

# Methodological Challenges in Analyzing Patient-reported Outcomes

Elizabeth A. Hahn

Center on Outcomes, Research and Education (CORE),  
Evanston Northwestern Healthcare, Evanston, IL

Dept. of Preventive Medicine and Institute for Healthcare  
Studies, Feinberg School of Medicine, Northwestern  
University, Chicago, IL

# Health-related Quality of Life as an Endpoint in a Phase III Clinical Trial

---

- ◆ Extend endpoint evaluation beyond response rate and survival
- ◆ Assess impact of treatment on multiple dimensions of daily life
- ◆ Examine aspects of HRQL that are relevant to clinical decision-making

# Methodological, statistical and interpretive challenges

---

- ◆ cross-cultural comparability
- ◆ longitudinal modeling
- ◆ missing data
- ◆ crossover
- ◆ clinical significance
- ◆ clinical interpretability

# Chronic Myeloid Leukemia (CML)

- ◆ Proliferative disorder of hematopoietic stem cells
- ◆ Linked to a single molecular abnormality (Philadelphia (Ph) chromosome)
- ◆ Well-characterized clinical course

# Therapeutic Options for CML

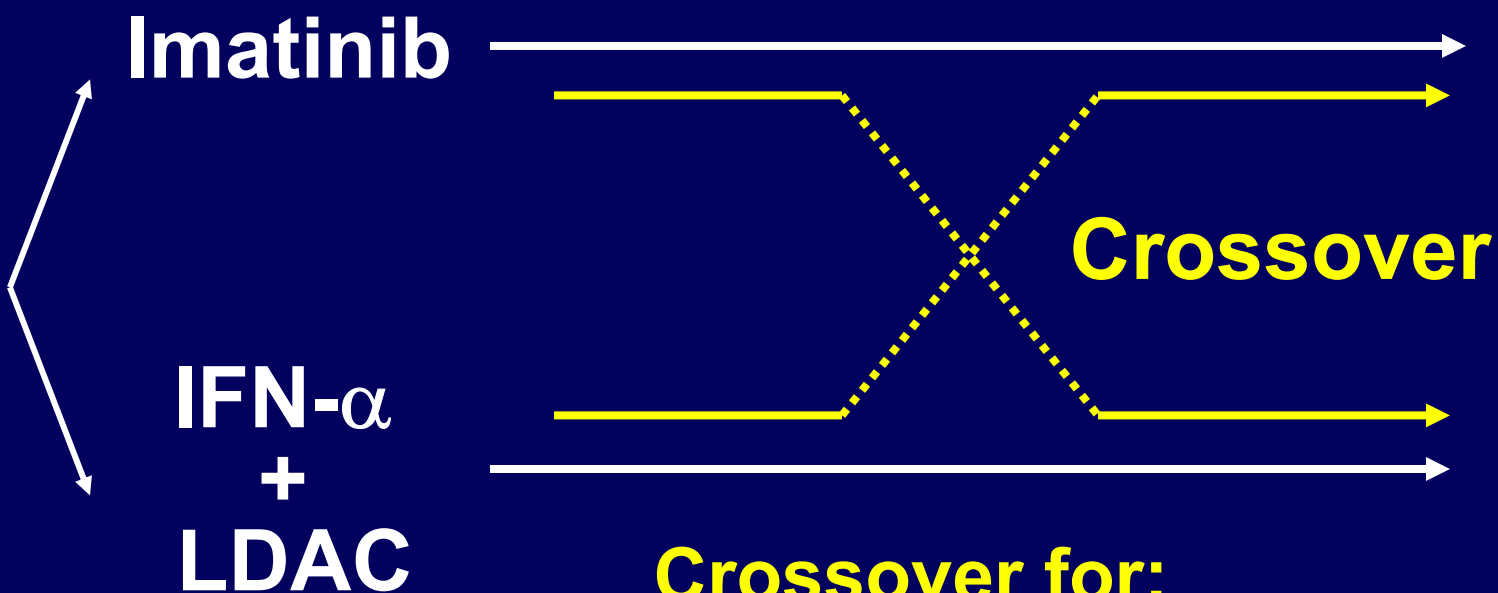
---

- ◆ **Allogeneic stem cell transplantation (SCT)**
- ◆ **Chemotherapy with hydroxyurea, busulfan**
- ◆ **IFN- $\alpha$ -based treatments**
- ◆ **Imatinib (targeted therapy)**

# IRIS Trial Design

(Phase III, Multicenter, Open Label)

