



# Studi sperimentali (parte 2)

Anna Castiglione  
anna.castiglione@cpo.it

19 marzo 2024

# SPIRIT 2013 checklist – 12. Outcomes

- Scelta e definizione dell' **outcome primario** e degli **outcomes secondari** (specificando metrica della variabile, metodo di aggregazione, timing della rilevazione)
- Giustificare le scelte in base a:
  - ✓ Rilevanza clinica
  - ✓ Validità, riproducibilità, responsività (sensibilità all'effetto dell'intervento)
  - ✓ Confrontabilità/Consistenza (Core Outcome Measures in Effectiveness Trials - COMET [www.comet-initiative.org](http://www.comet-initiative.org))

# Scelta degli endpoints negli studi clinici

- Dalla scelta dell'**endpoints primario** dipendono: numerosità campionaria, metodi statistici, utilità dei risultati
- Uno studio comparativo dovrebbe essere finalizzato a dimostrare un vantaggio su un **“Endpoint clinico”** (*“A characteristic or variable that reflects how a patient feels, functions, or survives.”* ad es. sopravvivenza, stato funzionale, sintomi e dolore, QoL)
- Quando l'effetto può agire su più endpoints, gli studi utilizzano **Endpoints Compositi** (> la potenza statistica, < interpretabilità)
- Spesso, per ragioni di efficienza, di tempi e di facilità di conseguimento di risultati favorevoli, gli studi utilizzando in alternativa degli **“Endpoint surrogati”**, non validati

# Gli endpoint surrogati

- Definizione:  
“... a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions or survives. Changes introduced by a therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint” (Temple, 1995)
- Dal 1992, per farmaci potenzialmente importanti, la FDA accetta a fini di registrazione provvisoria risultati preliminari basati su studi non randomizzati e con endpoint surrogati (**Accelerated Approval process**).

# Surrogate End Points in Clinical Research: Hazardous to Your Health

David A. Grimes, MD, and Kenneth F. Schulz, PhD, MBA  
*Family Health International, Research Triangle Park, North Carolina*

**Table 1.** Some Dangerous Discrepancies Between Surrogate End Points and Clinical Outcomes in Trials

Condition	Treatment	Surrogate End Point	What Happened to Trial Participants
Ventricular arrhythmia <sup>10</sup>	Encainide, flecainide	Arrhythmia	Increased sudden death
Osteoporosis <sup>11</sup>	Fluorides	Bone mineral density	Increased fractures
Heart disease in postmenopausal women <sup>12</sup>	Conjugated equine estrogen and medroxyprogesterone acetate	Serum lipids	Increased heart disease
Heart disease <sup>13</sup>	Clofibrate	Serum cholesterol	Increased mortality
Heart failure <sup>14</sup>	Epoprostenol	Exercise tolerance	Increased mortality
Heart failure <sup>15</sup>	Ibopamine	Symptoms of heart failure	Increased mortality

# Surrogate End Points in Clinical Research: Hazardous to Your Health

David A. Grimes, MD, and Kenneth F. Schulz, PhD, MBA  
*Family Health International, Research Triangle Park, North Carolina*

**Table 1.** Some Dangerous Discrepancies Between Surrogate End Points and Clinical Outcomes in Trials

Condition	Treatment	Surrogate End Point	What Happened to Trial Participants
Ventricular arrhythmia <sup>10</sup>	Encainide, flecainide	Arrhythmia	Increased sudden death
Osteoporosis <sup>11</sup>	Fluorides	Bone mineral density	Increased fractures
Heart disease in postmenopausal women <sup>12</sup>	Conjugated equine estrogen and medroxyprogesterone acetate	Serum lipids	Increased heart disease
Heart disease <sup>13</sup>	Clofibrate	Serum cholesterol	Increased mortality
Heart failure <sup>14</sup>	Epoprostenol	Exercise tolerance	Increased mortality
Heart failure <sup>15</sup>	Ibopamine	Symptoms of heart failure	Increased mortality



## Rituximab-dose-dense chemotherapy with or without high-dose chemotherapy plus autologous stem-cell transplantation in high-risk diffuse large B-cell lymphoma (DLCL04): final results of a multicentre, open-label, randomised, controlled, phase 3 study

*Annalisa Chiappella\*, Maurizio Martelli\*, Emanuele Angelucci, Ercole Brusamolino†, Andrea Evangelista, Angelo Michele Carella, Caterina Stelitano, Giuseppe Rossi, Monica Balzarotti, Francesco Merli, Gianluca Gaidano, Vincenzo Pavone, Luigi Rigacci, Francesco Zaja, Alfonso D'Arco, Nicola Cascavilla, Eleonora Russo, Alessia Castellino, Manuel Gotti, Angela Giovanna Congiu, Maria Giuseppina Cabras, Alessandra Tucci, Claudio Agostinelli, Giovannino Ciccone, Stefano A Pileri, Umberto Vitolo*

### Summary

**Background** The prognosis of young patients with diffuse large B-cell lymphoma at high risk (age-adjusted International Prognostic Index [aa-IPI] score 2 or 3) treated with R-CHOP (rituximab, cyclophosphamide, vincristine, doxorubicin, and prednisone) is poor. The aim of this study was to investigate the possible benefit of intensification with high-dose chemotherapy and autologous stem-cell transplantation as part of first-line treatment in these patients.

*Lancet Oncol* 2017; 18: 1076–88

Published Online

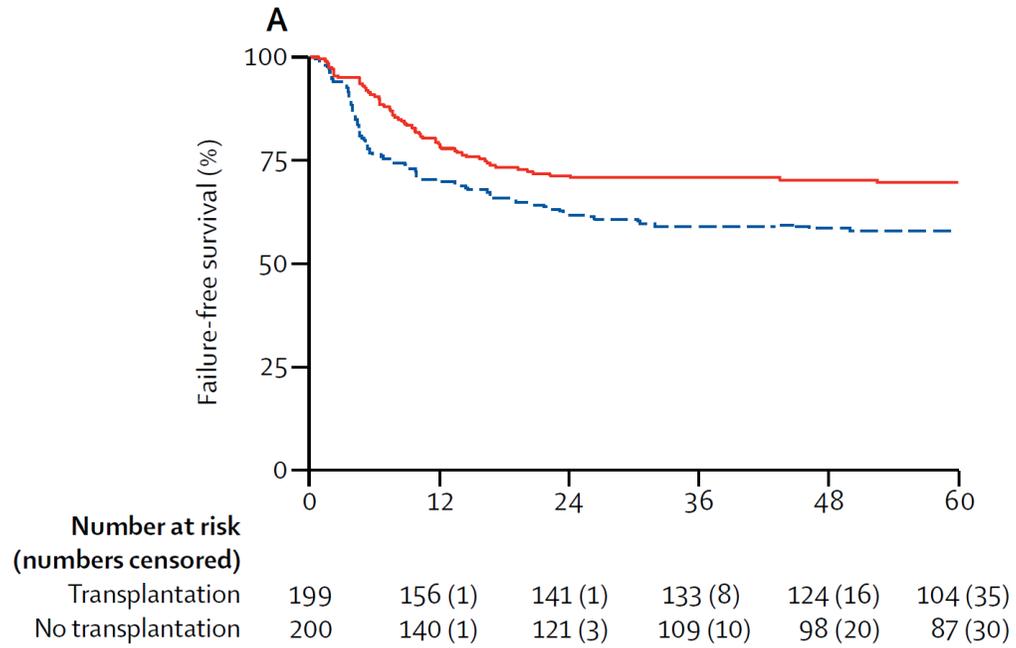
June 28, 2017

[http://dx.doi.org/10.1016/](http://dx.doi.org/10.1016/S1470-2045(17)30444-8)

[S1470-2045\(17\)30444-8](http://dx.doi.org/10.1016/S1470-2045(17)30444-8)

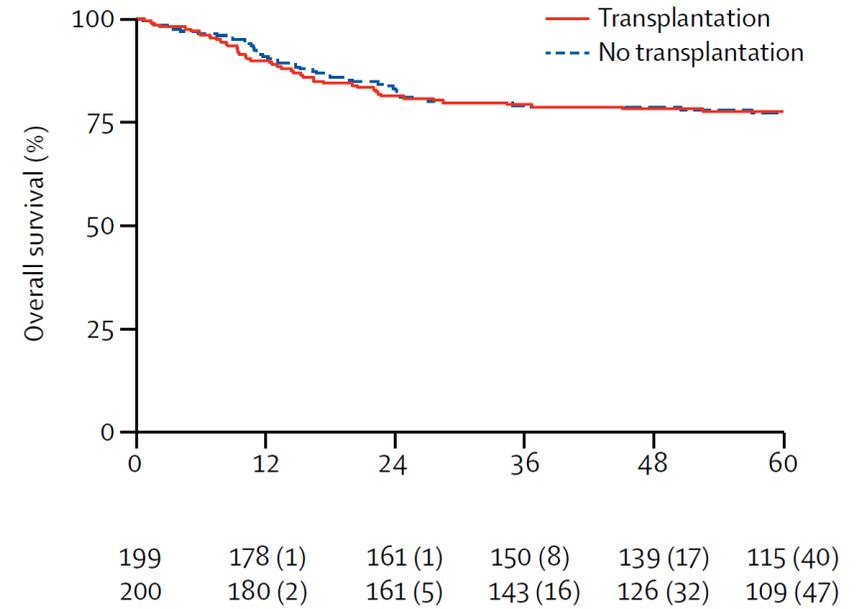
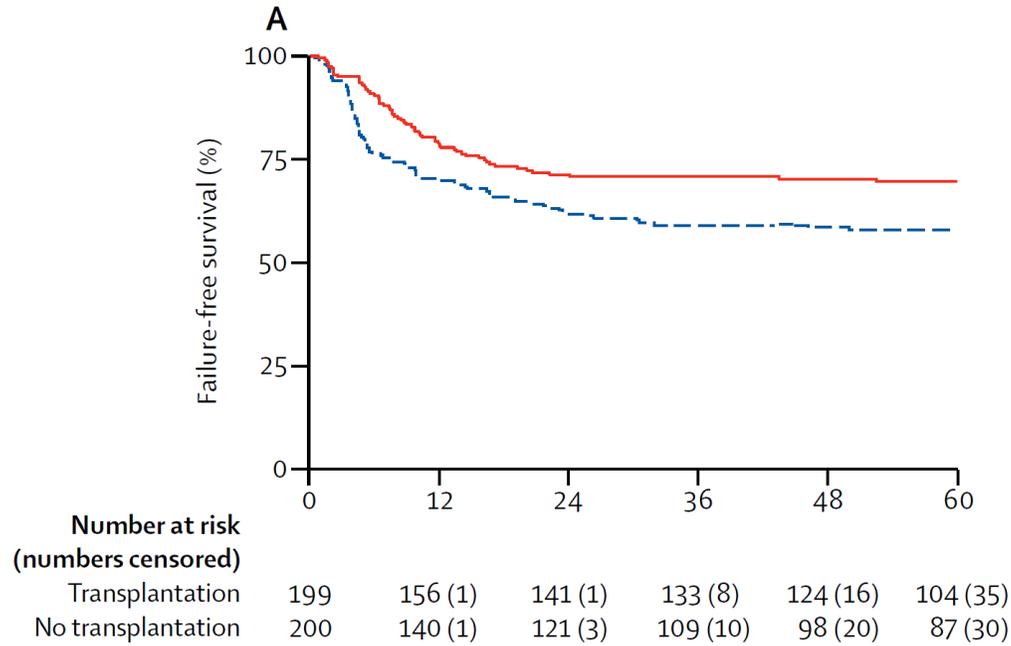
See [Comment](#) page 989

