

Methodik für die Auswertung von Lebensqualitätsdaten im Rahmen der Nutzenbewertung von Arzneimitteln

Friedhelm Leverkus

Director Health Technology Assessment &
Outcomes Research

Pfizer Deutschland GmbH

Köln, 29. November 2013



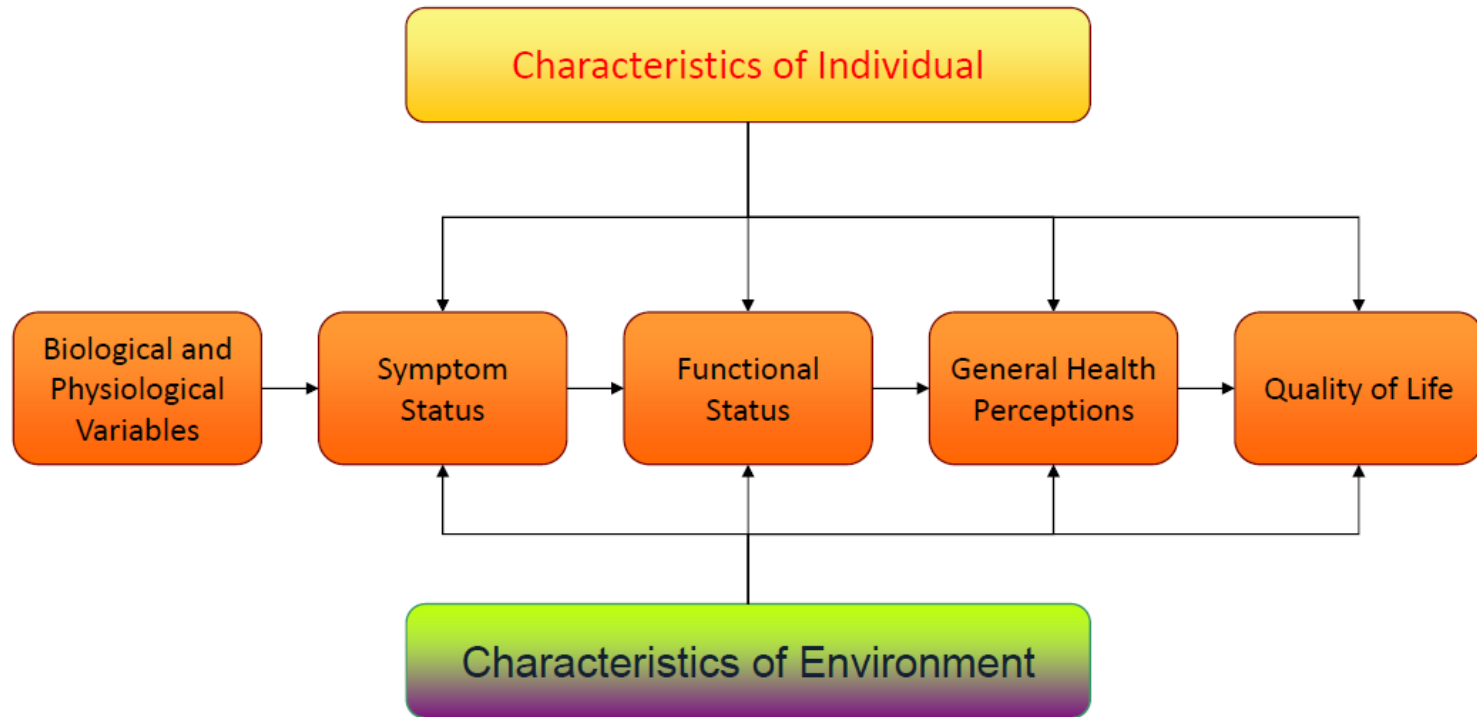
Inhalt

- Was ist Lebensqualität überhaupt – was ist es nicht ?
- Lebensqualität als primärer Endpunkt?
- Auswertung am Beispiel A12-15
 - Im Protokoll geplante Auswertung
 - Auswertung für die Nutzenbewertung

