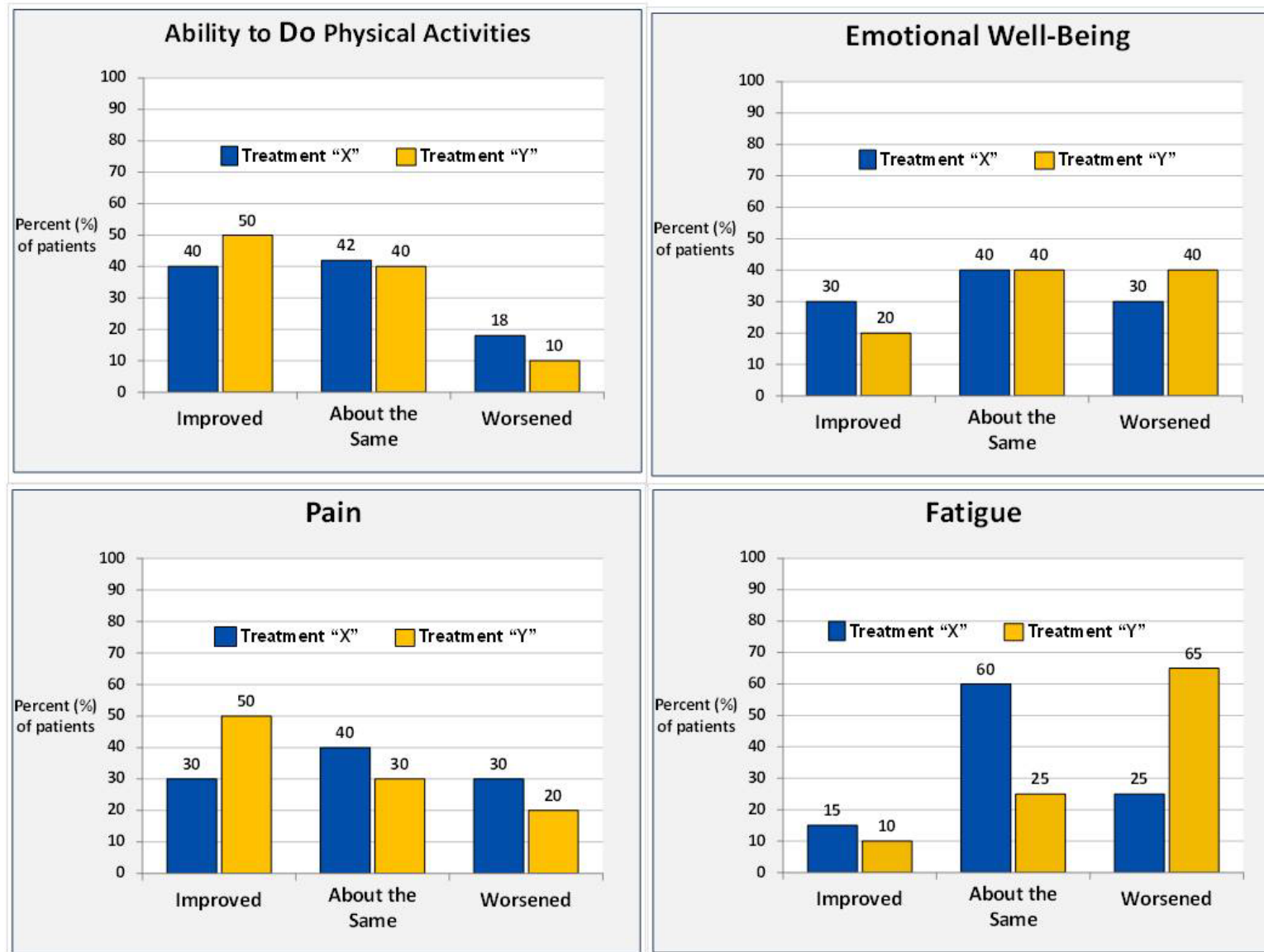


Fatigue



Bars

Status of 100 patients 9 months after starting treatment

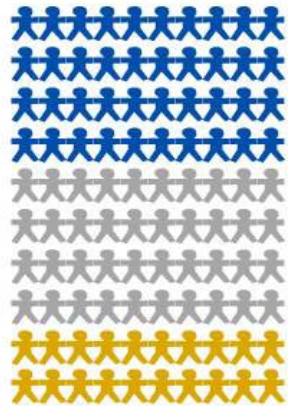


Icons

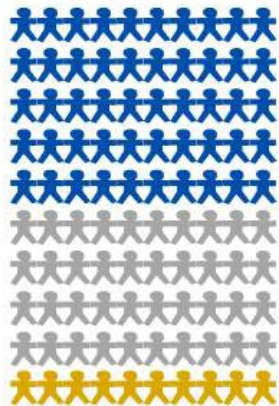
Status of 100 patients 9 months after starting treatment

Ability to Do Physical Activities

Treatment "X"

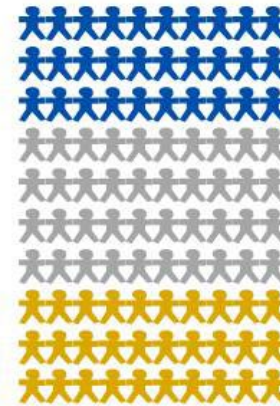


Treatment "Y"

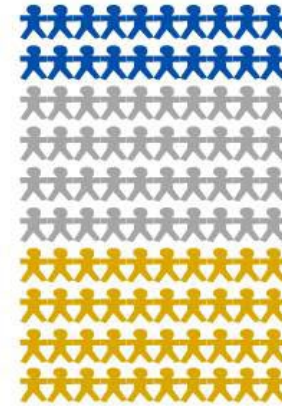


Emotional Well-Being

Treatment "X"



Treatment "Y"



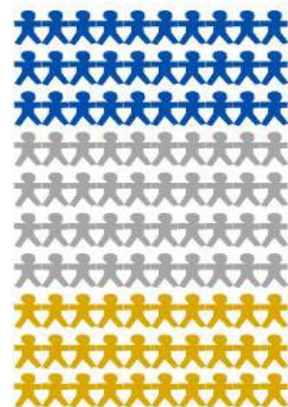
Improved

Stable

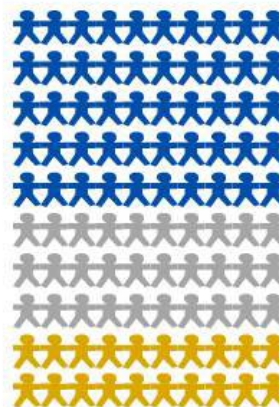
Worsened

Pain

Treatment "X"

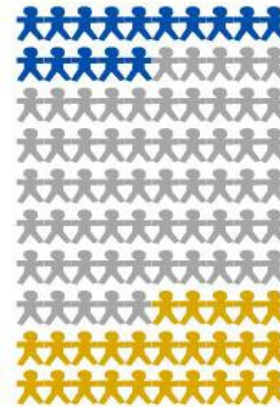


Treatment "Y"



Fatigue

Treatment "X"



Treatment "Y"



Key Findings

- **FOR PATIENTS:** Pie charts
 - More accurately interpreted than bar charts
 - Rated clearer than bar charts or icon arrays
 - Rated best for proportions
- **FOR CLINICIANS**
 - Clarity ratings did not significantly differ between pie charts and bar charts
 - Researchers tended toward picking bar charts as preferred



Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data

Claire Snyder^{1,2,3} · Katherine Smith^{2,3} · Bernhard Holzner⁴ · Yonaira M. Rivera² · Elissa Bantug³ · Michael Brundage⁵ · PRO Data Presentation Delphi Panel