R  
A  
N  
D  
O  
M  
I  
Z  
E



## Crossover for:

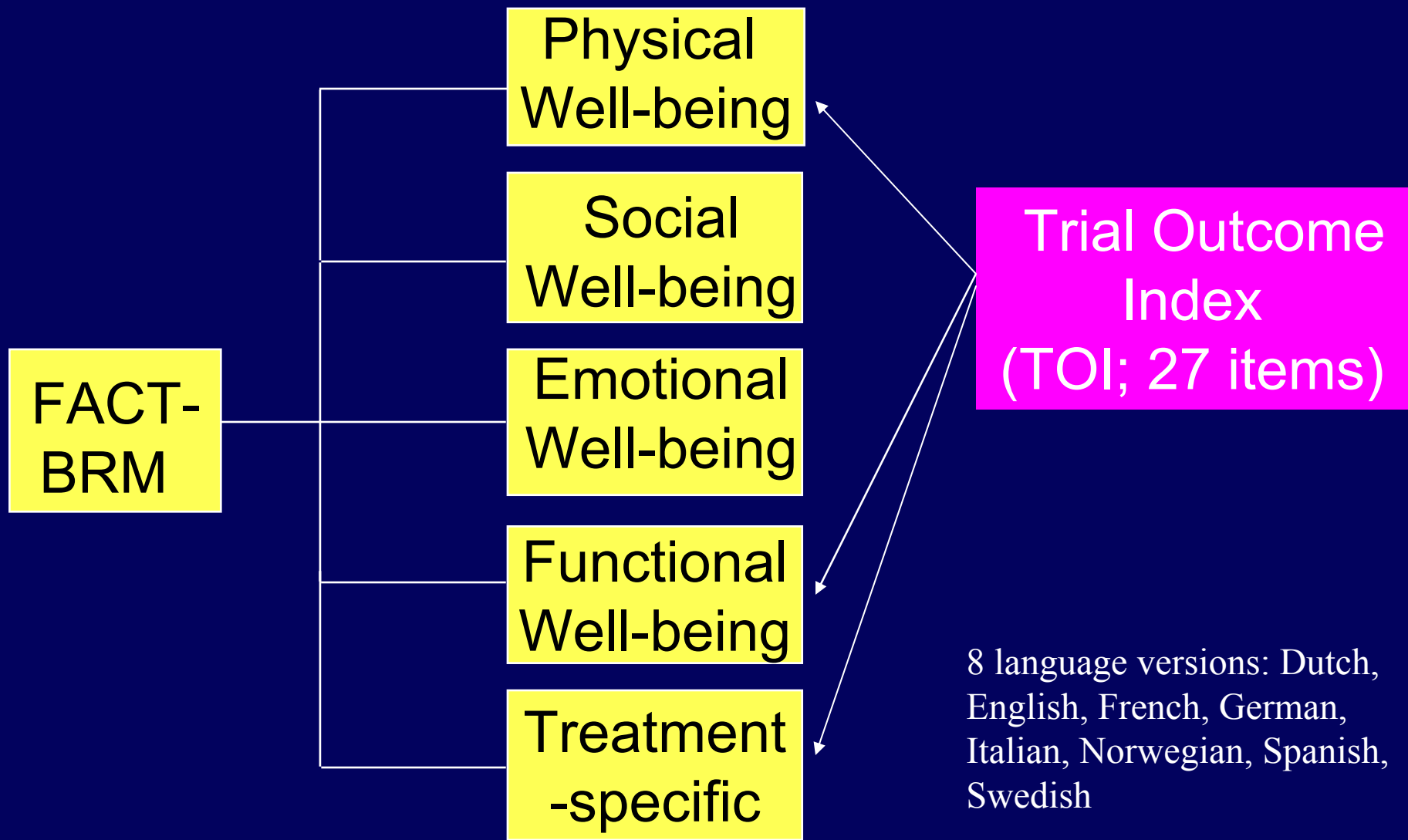
- Lack of response
- Loss of response
- Intolerance of treatment

# Study Endpoints

---

- ◆ **Primary Endpoint: Time to Progression**
- ◆ **Secondary Endpoints:**
  - Rate/duration of complete hematologic response & major cytogenetic response, safety, tolerability, molecular remission, pharmacogenomics, pharmacokinetics
  - **Health-related quality of Life** (measured at Baseline and Months 1-6, 9, 12, 18, 24)

# Functional Assessment of Cancer Therapy- Biologic Response Modifiers (FACT-BRM)



# FACT-BRM Sample Questions

---

	Not at all	A little bit	Some -what	Quite a bit	Very much
I have a lack of energy	0	1	2	3	4
I am forced to spend time in bed	0	1	2	3	4
I am able to work (include work in home)	0	1	2	3	4
I am content with the quality of my life right now	0	1	2	3	4
I get tired easily	0	1	2	3	4
I have pain in my joints	0	1	2	3	4
I get depressed easily	0	1	2	3	4

# FACT-BRM Status at Month 24

---

	Randomised Treatment Group	
	Imatinib ( <i>n</i> =530)	INF+LDAC ( <i>n</i> =519)
Dropout	67 (13%)	185 (36%)
Incomplete	132 (25%)	84 (16%)
Complete	331 (62%)	250 (48%)