FDA (Food and Drug Administration) PRO (Patient Reported Outcome) Guideline 2009

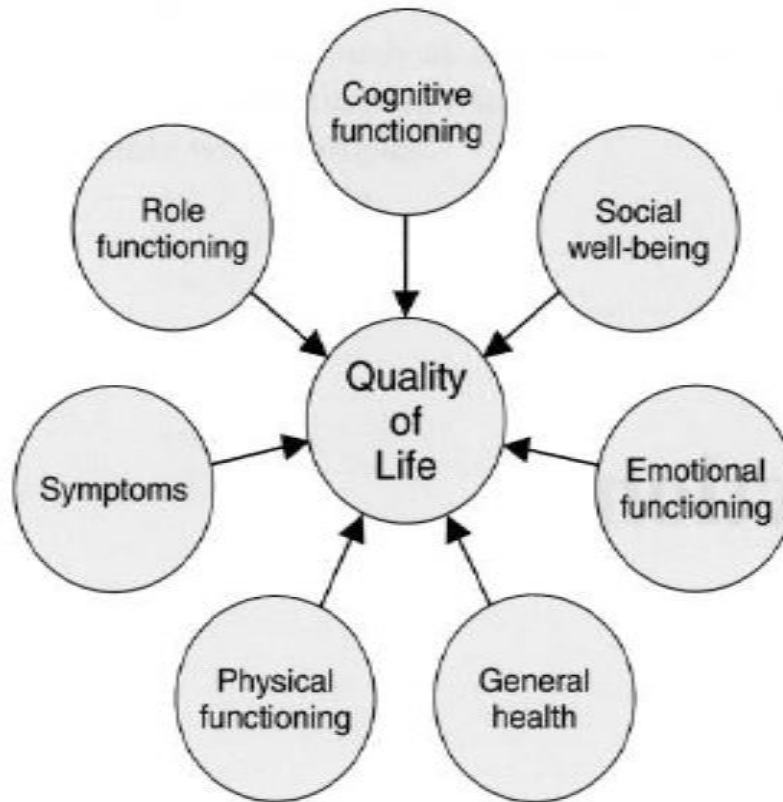
- „Directly from patients response without interpretation of clinician or elsewhere“
- „Sign, symptoms or on aspects of functioning directly related to disease status“
- Quality of Life - because this term implies the evaluation of non health related aspects of life..., it is too general and undefined

MODEL OF HEALTH OUTCOMES



Wilson & Cleary *JAMA* (1995)

Beispiele für Quality of Life Bereiche



Walters 2009

Figure 1.1 Examples of QoL domains.

QALY- Quality Adjusted Life Year sind keine HRQoL Messungen (Health Related Quality of Life)

Ziel ist der Nachbau der Nutzenfunktion
aus der klassischen ökonomischen Theorie
Funktionsmessung nur Mittel zum
Zweck



SP5A-CT-2007-044172

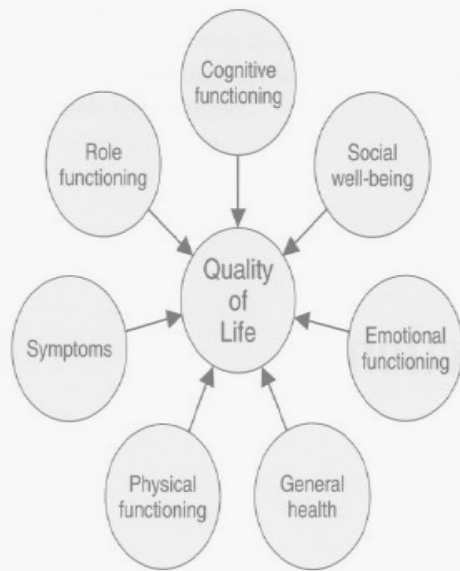
European Value of a Quality Adjusted Life Year

Instrument: Specific Targeted Research Project

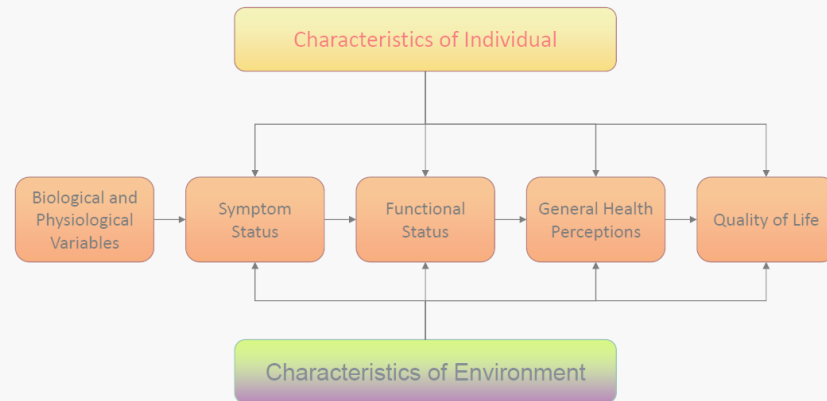
Final Publishable Report

EU Kommission gesponsertes Projekt mit
6 Universitäten und Instituten
1.300 Befragte

Gerard Duru, Emeritus Research Director in Mathematics at the French National Centre of Scientific research (CNRS) said: 'The underlying assumptions of the QALY outcome are very theoretical and are not verified in a real population. The QALY indicator is not a valid scientific scale. It is impossible to know what we are measuring, and therefore impossible to base a formula upon it. To be able to

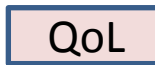
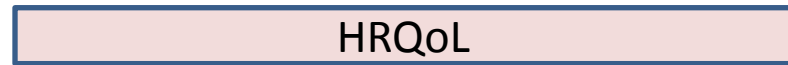


MODEL OF HEALTH OUTCOMES



Wilson & Cleary JAMA (1995)

Walters 2009 Examples of QoL domains.