HR=0.65 [95% CI, 0.47–0.91] p=0.012



HR=0.65 [95% CI, 0.47–0.91] p=0.012

HR= 0.98 [95% CI 0.65–1.48] p=0.91



---

# Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study

 OPEN ACCESS

Oriana Ciani *PhD candidate*<sup>1</sup>, Marc Buyse *chairman*<sup>2</sup>, Ruth Garside *senior lecturer*<sup>1</sup>, Toby Pavey *research fellow*<sup>3</sup>, Ken Stein *professor*<sup>1</sup>, Jonathan A C Sterne *professor*<sup>4</sup>, Rod S Taylor *professor*<sup>1</sup>

“On average the **treatment effect estimate was 47% higher in the trials using surrogate outcomes** than in the trials using final patient relevant outcomes (**ratio of odds ratios 1.47**, 95% confidence interval 1.07 to 2.01, P=0.02).”

**Conclusions** Trials reporting surrogate primary outcomes are more likely to report larger treatment effects than trials reporting final patient relevant primary outcomes. This finding was not explained by differences in the risk of bias or characteristics of two groups of trials.

# BMJ Open Protocol for the development of SPIRIT and CONSORT extensions for randomised controlled trials with surrogate primary endpoints: SPIRIT-SURROGATE and CONSORT-SURROGATE

---

Anthony Muchai Manyara ,<sup>1</sup> Philippa Davies,<sup>2</sup> Derek Stewart,<sup>3</sup> Christopher J Weir ,<sup>4</sup> Amber Young ,<sup>2</sup> Nancy J Butcher ,<sup>5,6</sup> Sylwia Bujkiewicz,<sup>7</sup> An-Wen Chan,<sup>8,9</sup> Gary S Collins ,<sup>10</sup> Dalia Dawoud,<sup>11</sup> Martin Offringa ,<sup>6</sup> Mario Ouwens,<sup>12</sup> Joseph S Ross,<sup>13,14</sup> Rod S Taylor,<sup>1,15</sup> Oriana Ciani<sup>16</sup>

# SPIRIT 2013 checklist – 3. Methods

## 13. Timeline

- Importante per presentare sinteticamente lo studio e la complessità organizzativa
- Includere uno schema con le procedure e i tempi previsti, dallo screening iniziale alla conclusione del trial (ad es. quello proposto da SPIRIT, BMJ 2013):

	STUDY PERIOD							
	Enrolment	Allocation	Post-allocation					Close-out
TIMEPOINT*	$-t_1$	0	$t_1$	$t_2$	$t_3$	$t_4$	<i>etc</i>	$t_x$
ENROLMENT:								
Eligibility screen	X							
Informed consent	X							
(List other procedures)	X							
Allocation		X						
INTERVENTIONS:								
(Intervention A)			←————→					
(Intervention B)			X		X			
(List other study groups)			←————→					
ASSESSMENTS:								
(List baseline variables)	X	X						
(List outcome variables)				X		X	<i>etc</i>	X
(List other data variables)			X	X	X	X	<i>etc</i>	X

\* List specific timepoints in this row

# SPIRIT 2013 checklist – 3. Methods

## 14. Sample size

- Un calcolo corretto e realistico è un aspetto chiave nella pianificazione di un trial per prevedere:
  - ✓ Tempi
  - ✓ Costi
  - ✓ Fattibilità
- Spesso il calcolo si basa solo sul confronto dell'outcome principale, ma è opportuno calcolare anche la numerosità richiesta per qualche importante outcome secondario

# SPIRIT 2013 checklist – 3. Methods

## 14. Sample size

**Table 4 | Outcome values to report in sample size calculation**

Element	Type of summary outcome		
	Binary	Continuous	Time to event
Assumed result for each study group	Proportion (%) with event	Mean and standard deviation	Proportion (%) with event at a given time point
Effect measure	Relative risk, odds ratio	Difference in means	Hazard ratio

Note: Although the sample size calculation uses the expected outcome value for each group, the corresponding contrast between groups (estimated effect) should also be reported.

# SPIRIT 2013 checklist – 3. Methods

## 14 Sample size

Gli elementi essenziali per il calcolo sono:

- per studi di superiorità: la **minima differenza clinicamente rilevante**, specificando i valori attesi dell'outcome nei gruppi a confronto ad un tempo definito
- per studi di non-inferiorità: il **margin di non inferiorità** (differenza massima considerata trascurabile)
- Gli **errori statistici** accettati:
  - ✓ Alfa (o di tipo I): di solito 0.05 (specificare se a 1 o 2 code)
  - ✓ Beta (o di tipo II): di solito 0.10-0.20 (o potenza statistica: 0.80-0.90)
- Il test statistico per il confronto principale (es. log-rank test x curve di sopravvivenza)
- Eventuali aggiustamenti per:
  - ✓ analisi ad interim
  - ✓ perdite di pazienti/dati durante lo studio

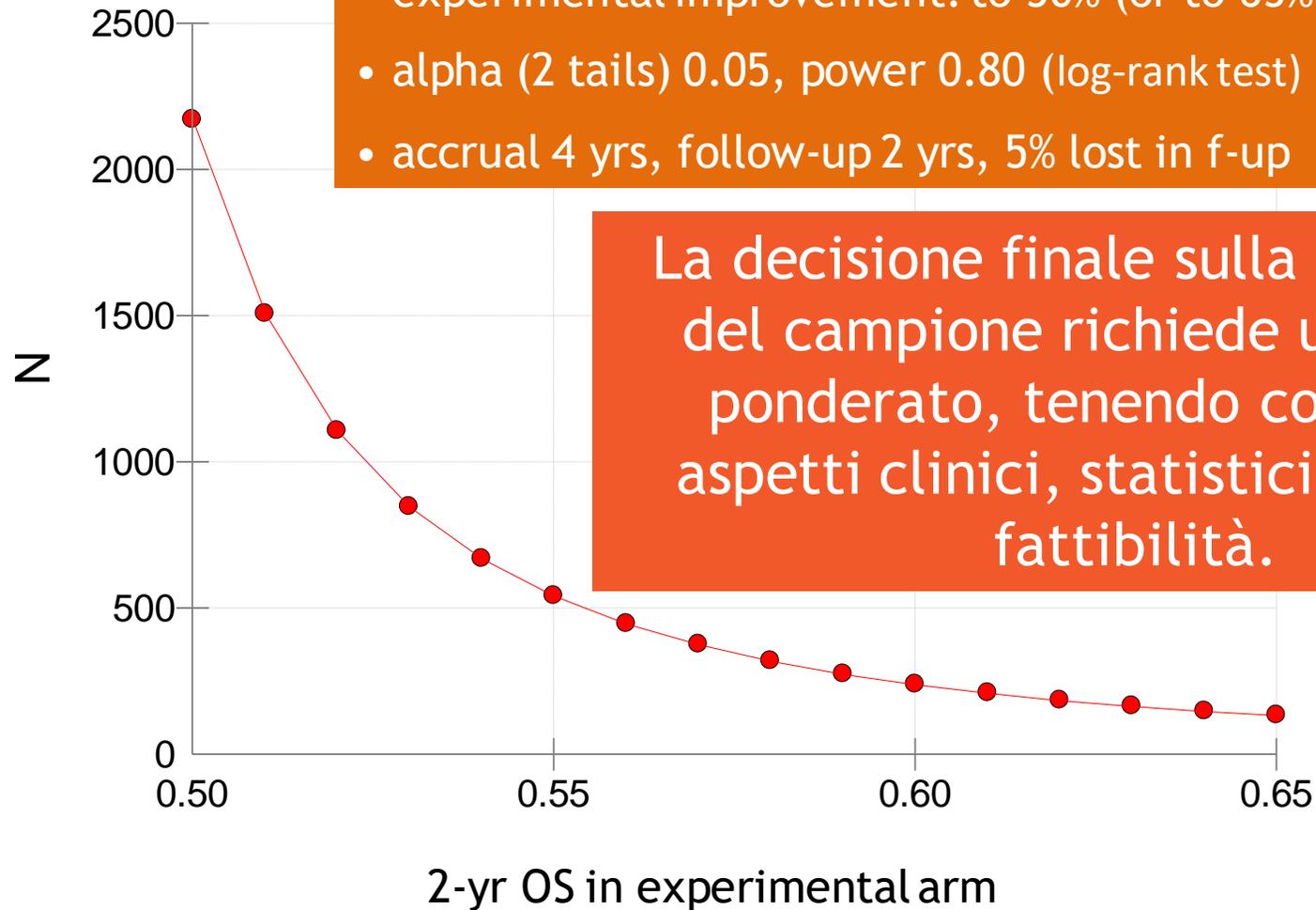
# SPIRIT 2013 checklist – 3. Methods

## 14-15 Sample size, Arruolamento

- Per studi adattivi: specificare esattamente i criteri per modificare le dimensioni durante lo studio
- La difficoltà di reclutamento è molto frequente:
  - ✓ Utile prevedere l'arruolamento atteso per centro
  - ✓ Prevedere strategie per favorire l'arruolamento