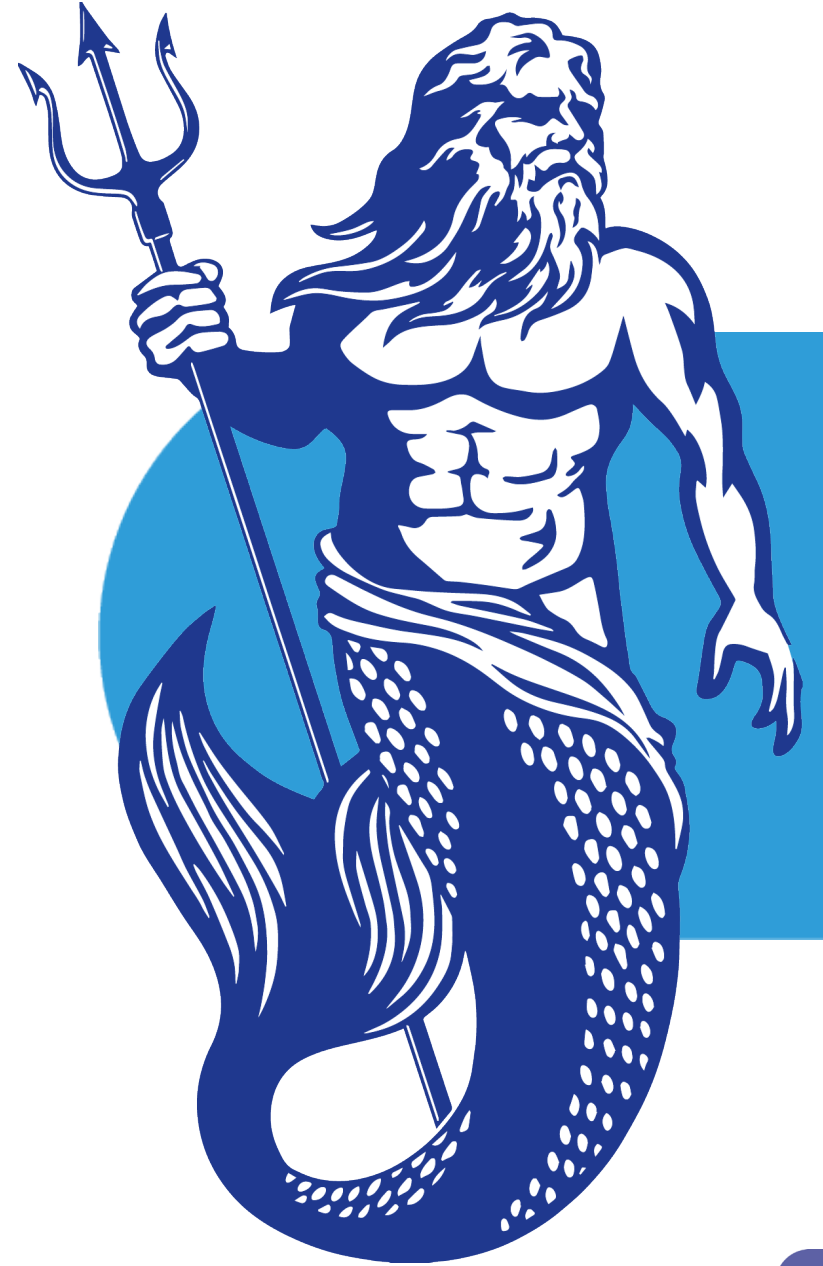
Accepted: 29 September 2018
© Springer Nature Switzerland AG 2018

Proportions changed:

- Pie charts are preferred for patients
- Bar charts, pie charts, or stacked bar charts are reasonable options for clinicians and researchers

Four Principles

1. Know Your Audience
 - Clinicians Like It Simple
 - Patients Like It Even Simpler
2. There Is No Agreement on What “Up” Means
3. Don’t Over-Estimate Statistical Savvy
4. Research Questions Drive Data Display





Helping you navigate the use of patient-reported outcomes (PROs)
in clinical trials and clinical practice



Visit online: TheProteusConsortium.org



Subscribe to the e-newsletter, PRO-cisely PROTEUS:
theproteusconsortium.org/subscribe

Follow on Twitter: [@proteuspros](https://twitter.com/proteuspros)



Connect on LinkedIn: linkedin.com/company/theproteusconsortium

Contact: TheProteusConsortium@gmail.com