AMNOG



Primärer Endpunkt

Figure 1. Endpoint Model: Treatment of Disease X

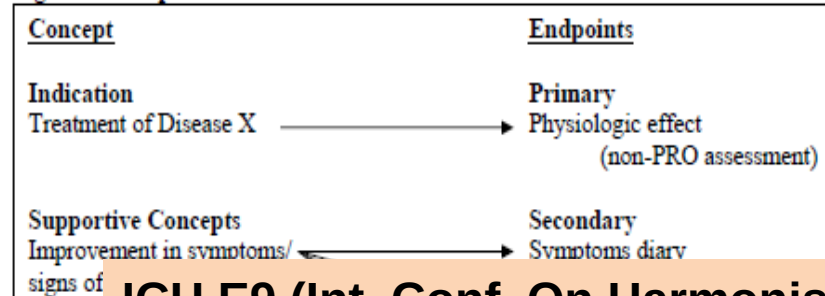
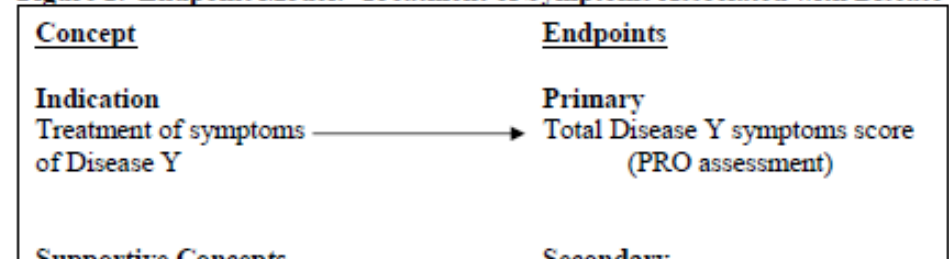


Figure 2. Endpoint Model: Treatment of Symptoms Associated with Disease Y



ICH E9 (Int. Conf. On Harmonisation) zu primären und sekundären Endpunkten

- The primary variable ... should be the variable the most ***clinically relevant and convincing evidence***. There should generally be only **one primary** variable.
- The primary variable should generally be the one used when estimating **the sample size**
- Secondary variables are either **supportive** measurements related to the primary objective or measurements of effects related to the secondary objectives.

Case Study

- Sie sind im Planungsteam einer klinischen Prüfung in einer onkologischen Indikation
- Ziel: aussagekräftige Studie, die zur Zulassung führt und von der Clinical Community anerkannt wird
(Wir befinden uns in der Zeit vor dem AMNOG Arzneimittelneuordnungsgesetz)
- Primärer Endpunkt PFS (progression free survival) oder HRQoL (Health Related Quality of Life)?

Regulatorische Überlegungen

- **Guideline on the evaluation of anti cancer medicinal products in man 2013**
 - **7.1.5. Endpoints**
 - Acceptable primary endpoints **include cure rate, OS** (overall survival) **and PFS/DFS** (progression free survival /disease free survival)
 - In patients with tumor-related symptoms at base line, **symptom control**, if related to anti-tumor effects, is a valid measure of therapeutic activity and may serve as primary endpoint in **late line therapy studies**, provided that sources of possible bias can be minimized.
 - **7.1.5.1. Secondary endpoints and exploratory analyses**
 - Irrespective of the choice of primary endpoint OS or PFS, ORR (overall response rate) and rate of tumor stabilisation for, e.g. 3 or 6 months should be reported. Especially in the palliative setting, **HRQoL/PRO** using generally accepted instruments might be informative

Medical Community

- Interesse an klinischen Daten in der Regel bei Ärzten höher als an PRO
- Lebensqualitätsmessungen keine ärztliche Routine
- PFS ist ein „generell anerkannter Endpunkt “

Methodische Probleme

Fehlende Werte (Fairclough)

TABLE 1.9 Study 3: Documentation of FACT-Lung assessments.

Status/ Reason Missing	Baseline N (%)	6 weeks N (%)	12 weeks N (%)	6 months N (%)
Completed[†]	513 (98)	362 (75)	276 (63)	183 (52)
Refusal	0 (0)	34 (7)	46 (10)	62 (17)
Patient feels too ill	0 (0)	3 (1)	3 (1)	3 (1)
Staff felt patient too ill	0 (0)	18 (4)	15 (3)	11 (3)
Staff oversight	7 (1)	35 (7)	59 (13)	56 (15)
Other	1 (0)	13 (2)	23 (5)	20 (6)
Unknown	4 (1)	19 (4)	19 (4)	20 (6)
Total expected	525	484	441	355
Patients expired [‡]	0 (0)	41 (8)	84 (16)	170 (32)
Total patients	525	525	525	525

[†] Percent of expected assessments (e.g. patient still alive).