---

Chi-square=77.11, 2 d.f.,  $p < 0.001$

# Methodological Issue

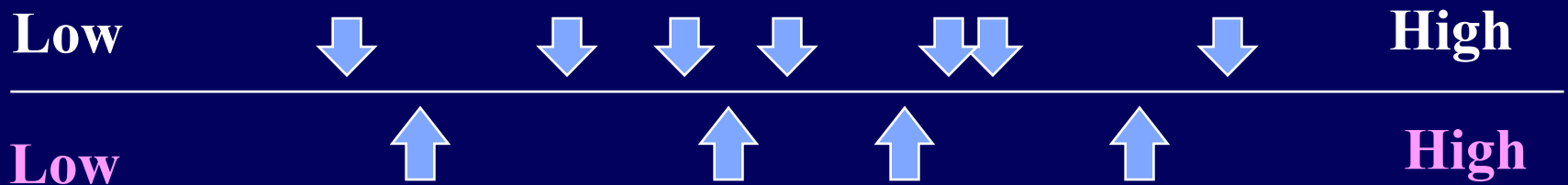
cross-cultural  
comparability

# Strategy

Rasch model

# Rasch Measurement Model

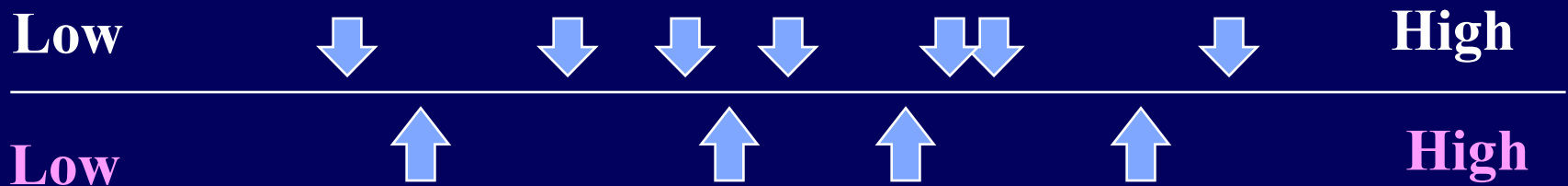
## Item Location



## Person HRQL

# Rasch Measurement Model

Item Location,  
English



Item Location,  
French

**Table 2.** Results of differential item functioning (DIF) analyses

Item abbreviation (27-item Trial Outcome Index; TOI)	20-item TOI	Nine-item TOI	English versus German		English versus French	
			Uniform DIF	Non-uniform DIF	Uniform DIF	Non-uniform DIF
			$t^a$	$\chi^{2b}$	$t^a$	$\chi^{2b}$
Lack of energy			4.13***	5.88	3.09**	18.63***
Nausea	√		2.98**	2.78	0.24	4.06
Trouble meeting family needs	√	√	0.07	2.99	- 0.72	10.24
Pain	√	√	- 0.49	8.50	1.23	2.28
Side effects of treatment	√	√	0.48	6.76	- 0.36	7.60
Feel ill	√		0.19	4.78	- 2.77**	10.78
Spend time in bed	√		2.43	1.38	2.86**	0.44
Able to work	√		- 5.16***	7.12	- 2.62**	14.46**
Work is fulfilling	√		- 3.37***	5.13	0.87	2.81
Enjoy life	√		- 1.60	7.81	- 5.05***	18.46**
Accepted illness	√		- 6.36***	5.07	- 7.87***	9.83
Sleeping well	√	√	- 1.42	6.04	1.60	0.65
Enjoying things	√	√	- 0.05	7.21	- 0.58	11.47
Content with quality of life	√		- 3.55***	27.60***	- 3.50***	9.68
Tired	√		3.25**	4.46	1.56	10.58
Weak	√		- 2.12	21.42***	- 3.14**	0.93
Good appetite	√		- 3.83***	4.34	- 0.88	7.53
Pain in joints	√	√	0.71	3.57	2.07	4.75
Chills	√		2.69**	2.72	5.97***	0.76
Fevers	√		2.46	2.94	4.13***	2.33
Sweating	√		1.31	16.85**	1.53	0.83
Trouble concentrating	√	√	- 0.56	9.98	- 1.40	4.09
Trouble remembering things	√	√	1.20	2.56	0.37	1.73
Depressed	√		2.91**	2.49	0.28	1.91
Get annoyed easily	√		1.03	6.27	2.93**	0.90
Emotional ups and downs	√	√	1.29	4.08	1.04	7.94
Feel motivated to do things	√		3.43***	6.63	1.48	17.67**

<sup>a</sup> $t$  = independent sample test of significant differences (negative values indicate lower item calibrations for the English reference group compared to the German or French focal group).

<sup>b</sup> $\chi^2$  = difference in nested chi-square values, four degrees of freedom.

\*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

√ = item met criterion for a 20-item or nine-item TOI subset (see text).

## Methodological Issue

## Strategy

longitudinal modeling

mixed-effects  
growth curve  
model

missing data

pattern-mixture  
sensitivity analyses

crossover

time-dep. covariate

# Statistical Analyses

---

## ◆ Mixed effects model:

$$Y_{ij} = \underbrace{X_{ij}\beta_j}_{\text{fixed effects}} + \underbrace{Z_{ij}d_{ij}}_{\text{random effects}} + \underbrace{e_{ij}}_{\text{residual error}}$$

Polynomial growth curve model to account for nonlinear change over time

- Fixed effects:  $\text{trt}$ ,  $\text{trt} \cdot \text{time}$ ,  $\text{trt} \cdot \text{time}^2$ , ...  $\text{trt} \cdot \text{time}^6$
- Random effects: patient-specific intercept and slope ( $\text{time}$ ,  $\text{time}^2$ )

# Statistical Analyses

---

## ◆ Nonignorable missing data

Pattern-mixture model to account for missing not at random (MNAR) data (Little, 1995):

- *stratify* patients into two groups (completers vs. dropouts)
- create *mixed effects model* in each stratum
- combine parameter estimates into a *weighted average*

# Pattern-mixture model (Little, 1995)

---

*Formula to calculate estimates of the population parameters ( $\tilde{\beta}_j$ ):*

$$\tilde{\beta}_j = \sum_{k=1}^K \pi_j^{(k)} \hat{\beta}_j^{(k)}$$

*where weights ( $\pi_j^{(k)}$ ) are estimated as the proportion of cases in each treatment group ( $j$ ) with the  $k$ th pattern, and  $\hat{\beta}_j^{(k)}$  are the estimates obtained within each of the  $k$  patterns*