# Esempio: Dimensionamento dello studio

- expected 2-yr OS in control arm: 45%
- experimental improvement: to 50% (or to 65%)
- alpha (2 tails) 0.05, power 0.80 (log-rank test)
- accrual 4 yrs, follow-up 2 yrs, 5% lost in f-up



# SPIRIT 2013 checklist – 3. Methods

## 16 Metodi di randomizzazione

### Box 1 | Key elements of random sequence to specify in trial protocols

---

- Method of sequence generation (eg, random number table or computerised random number generator)
- Allocation ratio (Item 8) (eg, whether participants are allocated with equal or unequal probabilities to interventions)
- Type of randomisation (box 2): simple versus restricted; fixed versus adaptive (eg, minimisation); and, where relevant, the reasons for such choices
- If applicable, the factors (eg, recruitment site, sex, disease stage) to be used for stratification (box 2), including categories and relevant cut-off boundaries

# SPIRIT 2013 checklist – 3. Methods

## 16 Metodi di randomizzazione

- Metodi di generazione della sequenza di randomizzazione (computer generated random numbers)
- Tipi di sequenze di randomizzazione (semplice, stratificata, a blocchi et c)
- L'algoritmo di generazione di randomizzazioni ristrette non deve essere esplicitato nel protocollo, ma tenuto protetto in un documento separato
- Metodi di protezione e di utilizzazione della sequenza di randomizzazione:
  - ✓ Buste opache per-numerate
  - ✓ Randomizzazione centralizzata telefonica o su sito web dedicato
  - ✓ Prevedere metodi per assicurare l'assegnazione irreversibile
- Separazione dei compiti tra:
  - ✓ chi ha generato (gestisce) la sequenza di randomizzazione
  - ✓ chi arruola, randomizza e tratta il paziente

# SPIRIT 2013 checklist – 3. Methods

## 16 Metodi di randomizzazione

### Example

---

“Participants will be randomly assigned to either control or experimental group with a 1:1 allocation as per a computer generated randomisation schedule stratified by site and the baseline score of the Action Arm Research Test (ARAT;  $\leq 21$  versus  $>21$ ) using permuted blocks of random sizes. The block sizes will not be disclosed, to ensure concealment.”<sup>224</sup>

# SPIRIT 2013 checklist – 3. Methods

## 17 Metodi per il Mascheramento

- Il mascheramento è un metodo per nascondere il gruppo di trattamento al paziente, allo sperimentatore e allo statistico
- Obiettivi del mascheramento:
  - ✓ Ridurre la misclassificazione dell'outcome clinici (da parte dello sperimentatore) e dei patient reported outcomes (da parte del paziente)
  - ✓ Aumentare l'aderenza del paziente al trattamento (**compliance**)
  - ✓ Prevenire l'interruzione precoce dello studio (da parte dello sperimentatore o da parte del paziente)
- Il placebo è una delle tecniche utilizzate per 'mascherare' il braccio di assegnazione

# SPIRIT 2013 checklist – 3. Methods

## 17 Metodi per il Mascheramento

- Specificare chi è “in cieco” (evitare “single”, “double” blinding):
  - ✓ Paziente
  - ✓ Medico
  - ✓ Valutatore
  - ✓ Statistico
  - ✓ Estensore dell’articolo
- Specificare quali condizioni giustificano l’apertura della cecità

# SPIRIT 2013 checklist – 3. Methods

## 18-19 Raccolta e gestione della raccolta dati

Protocollo di valutazione e raccolta dei dati (di baseline, di processo durante il trial, sugli outcomes)

- Gli strumenti (questionari, CRF) devono essere disponibili (in appendice, o attraverso referenze)
- Necessaria standardizzazione dei metodi (clinici, di laboratorio, di imaging) e degli strumenti
- Utile prevedere training e fasi pilota per standardizzare metodi
- Utile misurare la concordanza tra valutatori (in cieco)
- Utilizzo raccomandato di strumenti (questionari, scale, scores) già validati (da non modificare)
- Prevedere controlli di qualità dei dati (per ridurre incongruenze, missing)

# SPIRIT 2013 checklist – 3. Methods

## 18-19 Raccolta e gestione della raccolta dati

Prevedere misure per ridurre al minimo le perdite durante il follow-up

- Tutti i pazienti randomizzati costituiscono la popolazione da analizzare secondo l'approccio Intention To Treat
- I pazienti che “escono” dal trattamento previsto dal protocollo, “non escono” anche dallo studio
- Ai pazienti che rifiutano alcune valutazioni previste dal protocollo (es. esami invasivi) deve essere offerta la possibilità di continuare ad essere valutati per tutti gli altri outcomes

Descrivere le procedure per garantire il data entry, la codifica, la qualità, la sicurezza e l'archiviazione dei dati:

- Il data-entry in sistemi di eCRF su Internet è preferibile alla raccolta cartacea
- Utilizzare codifiche standard
- Attivare controlli di congruità automatici e di tipo statistico
- Verificare (in modo campionario) la qualità dei dati vs i «source records»

# SPIRIT 2013 checklist – 3. Methods

## 20 Metodi Statistici

Metodi di analisi degli outcomes (primario e secondari):

- Il piano di analisi statistica (SAP) deve essere chiaramente definito nel protocollo (con eventuale allegato per i dettagli)
- È necessario specificare:
  - ✓ La tecnica di analisi (grezza, aggiustata)
  - ✓ La misura di esito (risk difference, risk ratio, odds ratio, hazard ratio, ...)
  - ✓ Il test statistico e il metodo di calcolo dell'intervallo di confidenza della misura di effetto
- Limitare al minimo il numero di confronti (inflazione di risultati falsi positivi)
- Prevedere solo reporting descrittivo per eventi rari

# SPIRIT 2013 checklist – 3. Methods

## 20 Metodi Statistici

Metodi per analisi aggiuntive (di sottogruppo, aggiustate):

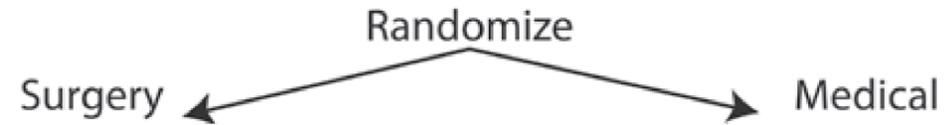
- Analisi di sottogruppo:
  - ✓ devono essere pre-spezifcate, con un rationale e di numero limitato
  - ✓ sono fonte di molti risultati spuri (nel 75% degli studi pubblicati non corrispondono a quelle previste nel protocollo)
  - ✓ devono essere limitate a variabili misurate al baseline (non ad eventi post-randomizzazione)
- Analisi stratificate/aggiustate:
  - ✓ È l'analisi corretta quando si usano variabili di stratificazione
  - ✓ Pre-definire le variabili da considerare per l'aggiustamento (almeno la logica da seguire)
  - ✓ Se sono previste sia analisi grezze, sia aggiustate, anticipare quale sarà considerato il risultato principale
  - ✓ Non aggiustare per variabili misurate dopo la randomizzazione