[‡] Percent of all possible assessments.

Methodische Probleme

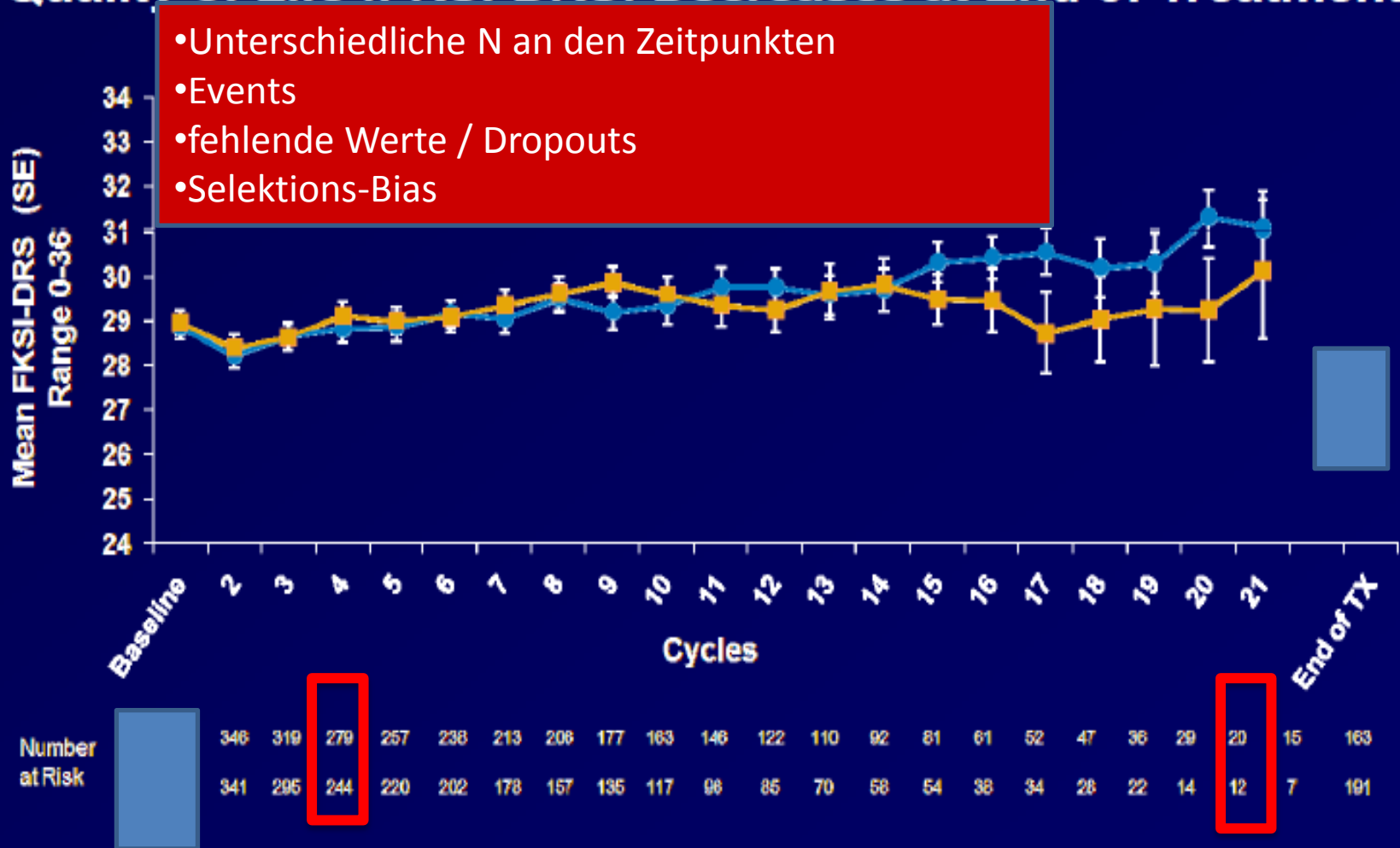
- Methodische Probleme
 - PFS eindimensional - PRO mehrdimensional-> Multiplizität
 - Verlängerung der PFS in Monaten besser interpretierbar als Skalenwert
 - Der erwartete Unterschied in PFS besser antizierbar als für PRO-> Fallzahl unsicher
 - Individuelle Variation erhöht die Varianz > Fallzahl höher
- Verblindung
 - PFS objektive Kriterien, Raterblinding => Verblindung nach Design nicht unbedingt notwendig
 - PRO subjektive Kriterien, unverblindet Biasmöglichkeit. Verblindung im Design aber nicht immer möglich.