# Statistical Analyses

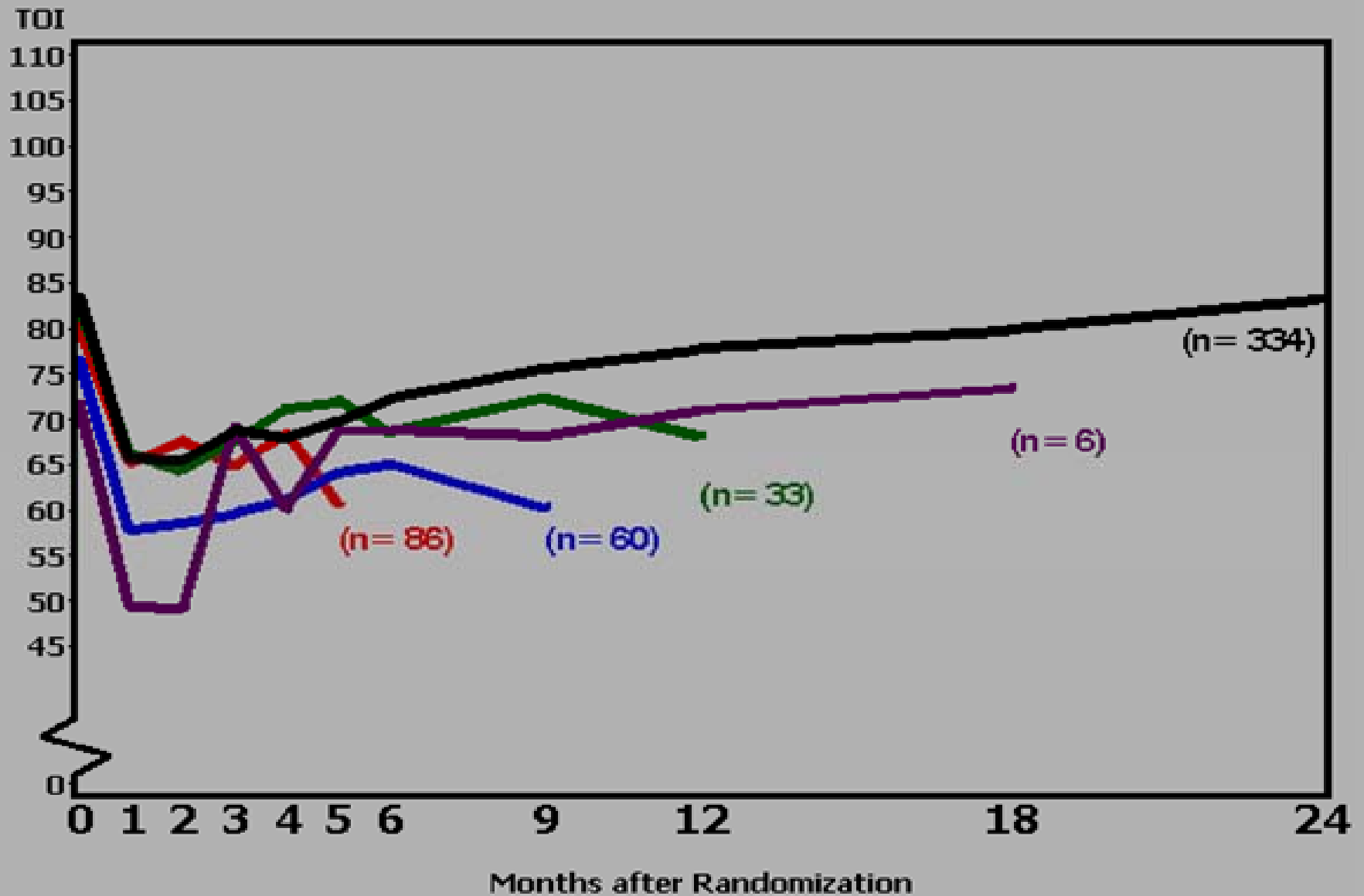
---

## ◆ Sensitivity analyses

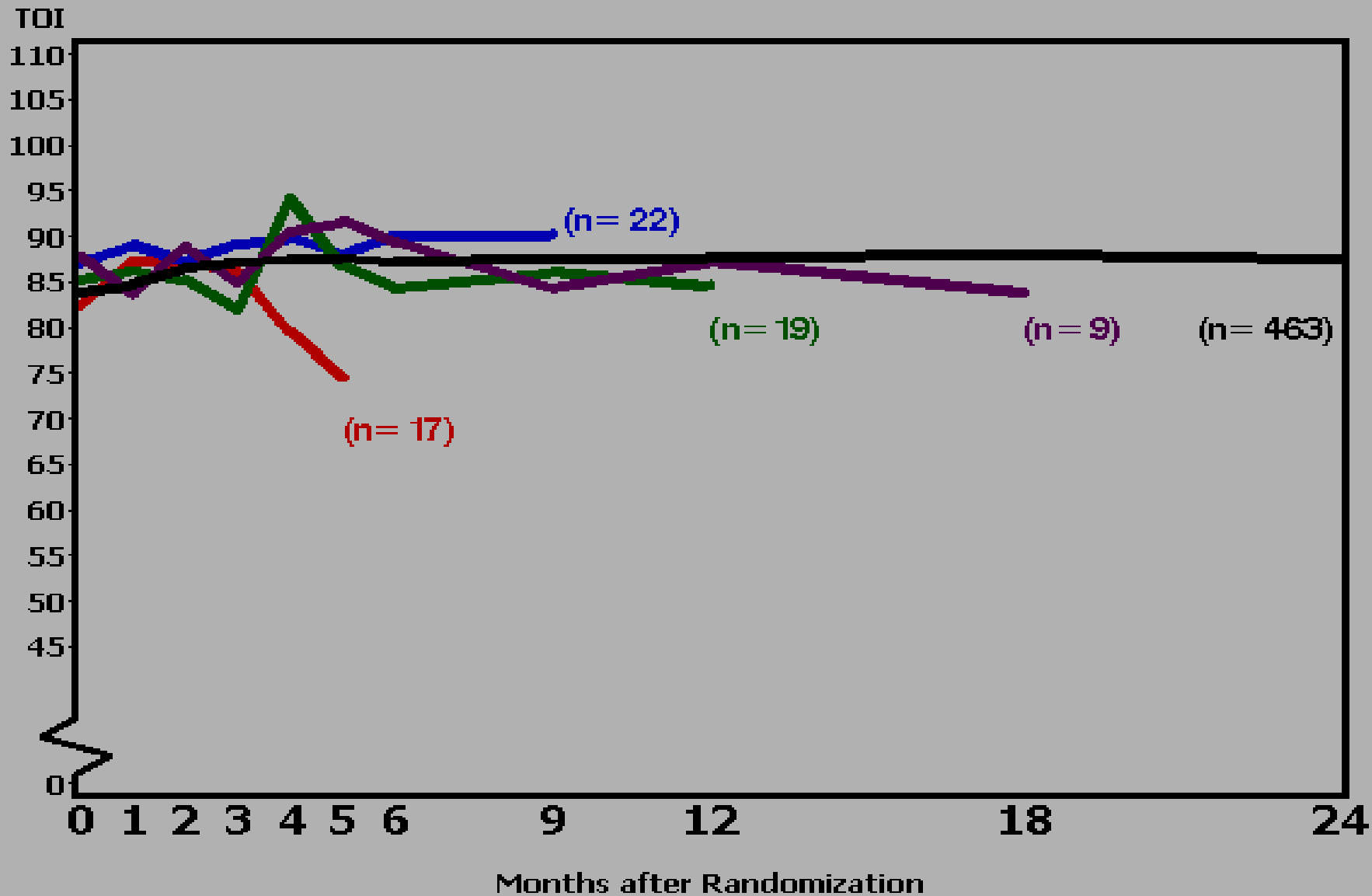
- completers vs. dropouts/incompletes
- completers/incompletes vs. dropouts
- all available data
- average score during treatment
- categories of clinically important change
- drop items with measurement bias