# SPIRIT 2013 checklist – 3. Methods

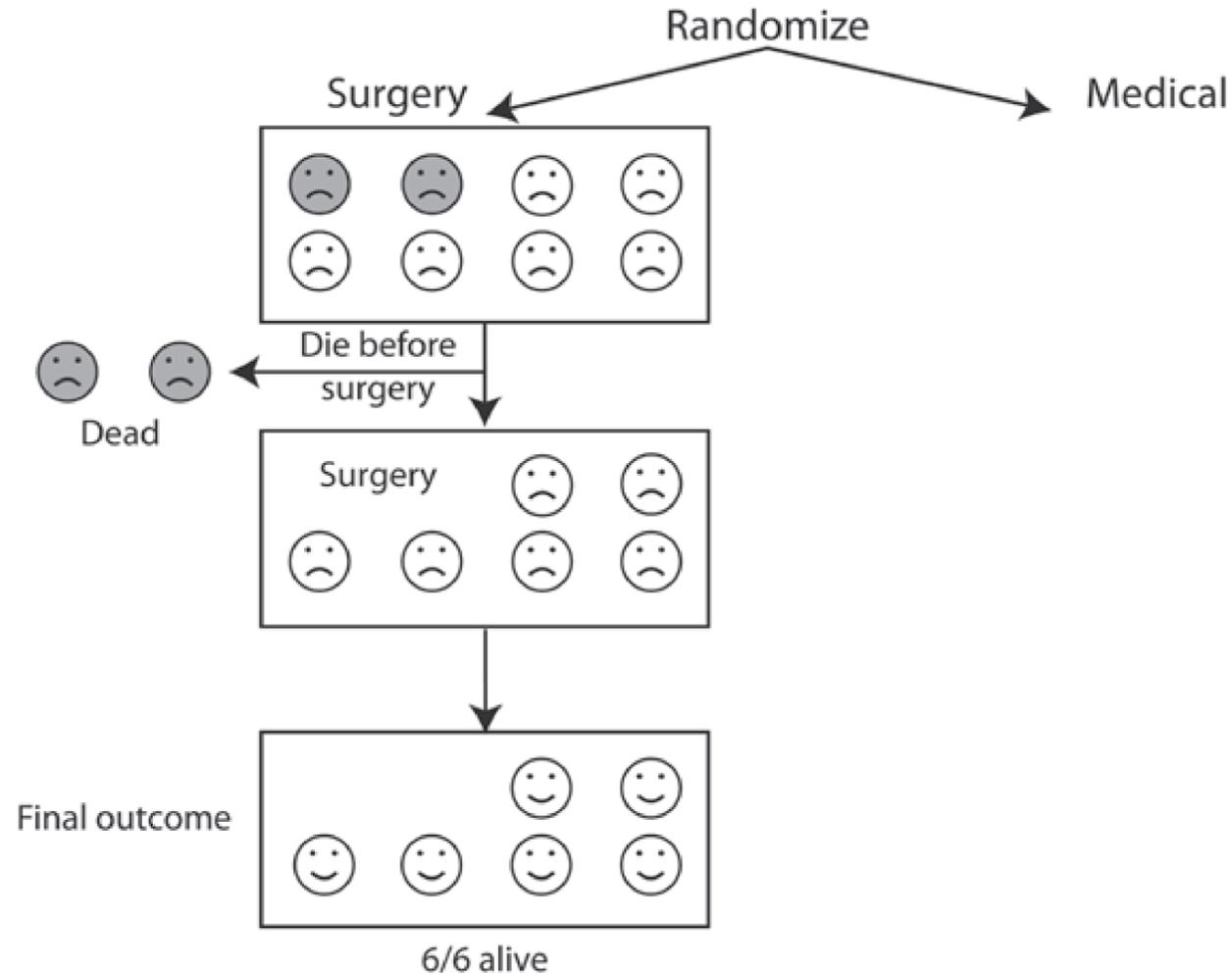
## 20 Metodi Statistici

- Esclusioni di pazienti post-randomizzazione:
  - Per “screening failures”
  - Revisioni istologiche
  - “Drop outs”!!!
- **Drop out:** nessun paziente dovrebbe “uscire dallo studio” per motivi clinici, di tossicità o di scarsa aderenza al protocollo (salvo esplicita richiesta del pz, con ritiro del consenso)
- **Aggiornamento** del follow-up completo a fine studio

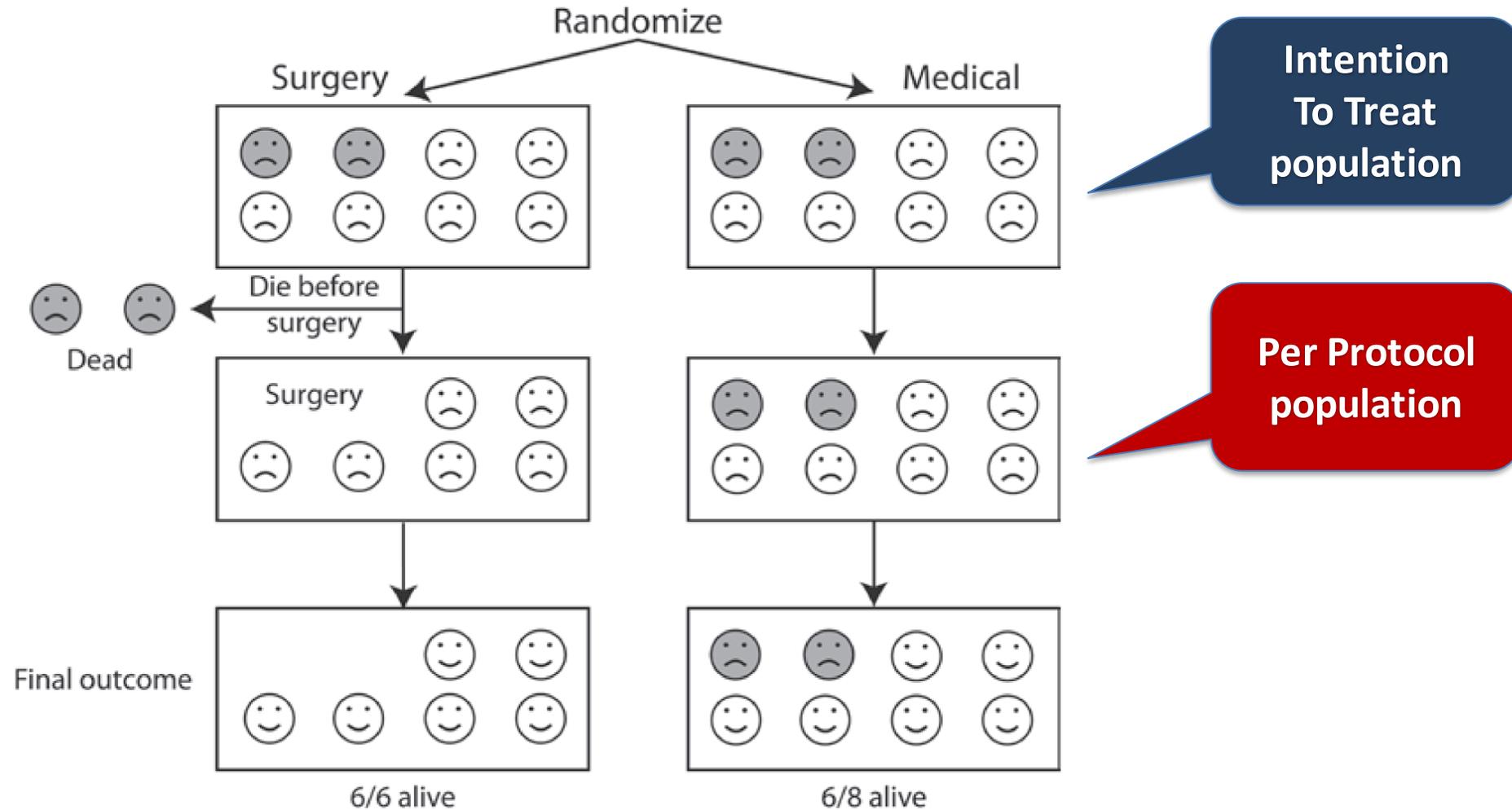
# Analisi *intention to treat* (ITT)



# Analisi *intention to treat* (ITT)



# Analisi *intention to treat* (ITT)



# Analisi di sottogruppo

- In un RCT le analisi di sottogruppo sono finalizzate a valutare l'omogeneità dell'effetto del trattamento rispetto a particolari caratteristiche e condizioni dei partecipanti.
- Usando una terminologia più epidemiologica, siamo interessati a valutare l'eventuale modificazione di effetto rispetto a particolari fattori della terapia sperimentale sugli outcome dello studio.
- Se lo studio non è formalmente dimensionato per evidenziare specifiche modificazioni di effetto, le analisi di sottogruppo hanno finalità esplorative e suggeriscono (anche importanti) spunti di ricerca.
- Sono molto più credibili quando sono esplicitamente dichiarate nel protocollo.
- Sono spesso condotte (e commentate) in maniera fuorviante utilizzando i p-value strato-specifici del confronto tra trattamenti

# Esempio (scorretto) di Analisi di sottogruppo



## Relapse and disability outcomes in patients with multiple sclerosis treated with fingolimod: subgroup analyses of the double-blind, randomised, placebo-controlled FREEDOMS study

*Virginia Devonshire, Eva Havrdova, Ernst Wilhelm Radue, Paul O'Connor, Lixin Zhang-Auberson, Catherine Agoropoulou, Dieter Adrian Häring, Gordon Francis, Ludwig Kappos, for the FREEDOMS study group\**

### Summary

*Lancet Neurol* 2012; 11: 420-28

Published Online

April 10, 2012

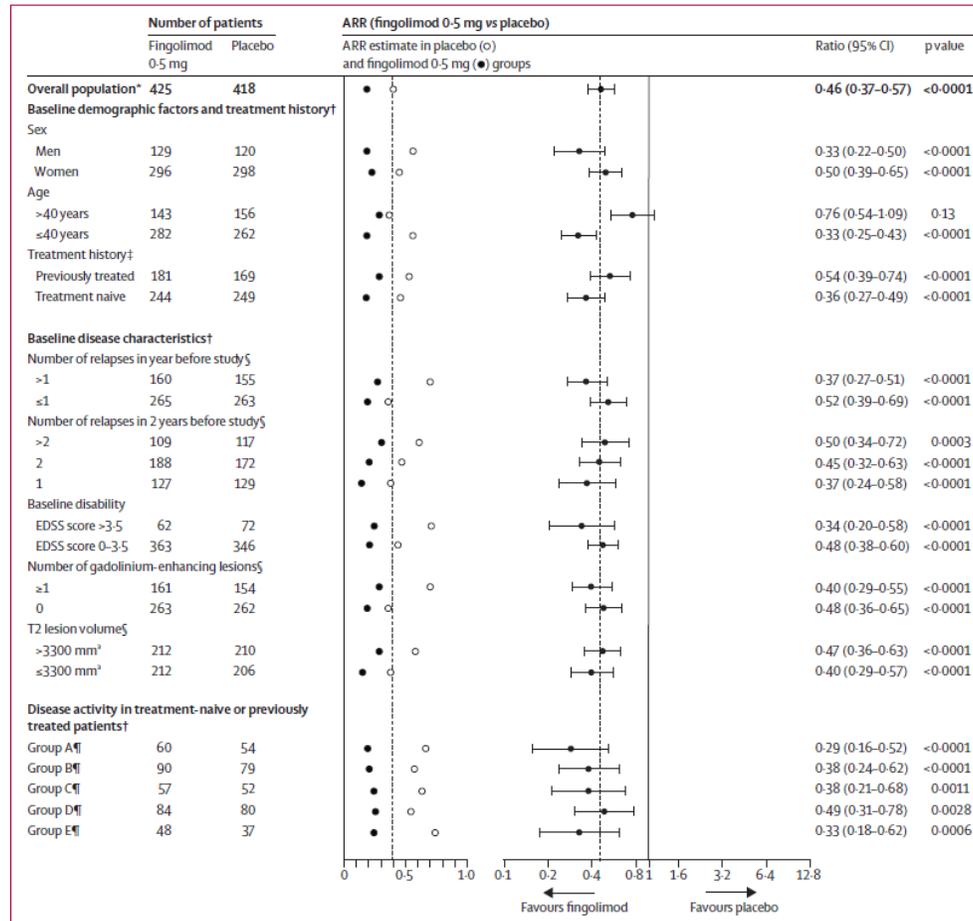
DOI:10.1016/S1474-

4422(12)70056-X

This online publication

**Background** Fingolimod 0.5 mg once daily is approved for treatment of relapsing multiple sclerosis (MS). In the phase 3, 2-year FREEDOMS (FTY720 Research Evaluating Effects of Daily Oral therapy in MS) study, fingolimod significantly reduced annualised relapse rates (ARRs) and the risk of confirmed disability progression compared with placebo. We aimed to investigate whether the beneficial treatment effect reported for the overall population is consistent in subgroups of patients with different baseline characteristics.