AMNOG Herausforderung andere Blick auf Studien

- Unterscheidung in primäre und sekundäre Endpunkte spielt keine Rolle
- Multiplizität spielt keine Rolle
- HRQoL ist patientenrelevanter Endpunkt - von großer Bedeutung für den Zusatznutzen

Auswertung

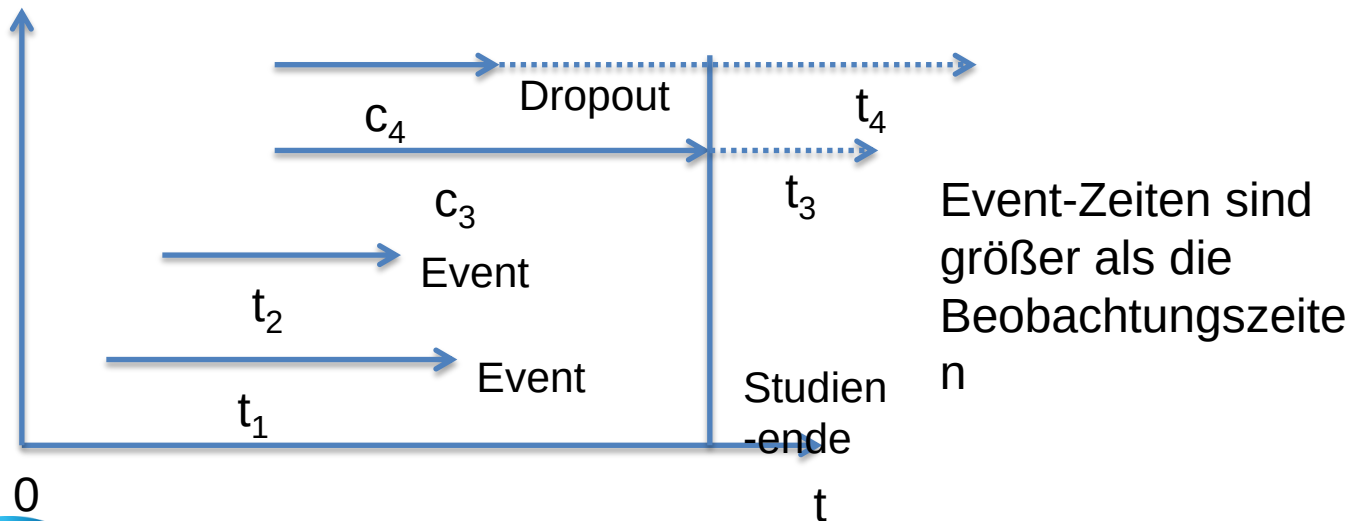
Quality of Life (FKSI-DRS) Decreases at End of Treatment



Fehlende Werte- Verzerrungspotential

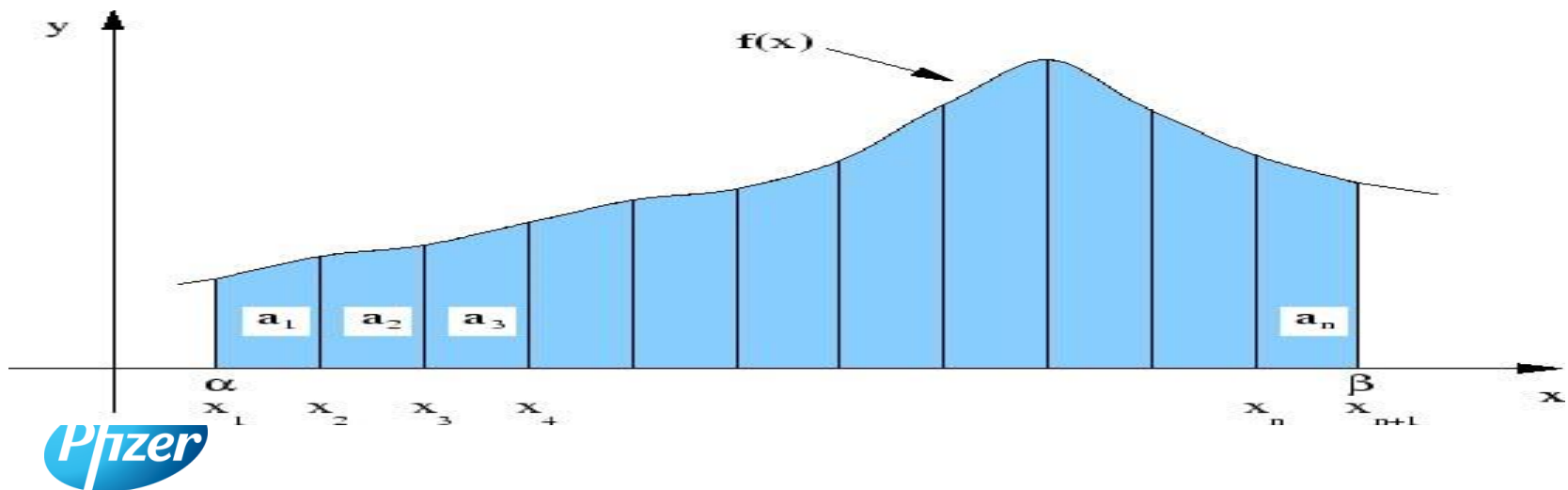
- Höhe des Anteils von den Patienten sind fehlende
- Grosse Unterschiede zwischen den Gruppen
- Grund für Dropout spielt eine Rolle
 - Abbruch wg. unerwünschte Ereignisse hängt sicher mit der LQ
 - =>informatives Zensierung
- Time to event Daten- führen zu unterschiedlichen N

