

I PROs in oncologia: dall'esperienza alla letteratura scientifica, le indicazioni per un corretto impiego dello strumento

Sono molto lieto di intervenire sull'uso dei PROs in oncologia, dal momento che nel corso degli ultimi anni ho coordinato vari studi clinici su farmaci anti-neoplastici (non terapie di supporto) in cui un PRO rappresentava l'*endpoint* primario o il *co-primary endpoint*. Penso ad esempio allo studio ELVIS sulla vinorelbina nei tumori del polmone dell'anziano⁽¹⁾, allo studio GEMVIN3 sulla chemioterapia senza platino nei tumori del polmone dell'adulto⁽²⁾; e poi a tre studi che confrontano modalità diverse di somministrazione degli stessi farmaci, settimanale contro tri-settimanale⁽³⁻⁵⁾.

Il mio intervento sarà strutturato in tre parti:

- a) quello che ho imparato coordinando studi non profit in oncologia;
- b) quello che mi ha colpito spulciando la letteratura;
- c) le riflessioni che ho fatto alla luce di quanto ho imparato e di quanto mi ha colpito (come dire $c=a+b+\text{sensò_critico}$).

Premettendo che non a tutte le domande sul tappeto mi sentirò di dare una risposta, men che meno definitiva.

Partiamo dagli studi che ho avuto l'onore di coordinare negli ultimi anni e che hanno rappresentato per me l'occasione per imparare molte cose.

Nello studio ELVIS⁽¹⁾ ci proponevamo di valutare se la vinorelbina migliorava la qualità di vita (QdV) di pazienti anziani affetti da un tumore del polmone non a piccole cellule (NSCLC). In realtà, oltre a trovare dei benefici di QdV, scoprimmo (e sicuramente rappresentò una fortuna per la popolarità dello studio) che la vinorelbina prolungava la sopravvivenza dei pazienti. Ma trovare una differenza di sopravvivenza significa che in un braccio (quello di controllo) i pazienti muoiono prima e questo condiziona notevol-

mente la compliance ai questionari di QdV. La lezione che ho imparato, in questo caso, è che i questionari *missing* non sono tali per caso, capitano soprattutto nel braccio dove si sta peggio, nel braccio dove si muore prima.

Nella *Figura 1*, la parte nera delle colonnine che progressivamente cresce, andando dal primo al sesto ciclo, è molto più rilevante nel braccio di controllo (segnalato dalla lettera C) che nel braccio con vinorelbina (V). Questa cosa non è insignificante, perché, come dimostrammo nel lavoro pubblicato su JNCI (*Figura 2*), i pazienti che non compileranno il prossimo questionario hanno degli score di QdV (in questo caso il punteggio nelle scale 29-30 del QLQC30 EORTC⁽⁶⁾) molto più bassi di quelli che invece andranno avanti. In fondo si tratta di un fenomeno del tutto prevedibile e ragionevole, che però rappresenta un notevole problema metodologico per la distorsione che si può avere nella analisi di QdV in uno studio in cui vi sia una differenza di sopravvivenza tra i bracci di confronto.

Nello studio GEMVIN3⁽²⁾ confrontavamo una chemioterapia con platino contro una chemioterapia senza platino nel NSCLC. L'assunto

Figura 1 - Risposte ai questionari sulla qualità di vita in base al trattamento (e alla sopravvivenza)

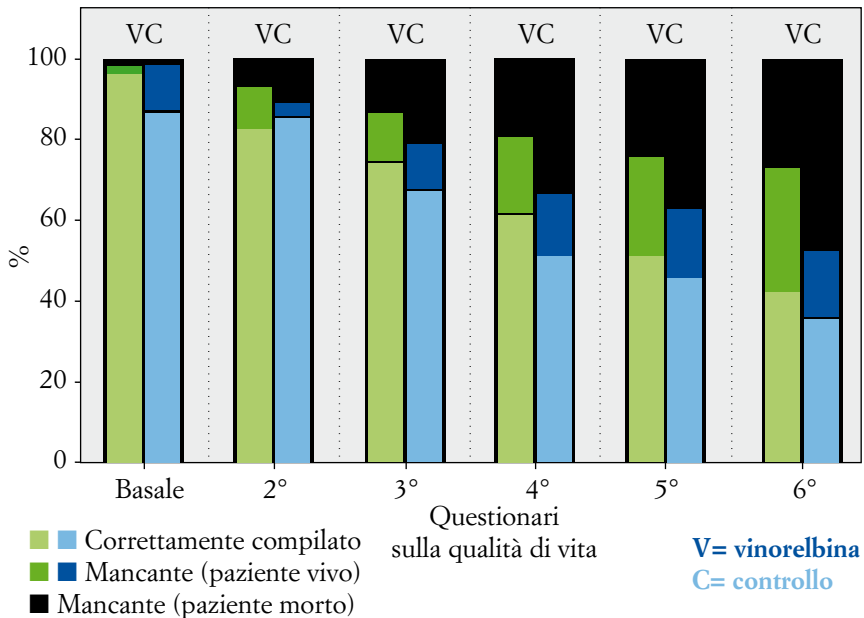
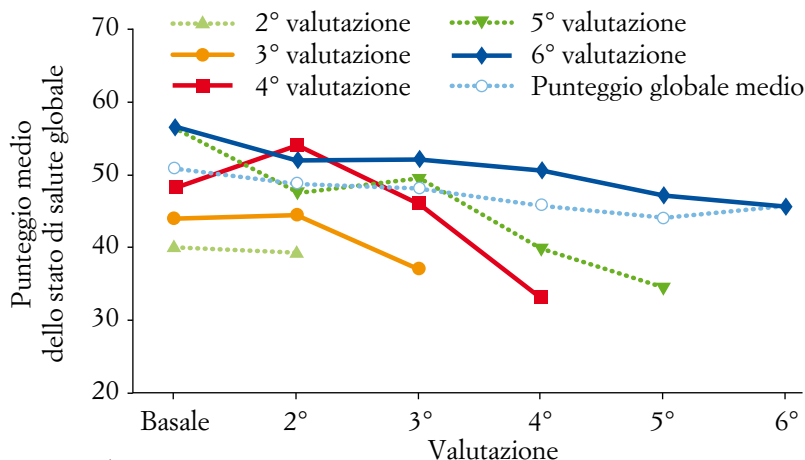