# Esempio (scorretto) di Analisi di sottogruppo



Devonshire et al, Lancet Neurol 2012; 11: 420-28

# Esempio (scorretto) di Analisi di sottogruppo

	Number of patients		ARR (fingolimod 0.5 mg vs placebo)		Ratio (95% CI)	p value
	Fingolimod 0.5 mg	Placebo	ARR estimate in placebo (○) and fingolimod 0.5 mg (●) groups			
Overall population*	425	418			0.46 (0.37-0.57)	<0.0001
<b>Baseline demographic factors and treatment history†</b>						
Sex						
Men	129	120			0.33 (0.22-0.50)	<0.0001
Women	296	298			0.50 (0.39-0.65)	<0.0001
Age						
>40 years	143	156			0.76 (0.54-1.09)	0.13
≤40 years	282	262			0.33 (0.25-0.43)	<0.0001
Treatment history‡						
Previously treated	181	169			0.54 (0.39-0.74)	<0.0001
Treatment naive	244	249			0.36 (0.27-0.49)	<0.0001

# Esempio (scorretto) di Analisi di sottogruppo

Fingolimod 0.5 mg significantly reduced ARR<sub>s</sub> over 2 years versus placebo in all subgroups (table 1, figure 1) except for patients aged over 40 years (24% decrease in ARR;  $p=0.13$ ). In the other patient groups defined by baseline demographic factors or treatment history, ARR<sub>s</sub> were significantly reduced by fingolimod 0.5 mg by up to 67% versus placebo over 24 months. The greatest reductions were noted in men and in patients aged 40 years or less (67% and  $p<0.0001$  for both).

# Esempio (scorretto) di Analisi di sottogruppo

In conclusion, it is desirable that high-quality subgroup analyses such as the present one adopt comparable standards for the reporting of their results, on the basis of the interaction tests, while subgroup-specific p values are finally buried in the cemetery of wrong statistics.

# Metodo raccomandato per Analisi di sottogruppo

- La valutazione dell'omogeneità dell'effetto del trattamento viene effettuata utilizzando modelli di regressione appropriati per gli outcome in studio, prevedendo tra le variabili indipendenti il trattamento sperimentale, la variabile che identifica il sottogruppo di interesse e il loro termine di interazione.
- La significatività del test associato al termine di interazione suggerisce la non omogeneità dell'effetto del trattamento rispetto ai livelli di quello specifico sottogruppo.
- Dato che entro sottogruppi lo studio non è più controllato dalla randomizzazione, è fortemente consigliato l'aggiustamento per noti fattori associati all'esito in studio

# Esempio di Analisi di sottogruppo

VOLUME 31 · NUMBER 27 · SEPTEMBER 20 2013

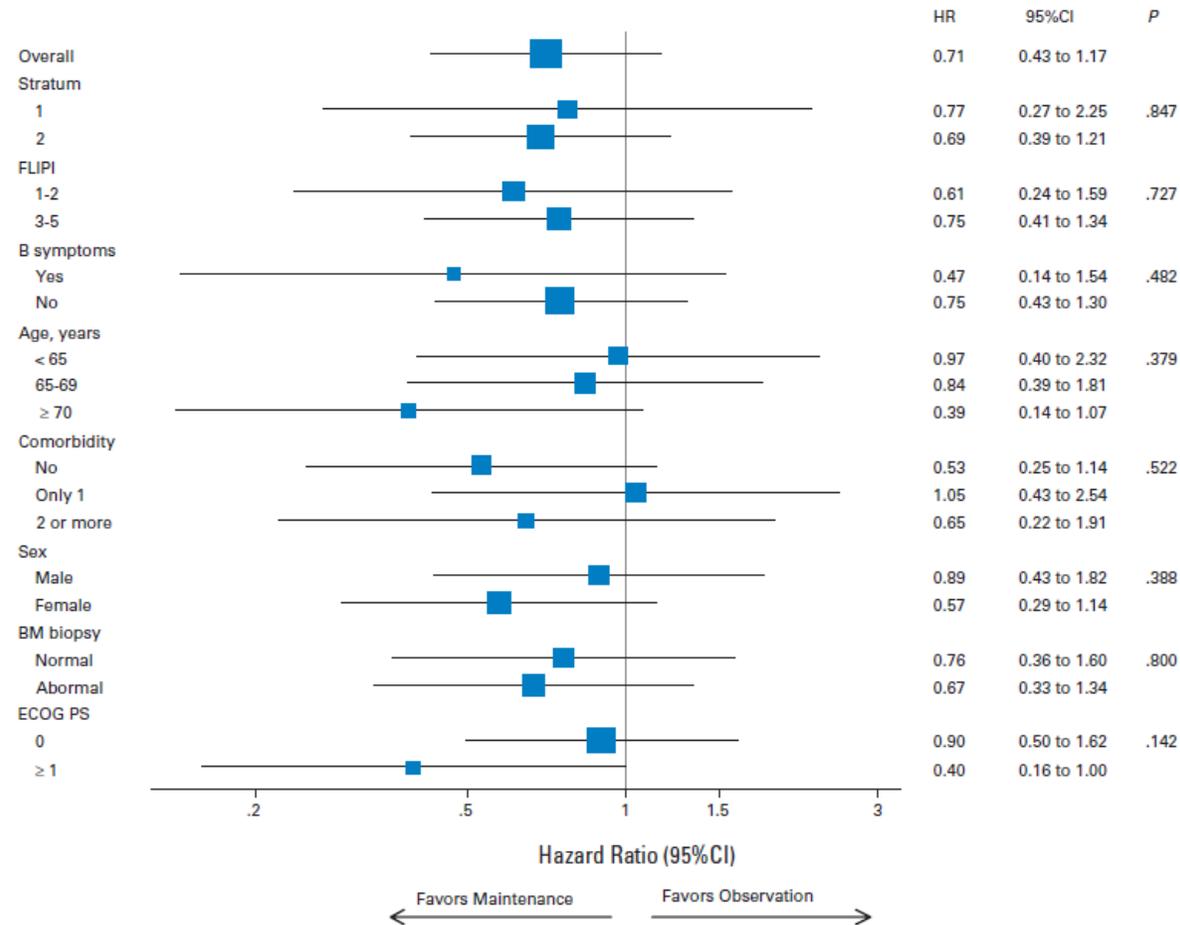
JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

## Rituximab Maintenance Compared With Observation After Brief First-Line R-FND Chemoimmunotherapy With Rituximab Consolidation in Patients Age Older Than 60 Years With Advanced Follicular Lymphoma: A Phase III Randomized Study by the Fondazione Italiana Linfomi

*Umberto Vitolo, Marco Ladetto, Carola Boccomini, Luca Baldini, Federico De Angelis, Alessandra Tucci, Barbara Botto, Annalisa Chiappella, Annalisa Chiarenza, Antonello Pinto, Amalia De Renzo, Francesco Zaja, Claudia Castellino, Alessia Bari, Isabel Alvarez De Celis, Andrea Evangelista, Guido Parvis, Enrica Gamba, Chiara Lobetti-Bodoni, Giovannino Ciccone, and Giuseppe Rossi*