Time to Event Daten - Zensierung



Time to event

- Variable Beobachtungszeiten – Time to event
 - Zielparameter OS/PFS bestimmt die Beobachtungsdauer
 - Overall Survival – Fortschreiben/ Ersetzung von HRQoL nach Ableben macht keinen Sinn Ersetzung von PFS fraglich
- AUC- Area under the Curve für jeden Patienten
 - Doppelgewichtung der PFS/OS
 - Nutzenbewertung nicht abbildbar- PFS wird als nicht patientenrelevant angesehen
 - Multiplikativer Composite Endpoint
 - Implizite Wertung – 1 Monat mit HRQoL 5 = 5 Monate mit Score 1

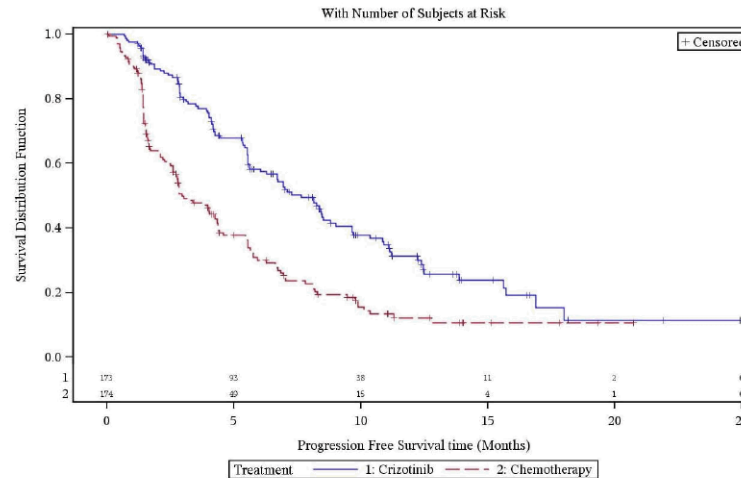


Nutzenbewertung A12-51- Profile 1007

- ALK + NSCLS Patienten (Onkologie-Patienten)
- RCT Crizotinib (N=172) vs Chemotherapie (n=171)
- PRO-auswertbar
 - Basline HRQoL und mindestens ein Wert unter Therapie
 - Crizotinib (n=162) - Chemotherapie (n=151)
- Messung der Lebensqualität
 - QLQ-C30 , QLQC13 , (EQ5D)
 - Messung bis Progression , Abbruch der Studie
 - Tag 1 in jedem Zyklus vor jeder Studie/Medizinischen Untersuchung
- Umgang mit Fehlenden Werten
 - **Messung bei Studienabbruch/Studienende => kein informatives Censoring**

Kleine Tabelle aus den Daten

Figure 1. Kaplan-Meier Plot of Progression-Free Survival Based on the Independent Radiology Review Assessment by Treatment Arm – Full Analysis



Patient evaluation groups by Treatment Arm – Full Analysis

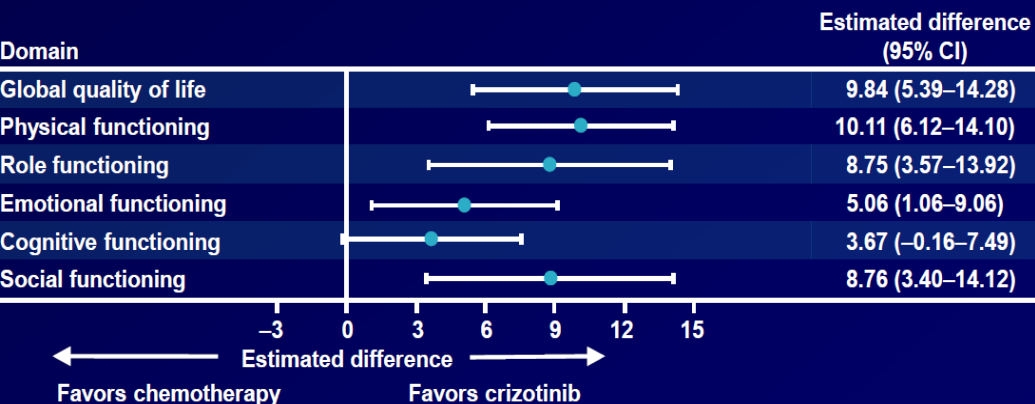
	Crizotinib	Chemotherapy
Randomized to study treatment, n		
-Discontinued, n (%)	87 (50.3)	143 (82.2)
Duration of study treatment (weeks)		
-Median	31.0	12.3