**Longitudinal Analyses of the  
Primary QOL Endpoint:  
Trial Outcome Index (TOI)**

# IFN+LDAC ( $n=519$ )



# Imatinib ( $n=530$ )



# Estimated Mean TOI (Intention-to-treat)

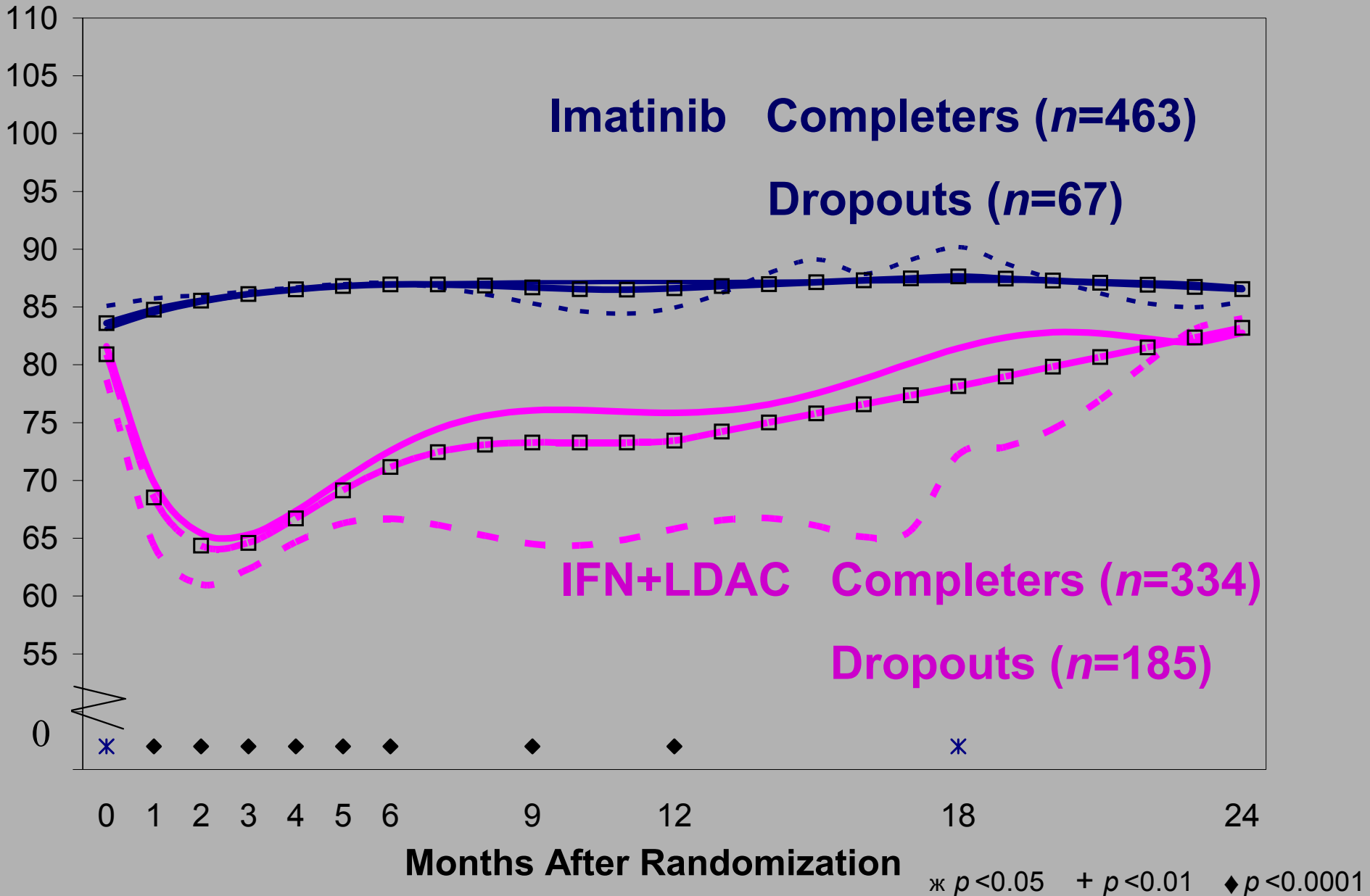
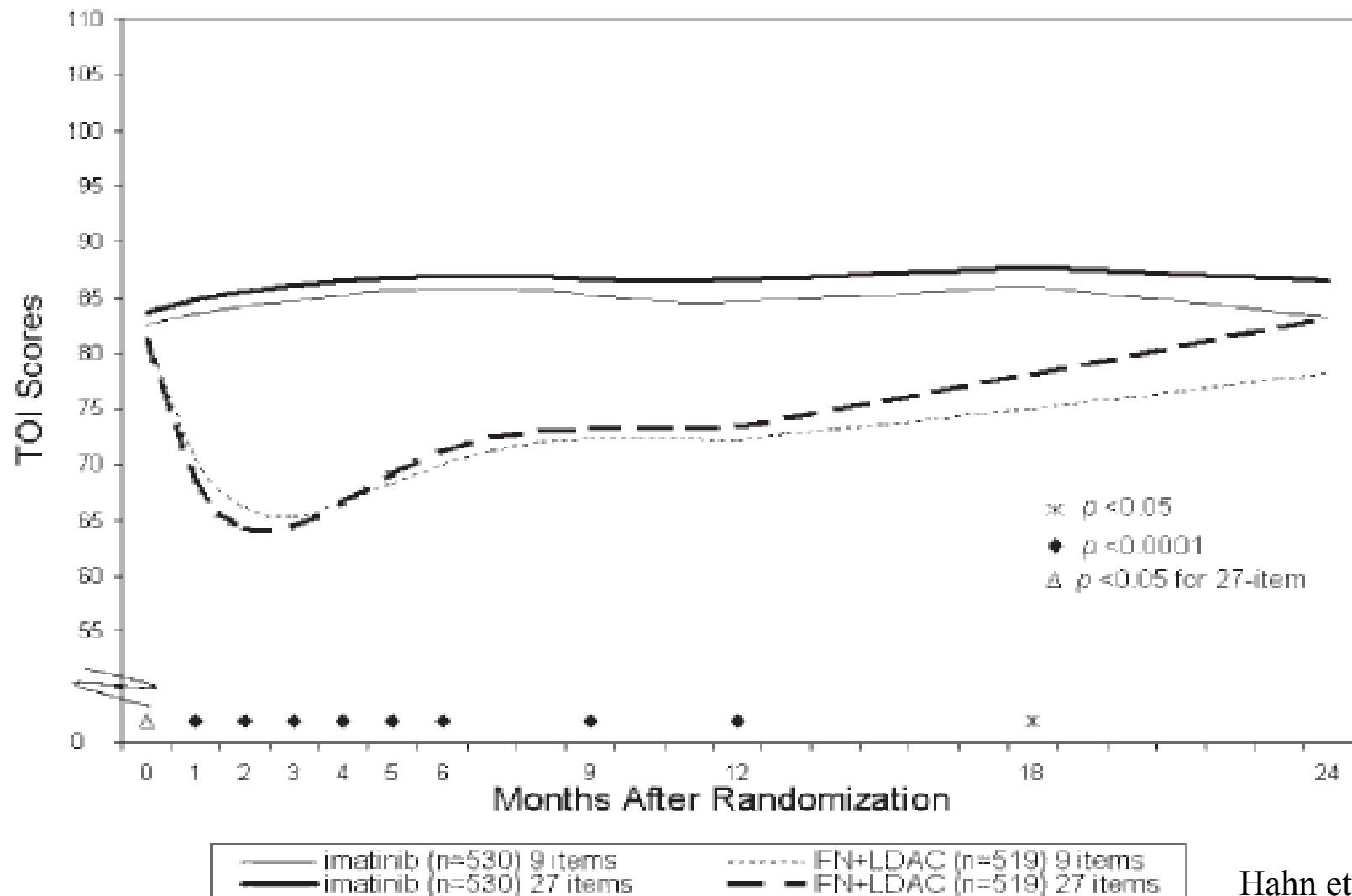
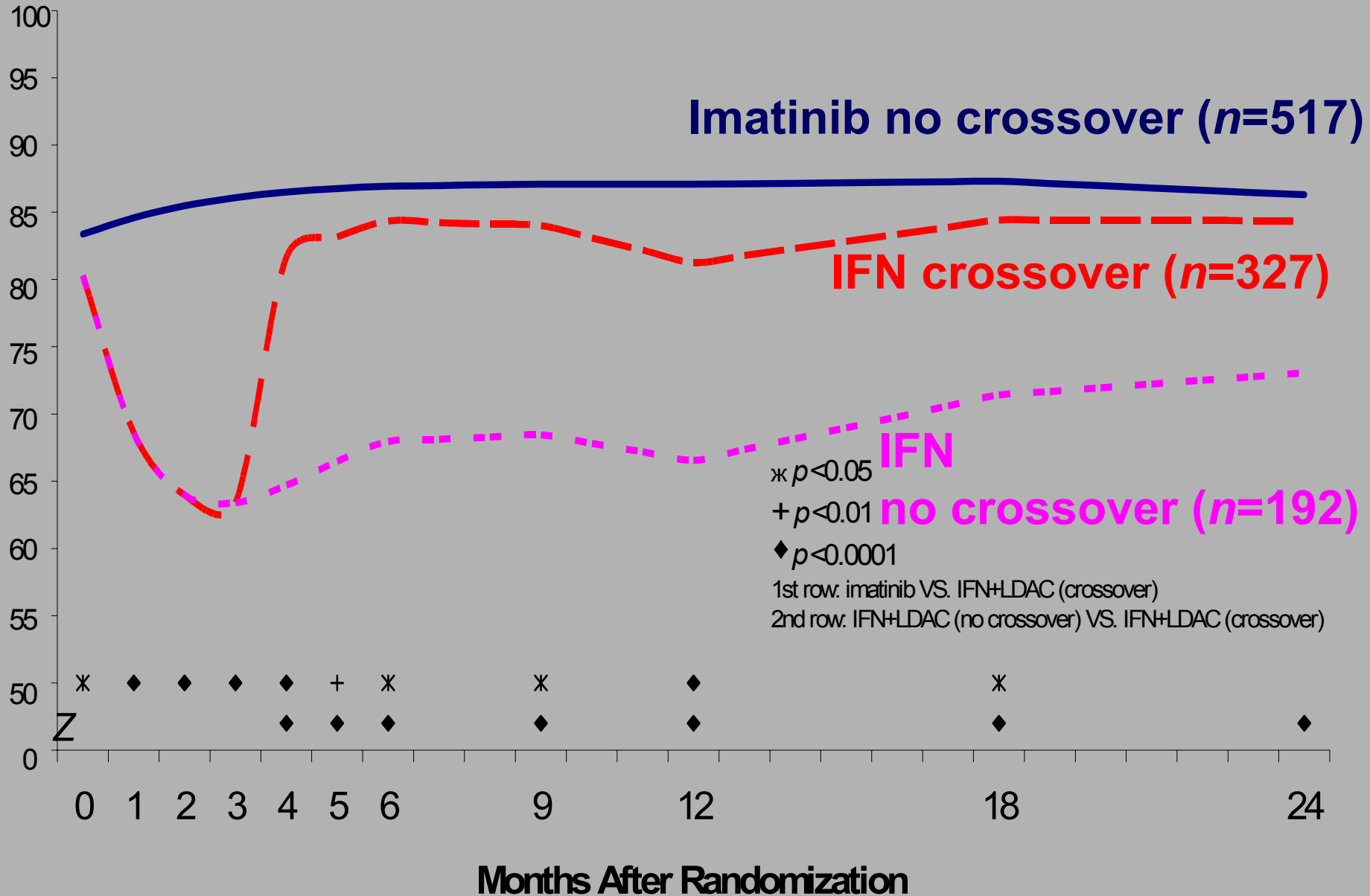