# Esempio di Analisi di sottogruppo



# Analisi intermedie

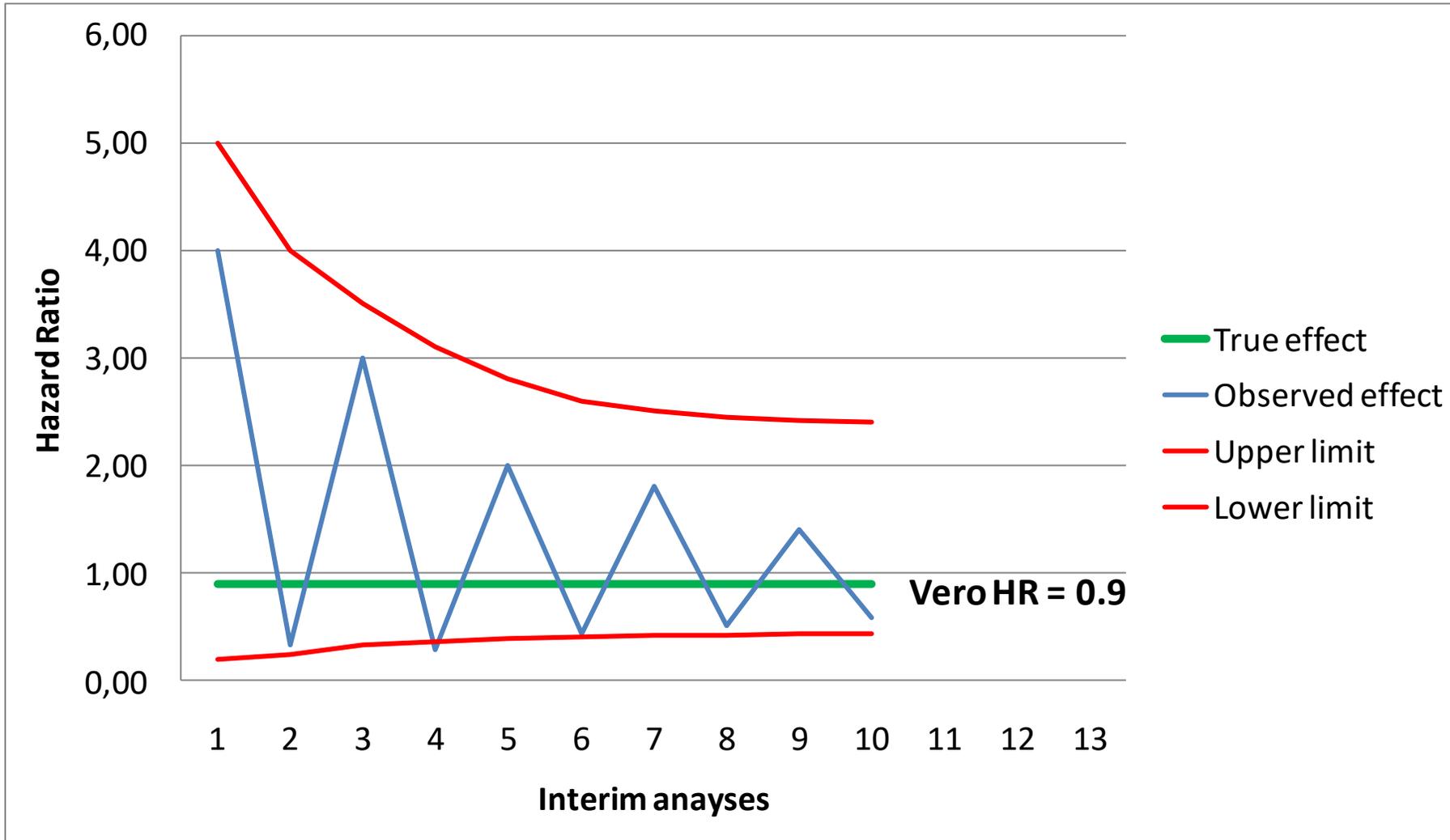
- Analisi comparative sull' endpoint primario per valutare se si osserva una differenza "significativa" tra i due gruppi, tale da richiedere l'interruzione dello studio.
- Scopi:
  - ✓ Interrompere l'arruolamento (non e' più etico randomizzare pazienti nello studio)
  - ✓ Passare i pazienti in trattamento alla terapia più efficace
- Metodologia consigliata:
  - ✓ Si decidono in anticipo le occasioni in cui si faranno le analisi (es. 1/2, 1/3 dell'arruolamento)
  - ✓ Si stabilisce il livello di significativita' richiesto per interrompere lo studio ( $\ll 0.05$ )
  - ✓ Di solito livelli molto stringenti alle analisi intermedie per salvaguardare l'analisi finale

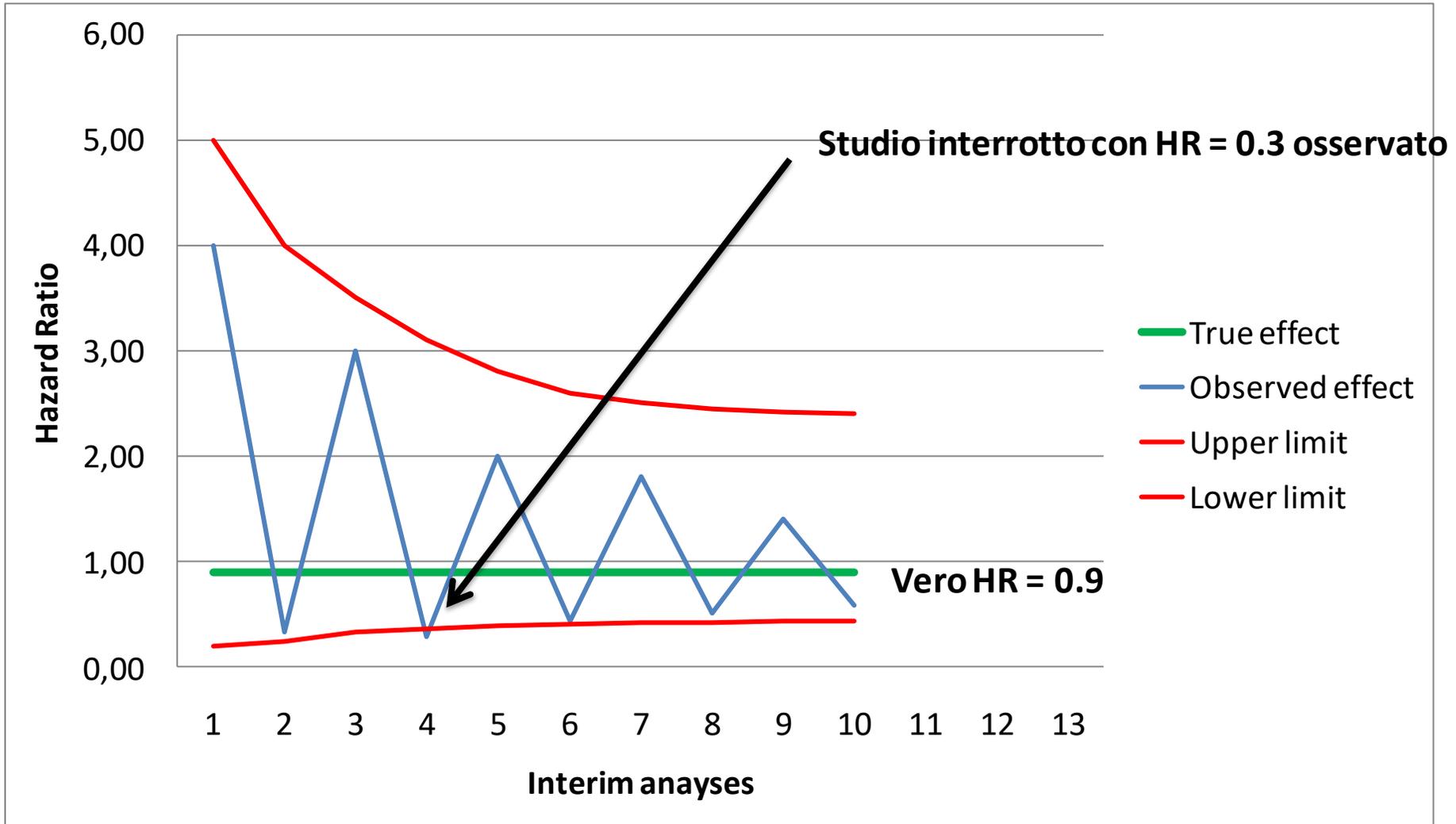
# Esempio di Analisi intermedie protocollo MCL0208

- According to O'Brien and Fleming group sequential design with a maximum of three stages, a total of 200 complete or partial responders (100 per group) should be randomised [...].
- Under these assumptions, two interim analyses will be performed to timely detect larger and unexpected differences in the PFS between arms, when 1/3 and 2/3 of the expected events will be observed.
- To preserve the overall type I error, critical p values are 0.0006 for the first interim analysis, 0.0142 for the second and 0.045 for the final analysis.

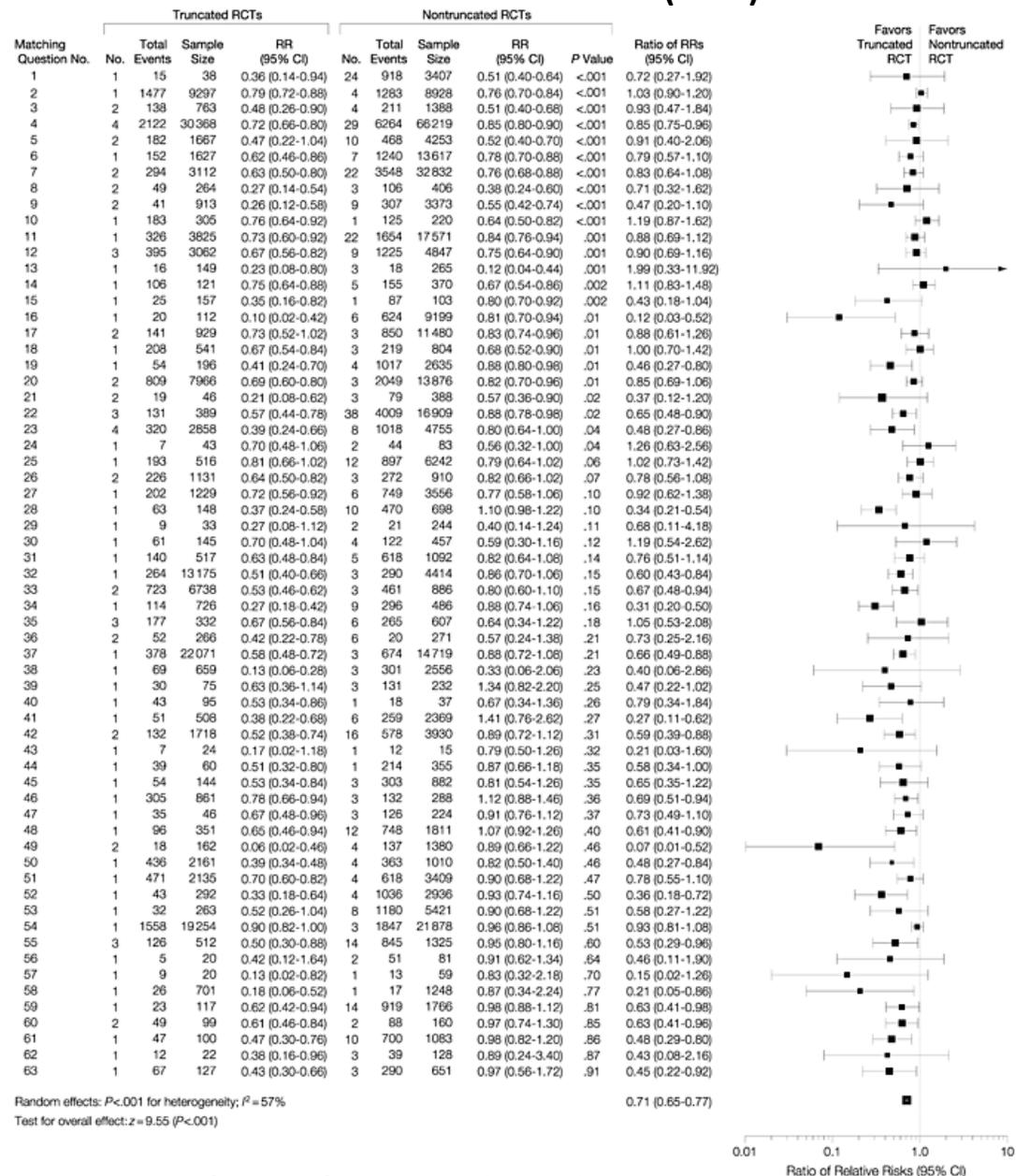
# Nota bene

- Le 'spending functions' vengono calcolate in modo da preservare **l'errore alfa** complessivo
- Per motivi statistici, gli studi interrotti precocemente forniscono sistematicamente una **sovrastima** del beneficio del trattamento piu' efficace
- Non si dovrebbero quindi utilizzare queste stime per confronti e valutazioni cliniche