Geplante Analyse - Mixed Linear Model

Patient Reported Outcomes Symptoms and Quality of Life^a

- **SYMPTOMS:** Greater improvement from baseline in cough, dyspnea, fatigue, alopecia, insomnia, and pain with crizotinib (statistically significant: all $P < 0.0001$)^b
- **QUALITY OF LIFE:** Greater improvement from baseline in global quality of life in patients treated with crizotinib (statistically significant: $P < 0.0001$)^b



^aTEORTC QLQ-C30 and QLQ-LC13; ^bbased on a repeated measures mixed-effects model with an intercept, treatment, treatment by time interaction, and subscale baseline score; not adjusted for multiplicity of testing

Standard Analyseverfahren für Repeated Measurement (auch bei zufälligen Fehlenden Werten einsetzbar)

- Effekte: Drug, baseline, time und Interaktion Drug Time
- Therapie-Effekte erfolgt auch unter der Möglichkeit unterschiedlicher Zusammenhänge mit der Zeit in den Therapiegruppen
- Für die Nutzenbewertung nur supportiv nutzbar da Ausmaß des Zusatznutzen nicht nach IQWiG-Methodik bestimmbar
- Abschätzung der Relevanz erfolgt Umrechnung nach Hedges G.

Responder Analyse

- D_{it} = Veränderung eines Symptoms oder HRQoL Domäne zum Zeitpunkt j des Patienten
- $t=1,...,k_i$ $L_i = (1/k_i) \sum_{j=1-k_i} d_{ij}$
- $L_i \geq \text{MID}$ Responder $i=1$, $L_i < \text{MID}$ Responder $=0$
- MID – minimum important Difference - 10 Punkte validierte MID
- Berechne Relatives Risiko mit 95% Konfidenzintervall => Ausmaß des Zusatznutzens nach Matrix
- Beobachtungslänge spielt keine Rolle
- Supportive : Anteil der Zyklen in der ein PatientenRespondiert/ Gesamtzahl der Zyklen eines Patienten

Ergebnisse Beispiel

- Körperliche Funktion
 - **Responder Analyse $RR=0,44$ 95%CI=(0,27;0,72)=> ZN Ausmaß beträchtlich**
 - MLM LSMeans=10,11 95% CI=(6,12;14,0)
 - MLM Hedges $g= 0,56$ 95% CI=(0,34;0,79)
 - Responder Zyklus 16,4% 95% CI=(8,3%;24,5%)