Figure 1 Estimated mean Trial Outcome Index scores by treatment arm, adjusted for missing data (intention-to-treat approach). *P*-values are for difference in treatment arm means at each scheduled administration of the FACT-BRM.



# Estimated Mean TOI (Adjusted for crossover)



## Methodological Issue

## Strategy

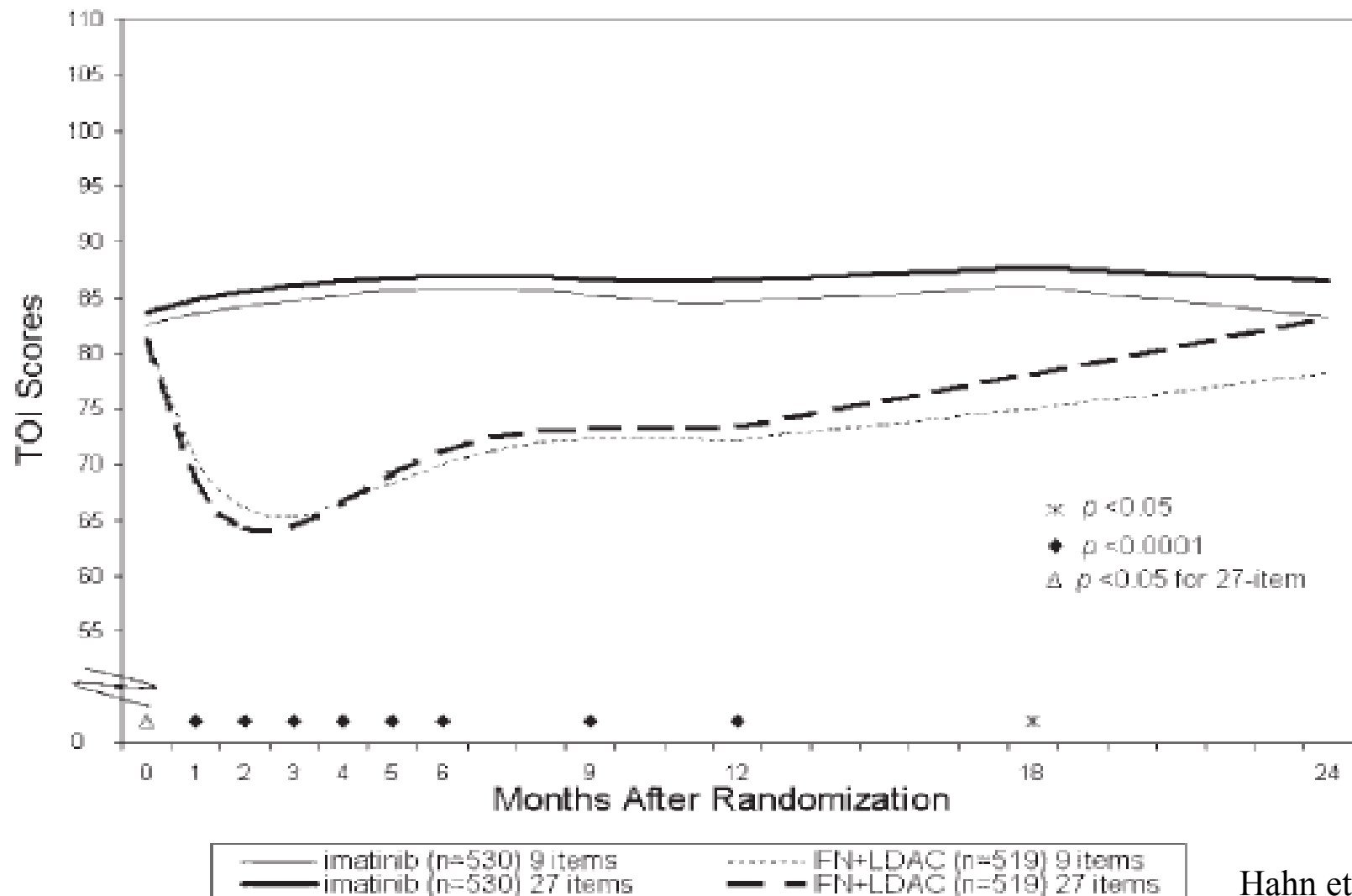
clinical significance

anchor- and  
distribution-based

### ◆ “minimally important difference” (MID):

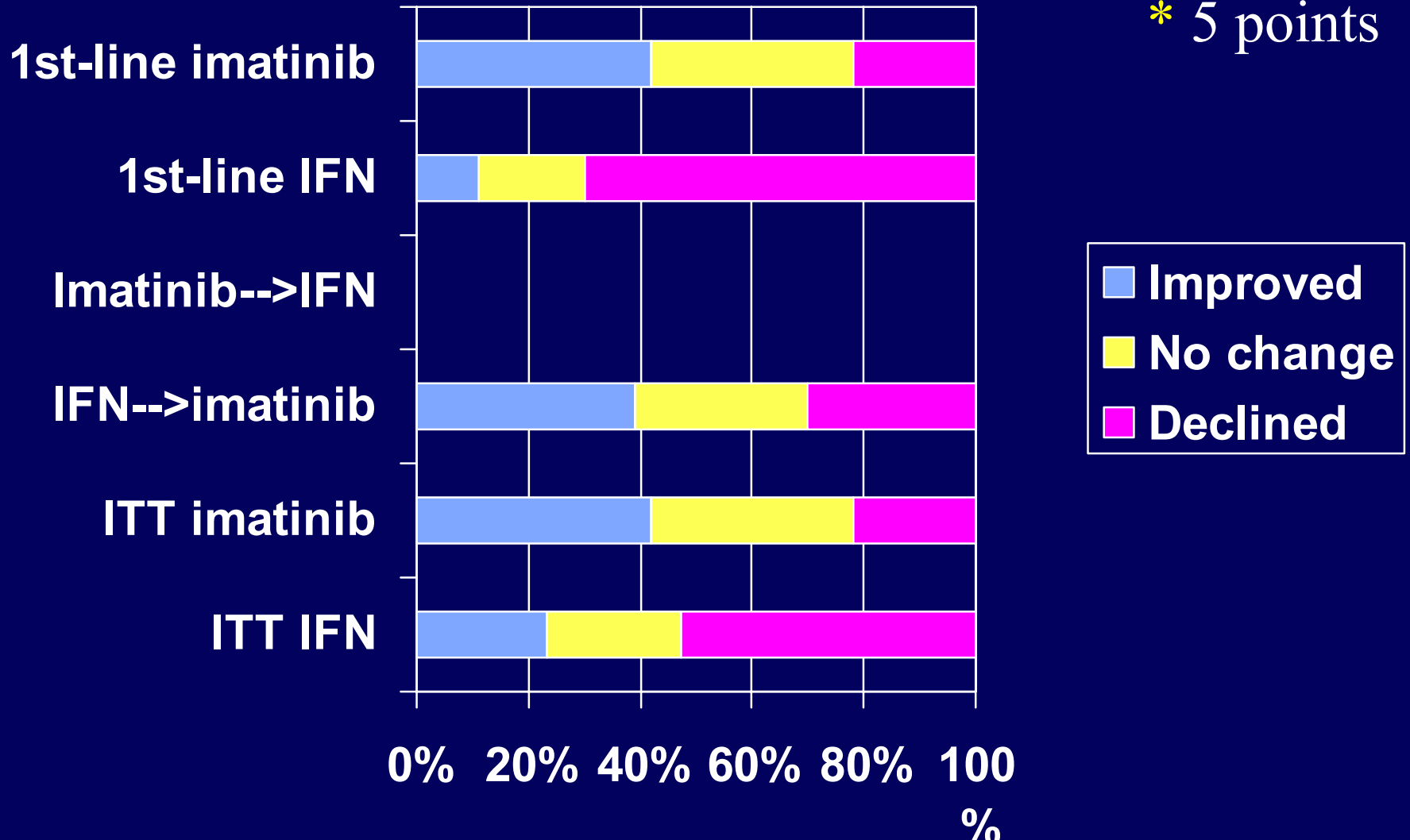
“the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management” (Jaeschke, Singer & Guyatt, 1989, p.408)

Figure 1 Estimated mean Trial Outcome Index scores by treatment arm, adjusted for missing data (intention-to-treat approach). *P*-values are for difference in treatment arm means at each scheduled administration of the FACT-BRM.



# Meaningful TOI Change\* from Baseline to Month 12 (n=736)

\* 5 points



## Methodological Issue

clinical interpretability

## Strategy

Rasch model

modified forest  
plot

**Not at all**      **A little bit**      **Some -what**      **Quite a bit**

■=imatinib    ○=IFN+Ara-C    ◼=both

○      ■      **Motivated to do things**  
 ○      ■      **Contented, enjoy life**  
 ○      ■      **Accept my illness**  
 ○      ◼      **Meet family needs**  
 ○      ■      **Sleep, work, appetite**

**Better daily functioning & well-being on imatinib**

■      ○      **Lack of energy, weak, tired**

**Less fatigue on imatinib**

■      ○      **Depressed, ups & downs**

■      ○      **Easily annoyed**  
 ■      ○      **Trouble remembering & concentrating**