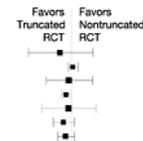


# Pooled Ratio of Relative Risks (RRs) and 95% Confidence Intervals (CIs) for Truncated vs Nontruncated Randomized Controlled Trials (RCTs)



## Pooled Ratio of Relative Risks (RRs) and 95% Confidence Intervals (CIs) for Truncated vs Nontruncated Randomized Controlled Trials (RCTs)

Matching Question No.	Truncated RCTs			Nontruncated RCTs			P Value	Ratio of RRs (95% CI)	
	No.	Total Events	Sample Size	No.	Total Events	Sample Size			
1	1	15	38	24	918	3407	0.51 (0.40-0.64)	<.001	0.72 (0.27-1.92)
2	1	1477	9297	4	1283	8928	0.79 (0.72-0.88)	<.001	1.03 (0.90-1.20)
3	2	138	763	4	211	1388	0.48 (0.26-0.90)	<.001	0.93 (0.47-1.84)
4	4	2122	30368	29	6264	66219	0.72 (0.66-0.80)	<.001	0.85 (0.75-0.96)
5	2	182	1957	10	408	4253	0.47 (0.22-1.04)	<.001	0.91 (0.40-2.06)
6	1	152	1627	7	1240	13617	0.62 (0.46-0.86)	<.001	0.79 (0.57-1.10)
7	2	294	3112	22	3545	32832	0.63 (0.50-0.80)	<.001	0.83 (0.64-1.08)

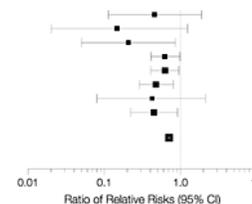


**Results** The analysis included 91 truncated RCTs asking 63 different questions and 424 matching nontruncated RCTs. The pooled ratio of relative risks in truncated RCTs vs matching nontruncated RCTs was 0.71 (95% confidence interval, 0.65-0.77). This difference was independent of the presence of a statistical stopping rule and the methodological quality of the studies as assessed by allocation concealment and blinding. Large differences in treatment effect size between truncated and nontruncated RCTs (ratio of relative risks <0.75) occurred with truncated RCTs having fewer than 500 events. In 39 of the 63 questions (62%), the pooled effects of the nontruncated RCTs failed to demonstrate significant benefit.

**Conclusions** Truncated RCTs were associated with greater effect sizes than RCTs not stopped early. This difference was independent of the presence of statistical stopping rules and was greatest in smaller studies.

56	1	5	20	0.42 (0.19-1.04)	2	51	61	0.91 (0.62-1.34)	.24	0.46 (0.11-1.90)
57	1	9	20	0.13 (0.02-0.82)	1	13	69	0.83 (0.32-2.16)	.70	0.15 (0.02-1.26)
58	1	26	701	0.18 (0.06-0.52)	1	17	1248	0.87 (0.34-2.24)	.77	0.21 (0.05-0.86)
59	1	23	117	0.62 (0.42-0.94)	14	919	1766	0.98 (0.88-1.12)	.81	0.63 (0.41-0.98)
60	2	49	99	0.61 (0.46-0.84)	2	88	160	0.97 (0.74-1.30)	.85	0.63 (0.41-0.96)
61	1	47	100	0.47 (0.30-0.76)	10	700	1083	0.98 (0.82-1.20)	.86	0.48 (0.29-0.80)
62	1	12	22	0.38 (0.16-0.96)	3	39	128	0.89 (0.24-3.40)	.87	0.43 (0.08-2.16)
63	1	67	127	0.43 (0.30-0.66)	3	290	651	0.97 (0.66-1.72)	.91	0.45 (0.22-0.92)

Random effects:  $P < .001$  for heterogeneity;  $I^2 = 57\%$   
 Test for overall effect:  $z = 9.55$  ( $P < .001$ )



Bassler, D. et al. JAMA 2010;303:1180-1187.

# SPIRIT 2013 checklist – 3. Methods

## 21. Data monitoring committee (DMC)

21a. Composizione, ruolo, indipendenza del DMC:

- *Quando?* Trials di lunga durata e con possibili rischi è necessario un monitoraggio attento dell'andamento dei risultati (Il 40% dei trials registrati su ClinicalTrials.gov ha un DMC)
- *Composizione:* esperti con diverse competenze (cliniche, statistiche, etiche) del tutto indipendenti dallo sponsor, dagli investigatori e senza conflitti di interesse rispetto ai risultati del trial (da dichiarare formalmente)
- *Perche?:*
  - ✓ sorvegliare periodicamente i risultati che si accumulano
  - ✓ proporre eventuali modifiche al protocollo o la sospensione dello studio

# SPIRIT 2013 checklist – 3. Methods

## 21. Data monitoring committee (DMC)

### 21b. Analisi ad interim:

- Sono previste nei trials considerati rischiosi, secondo diversi approcci statistici
- Aiutano a decidere se interrompere prematuramente lo studio:
  - ✓ Per eccesso di rischio (in genere senza valutazioni statistiche formali)
  - ✓ Per eccesso di beneficio (anche in base a stopping rules statistiche)
  - ✓ Per futilità (scarsa probabilità di rifiutare l'ipotesi nulla)
- Necessario specificare le regole per l'interruzione dello studio (statistiche e non), chi può essere a conoscenza dei risultati in itinere e chi prende le decisioni

# SPIRIT 2013 checklist – 3. Methods

## 21. Data monitoring committee (DMC)

### Example

---

#### **“Premature termination of the study**

An interim-analysis is performed on the primary endpoint when 50% of patients have been randomised and have completed the 6 months follow-up. The interim-analysis is performed by an independent statistician, blinded for the treatment allocation. The statistician will report to the independent DSMC [data and safety monitoring committee]. The DSMC will have unblinded access to all data and will discuss the results of the interim-analysis with the steering committee in a joint meeting. The steering committee decides on the continuation of the trial and will report to the central ethics committee. The Peto approach is used: the trial will be ended using symmetric stopping boundaries at  $P < 0.001$  [reference]. The trial will not be stopped in case of futility, unless the DSMC during the course of safety monitoring advises [sic] otherwise. In this case DSMC will discuss potential stopping for futility with the trial steering committee.”<sup>332</sup>

# SPIRIT 2013 checklist – 3. Methods

## 22-23 Eventi avversi e Auditing

22. Necessario rilevare sistematicamente tutti gli eventi avversi che si verificano durante il trial, indipendentemente dalla causa:

- Distinzione tra eventi avversi seri prevedibili (SAE) o non prevedibili (SUSAR)
- Deve essere prevista una organizzazione per assicurare la segnalazione tempestiva degli eventi avversi alle autorità e agli investigatori

23. Auditing:

- L'auditing indipendente di un trial è previsto dalle GCP guidelines e da diverse agenzie
- Ha lo scopo di verificare il corretto svolgimento dello studio, il rispetto del protocollo e la qualità, anche formale, delle procedure e della documentazione (es. consenso informato firmato, altri documenti)
- Può essere svolto con “site visits” a campione o secondo un approccio basato sul rischio (centri con elevato arruolamento, altri criteri)

# SPIRIT 2013 checklist – 3. Methods

## 24-25 Approvazione CE. emendamenti

24. Approvazione da parte di comitato etico o IRB

- Nonostante l'approvazione di un trial sia prevista per legge, revisioni di studi hanno verificato che ciò non avviene per la totalità dei trials
- Nel protocollo deve essere chiaramente indicata l'approvazione ricevuta dal comitato etico

25. La necessità di modificare un protocollo dopo l'approvazione di un CE è frequente (circa il 50% degli studi) e la procedura deve essere prevista

- Gli emendamenti sostanziali devono essere approvati da CE e comportano impegni gravosi
- Necessario rivedere con attenzione il protocollo per prevenire alcune esigenze di emendamento
- Gli emendamenti su aspetti importanti del disegno, specie se suggeriti dai risultati in itinere, possono essere fonte di bias
- Le versioni precedenti del protocollo devono restare disponibili
- Nelle pubblicazioni è necessario descrivere correttamente gli eventuali emendamenti del protocollo

# SPIRIT 2013 checklist – 3. Methods

## 24-25 Approvazione CE. emendamenti

### Example

---

#### **“13.10 Modification of the Protocol**

Any modifications to the protocol which may impact on the conduct of the study, potential benefit of the patient or may affect patient safety, including changes of study objectives, study design, patient population, sample sizes, study procedures, or significant administrative aspects will require a formal amendment to the protocol. Such amendment will be agreed upon by BCIRG [Breast Cancer International Research Group] and Aventis, and approved by the Ethics Committee/IRB [institutional review board] prior to implementation and notified to the health authorities in accordance with local regulations.