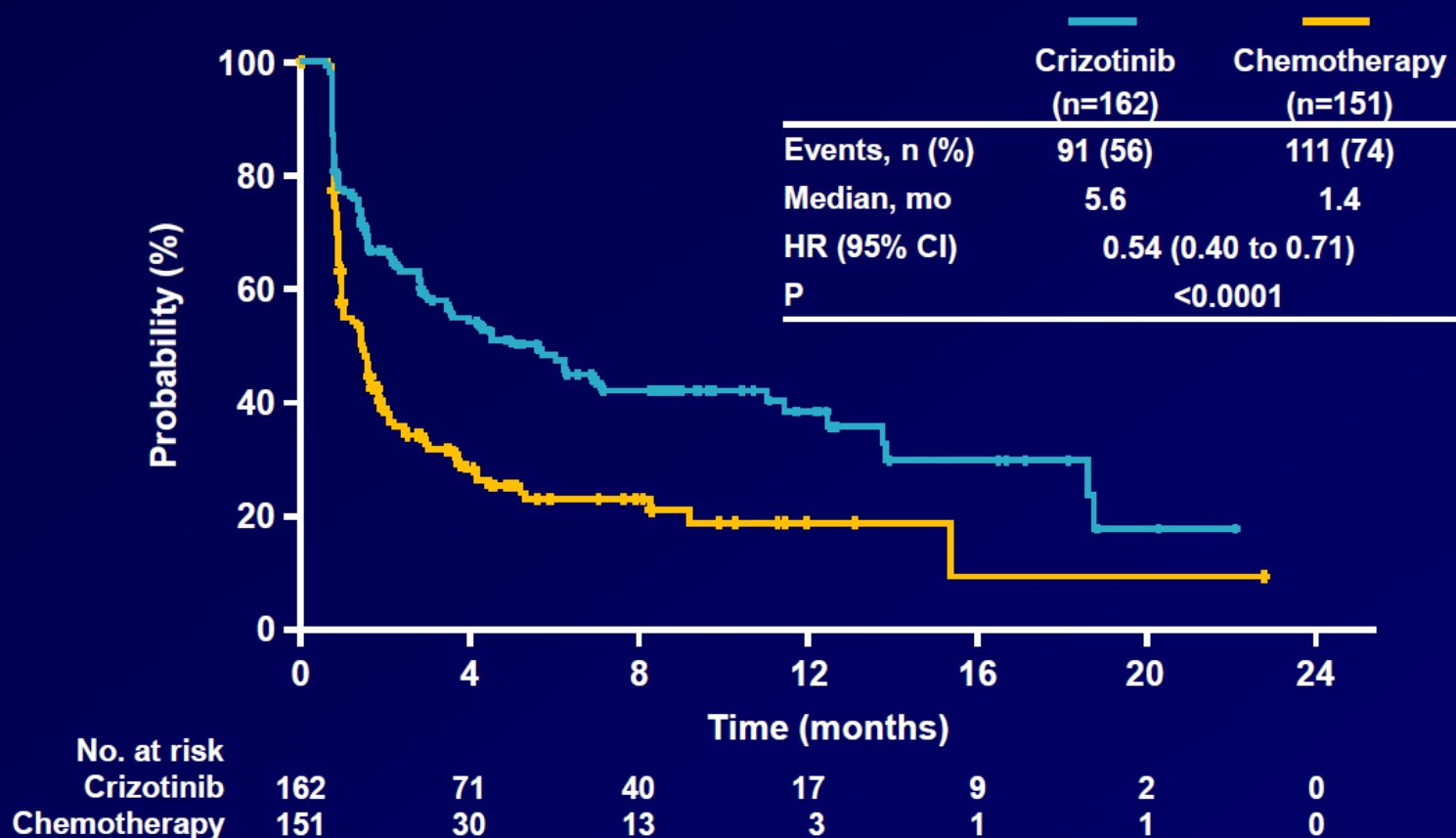
Zeit bis zur Verschlechterung

Time to Deterioration - geplante Analyse

- „Time to deterioration (TTD) “ (nonresponder Analyse)
 - pain, dyspnea or cough from the EORTC-LC13, is defined as the time from randomization to the earliest time the patient’s score shows a 10-point or higher increase after baseline.
 - A 10-point or higher change in the score is perceived by patients as clinically significant
 - For those who have not shown deterioration, patients will be censored at the last date when they completed an assessment (QLQ-LC13) for pain, dyspnea or cough. “
- Zusammengesetzter Endpunkt - alle Komponenten sollten eine ähnliche Bedeutung haben
- Passt zur Time to event Methodik
- Zensierungsgründe
 - Progression oder Tod
 - Abbruch des Patienten
 - Ende der Beobachtung

Patient Reported Outcomes

Time to Deterioration in Lung Cancer Symptoms^a



^aComposite of chest pain, cough, and dyspnea



Verzerrungspotential- informative Zensierung

- ***Profile 1007- Bei Abbruch musste eine HRQoL Fragebogen ausgefüllt werden; von daher ist keine Verzerrung zu erwarten***
- Generell nicht 1007
 - PFS nicht patientenrelevant
 - HRQoL verschlechtert sich im Verlauf-> Effekt wäre grösser bei länger Behandlungsdauer
 - PFS patientenrelevant
 - LQ und PFS korrelieren, d.h. HRQoL sinkt um den Zeitpunkt der Progression
 - Deterioration kann bei einem Teil der Patienten durch die PFS verdeckt sein. Progression häufiger bei Kontrolle > Unterschätzung des Therapieeffekts
 - Wenn keine Messung bei Abbruch (wegen UE) vorliegt, Sensitivitätsanalyse mit komplexeren statistischen Modellen wie Competing Risk-Modelle und Pattern Mixture Modell

Fazit

- HRQoL gibt wichtige patientenrelevante Information
- Anforderung von Regulatoren und HTA-Behörden sind unterschiedlich
- Durch die Nutzbewertung wird HRQoL noch wichtiger
- Methodische Herausforderungen setzen aber auch Grenzen, die verstanden werden sollten.
- Mehr Versorgungsforschung über den Verlauf der HRQoL ist notwendig.
- Blick auf Phase III Programm
 - Früherer Input in Protokoll und Analyseplan soweit wie möglich bei multinationalen Studien
 - Handling und Vermeidung von Missing Values
 - Nutzenbewertungsrelevante Analysen, Sensitivitätsanalysen und Subgruppenanalysen
 - Responder-Analysen mit validierten MID
 - Time To Deterioration



**VIELEN DANK FÜR IHRE
AUFMERKSAMKEIT**