**Milder emotional/ cognitive complaints on imatinib**

■      ○      **Fevers, chills, nausea**

■      ○      **Bedrest, side effects**

◼      **Pain, sweating, feel ill**

**Fewer side effects on imatinib**

# Summary and Conclusions

# Rasch Measurement Model

---

- ◆ Useful for evaluating cross-cultural measurement equivalence
  - no evidence of systematic measurement bias, thus permitting pooling of international clinical trial data
- ◆ Provides innovative approach to interpreting meaning of HRQL outcomes

# Mixed-effects & growth curve models

---

- ◆ unbiased estimates if missing data are MCAR / MAR
- ◆ model the covariance structure
- ◆ account for nonlinear time trends
- ◆ time-dependent covariates
- ◆ standard software (SAS)

# Pattern-mixture model

---

## ◆ Advantages:

- useful if data are MNAR
- descriptive info. about treatment effects
- use standard software (SAS)

## ◆ Challenges:

- how to create the strata? (duration of follow-up, number of treatment cycles completed, disease progression versus no progression)
- underidentified models (extrapolation or restrictions on parameters required)

# Clinical Significance

---

- ◆ focus on clinical meaningfulness vs. statistical significance testing
- ◆ enhances provider understanding of HRQL results

# Acknowledgements

---

- ◆ The IRIS Study and statistical analyses were supported by Novartis Pharma
- ◆ Additional statistical analysis support provided by Hongyan Du (CORE)
- ◆ Hahn et al., *J Clin Oncol* 21:2138-2146, 2003
- ◆ Hahn et al., *Clinical Trials* 3:280-290, 2006