Administrative changes of the protocol are minor corrections and/or clarifications that have no effect on the way the study is to be conducted. These administrative changes will be agreed upon by BCIRG and Aventis, and will be documented in a memorandum. The Ethics Committee/IRB may be notified of administrative changes at the discretion of BCIRG.”<sup>345</sup>

# SPIRIT 2013 checklist – 3. Methods

## 26-27 Consenso informato, Confidenzialità

26a. La **procedura e le modalità** per informare i partecipanti e **per la richiesta del consenso** devono essere descritte nel protocollo e la modulistica usata deve farne parte integrante

- Le informazioni scritte devono essere complete e comprensibili
- Il soggetto deve poter ricevere tutte le informazioni e i chiarimenti aggiuntivi
- Devono essere previsti i casi nei quali il consenso va richiesto ad altre figure
- Nei trials con disegno a cluster deve essere dichiarata e giustificata l'impossibilità di richiesta del consenso individuale

26b. Per la **raccolta di campioni biologici** e per gli studi ancillari deve essere prevista una **procedura specifica per la richiesta del consenso**

27. Il protocollo deve specificare come viene **garantita la confidenzialità** dei dati raccolti e descrivere adeguate misure di protezione e sicurezza dei dati

# SPIRIT 2013 checklist – 3. Methods

## 28-30. Dichiarazione di interessi, accesso ai dati e alle cure

28. Il PI e gli investigatori responsabili dello studio dovrebbero dichiarare nel protocollo l'esistenza di potenziali conflitti di interesse

- La presenza di conflitti di interesse non è necessariamente causa di distorsioni, ma deve essere dichiarata (e aggiornata) nel protocollo in modo trasparente, distinguendo motivi:
  - ✓ Economici
  - ✓ Professionali
  - ✓ Istituzionali

29. Accesso ai dati

- Nel protocollo deve essere specificato chi ha il completo accesso ai dati e se ci sono eventuali restrizioni

30. L'accesso alle cure ancillari per conseguenze dei trattamenti e dopo la conclusione del trial deve essere esplicitato nel protocollo:

- queste garanzie sono di particolare importanza per i partecipanti a sperimentazioni in paesi dove il diritto alle cure è limitato dal costo o dall'accessibilità

# SPIRIT 2013 checklist – 3. Methods

## 31. Regole di diffusione dei risultati

- 31a. Nel protocollo devono essere previste le regole di pubblicazione dei risultati del trial.
- Siccome solo il 50% circa degli studi viene pubblicato, e la velocità di pubblicazione è associata ai risultati, questo aspetto ha un'importanza fondamentale.
  - Se sono presenti limitazioni o restrizioni alla pubblicazione (es. da parte dello sponsor), queste devono essere dichiarate nel protocollo.
  - Il protocollo dovrebbe contenere il piano di pubblicazione dei risultati nei confronti dei diversi stakeholders, specificando anche il tempo previsto.
- 31b. Il protocollo deve specificare le regole concordate per l'authorship nelle pubblicazioni per evitare l'esclusione di soggetti che avevano contribuito allo studio (ghost authors), l'inclusione di soggetti che non soddisfano i criteri di authorship (secondo l'ICMJE) (guest authors) e il coinvolgimento di eventuali professional writers deve essere dichiarato e il loro ruolo e modalità di pagamento esplicitati
- 31c. Il protocollo dovrebbe indicare se il report con i risultati completi dello studio, il protocollo ed eventualmente i dati e i programmi di analisi saranno resi pubblici e con quali tempi

# SPIRIT 2013 checklist – 3. Methods

## 31. Regole di diffusione dei risultati

### Example

---

“**Data sharing statement** No later than 3 years after the collection of the 1-year postrandomisation interviews, we will deliver a completely deidentified data set to an appropriate data archive for sharing purposes.”<sup>393</sup>

# SPIRIT 2013 checklist – 3. Methods

## 32-33. Consenso, campioni biologici

In appendice al protocollo devono essere previste:

- copie della documentazione informativa per i partecipanti
- moduli per la raccolta del consenso informato
- Per trial che coinvolgono soggetti di diversa età e capacità di comprensione, può essere opportuno predisporre documenti informativi diversi

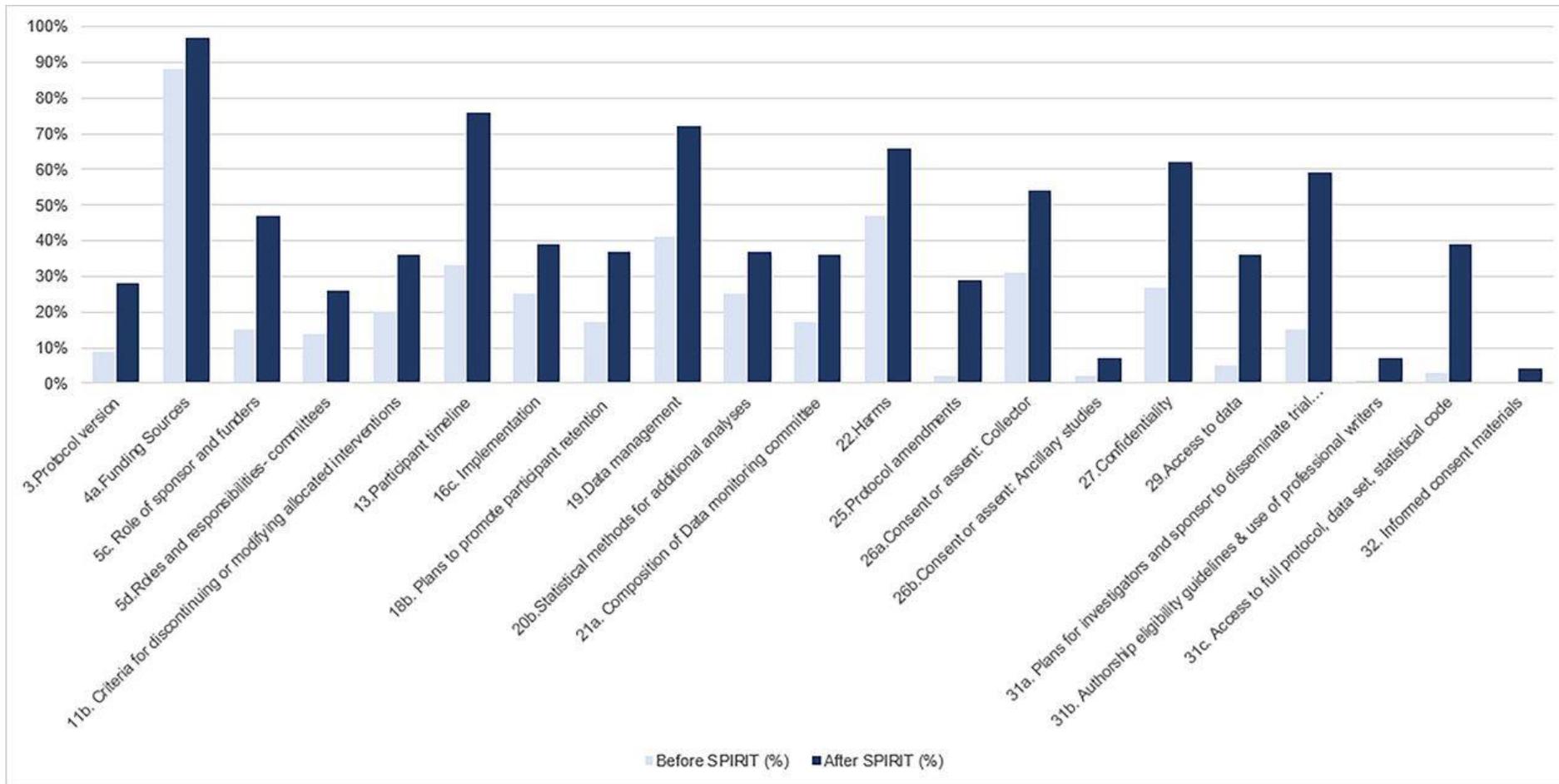
Negli studi che prevedono la raccolta e la conservazione di campioni biologici per studi ancillari o per usi futuri (previsti o non previsti) il protocollo deve specificare le modalità operative di raccolta, stoccaggio, conservazione e protezione della confidenzialità dei pazienti

**Table 2. Potential Benefits and Proposed Stakeholder Actions for Supporting Adherence to SPIRIT 2013**

Stakeholder	Proposed Actions	Potential Benefits
Clinical trial groups, investigators, sponsors	Adopt SPIRIT as standard guidance Use as tool for writing protocols	Improved quality, completeness, and consistency of protocol content Enhanced understanding of rationale and issues to consider for key protocol items Increased efficiency of protocol review
Research ethics committees/institutional review boards, funding agencies, regulatory agencies	Mandate or encourage adherence to SPIRIT for submitted protocols Use as training tool	Improved quality, completeness, and consistency of protocol submissions Increased efficiency of review and reduction in queries about protocol requirements
Educators	Use SPIRIT checklist and explanatory paper as a training tool	Enhanced understanding of the rationale and issues to consider for key protocol items
Patients, trial participants, policymakers	Advocate use of SPIRIT by trial investigators and sponsors	Improved protocol content relevant to transparency, accountability, critical appraisal, and oversight
Trial registries	Encourage SPIRIT-based protocols Register full protocols to accompany results disclosure	Improved quality of registry records Prompt for trialists to update registry record when SPIRIT checklist item 2b (Registration Data Set) is updated Improved quality, completeness, and consistency of protocol content for registries that house full protocols and results
Journal editors and publishers	Endorse SPIRIT as standard guidance for published and unpublished protocols Include reference to SPIRIT in instructions for authors Ask that protocols be submitted with manuscripts, circulate them to peer reviewers, and encourage authors to make them available as Web appendices	Improved quality, completeness, and consistency of protocol content Enhanced peer review of trial manuscripts through improved protocol content, which can be used to assess protocol adherence and selective reporting Improved transparency and interpretation of trials by readers

SPIRIT = Standard Protocol Items: Recommendations for Interventional Trials.

# Has the reporting quality of published randomised controlled trial protocols improved since the SPIRIT statement? A methodological study (Tan, BMJ Open 2020)



Checklist items with a significant increase in adequate reporting after the SPIRIT statement.