Figura 2 - Compliance alla valutazione della qualità di vita in base al trattamento



Questionari completati

Vinorelbina (n=78)#	75	65	58	48	40	33
%	96.1	82.9	75.0	61.8	51.3	42.1
Controllo (n=76)#	66	65	52	39	35	27
%	87.2	85.9	67.9	51.3	46.1	35.9

Punteggio medio dello stato di salute globale (globale e in base al tempo di uscita).

La compliance dei pazienti (%) alla valutazione della qualità di vita in base al trattamento è mostrato in fondo

mod. da J Natl Cancer Inst 1999

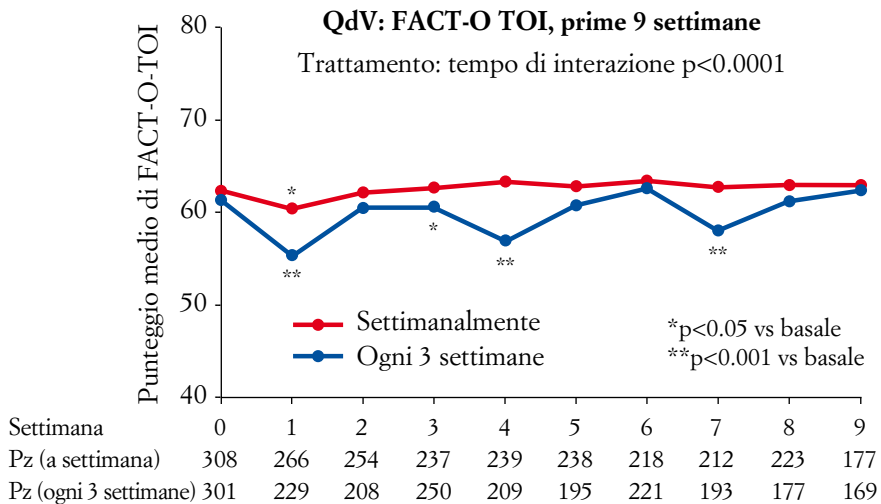
alla base dello studio era che siccome la chemioterapia con platino avrebbe dovuto essere più tossica di quella senza platino (gemcitabina + vinorelbina), quest'ultima avrebbe potuto migliorare la QdV o almeno consentire che essa peggiorasse di meno. L'assunto si confermò vero, cioè la chemioterapia con platino risultò molto più tossica di quella senza platino, ma questo non si tradusse assolutamente in qualcosa di comprensibile con le scale 29 e 30 del questionario EORTC, quelle che tendono a dare una stima generale della QdV. In realtà, qualche segnale lo si poteva cogliere solo guardando uno dei numerosi *items* singoli, che sono pure presenti nel questionario. La lezione che ho imparato in questo caso è che la relazione inversa tra tossicità e QdV non si può dare per scontata e comunque può non essere colta dagli indicatori complessivi, ma solo dai singoli *items*.

Gli studi di confronto tra schemi settimanali e tri-settimanali ci hanno da subito fatto interrogare su quale fosse lo strumento più ragionevole da utilizzare. Nello studio DISTAL-1⁽³⁾ e nello studio BREAST-10⁽⁴⁾

abbiamo utilizzato una Daily Diary Card⁽⁷⁾, che il paziente compilava a casa, giorno per giorno, perché abbiamo considerato necessario uno strumento sensibile a quello che accadeva quotidianamente; un classico questionario da compilare solo nel momento in cui si accedeva all'ospedale poteva non funzionare; inoltre, in una terapia settimanale il paziente accede all'ospedale tutte le settimane, mentre in una tri-settimanale accede solo ogni 3 settimane e, pertanto, il rischio di perdere dei questionari settimanali nel caso di terapia tri-settimanale è molto alto. Ci siamo così ritrovati ad analizzare diversi strumenti, con un numero notevole di items. La lezione che ho imparato è che la molteplicità di *items* e strumenti produce una notevole difficoltà di comprensione dei risultati, poiché può accadere (spesso) che un trattamento vinca in alcuni items ma perda in altri e che in altri ancora non vi siano differenze. Come comprendere risultati di questo tipo? Come tradurli in uno strumento che aiuti la decisione clinica?

Nello studio MITO7⁽⁵⁾, dedicato alle pazienti con tumore dell'ovaio, abbiamo confrontato lo schema carboplatino + paclitaxel settimanale verso tri-settimanale, utilizzando il questionario FACT-O⁽⁸⁾. Si può vedere

Figura 3 - Qualità di vita in pazienti in trattamento con carboplatino + paclitaxel: schema settimanale vs tri-settimanale



In tutte le scale, valori più elevati rappresentano outcome migliori
Tutti i test sono aggiustati per performance status, stadio, malattia residua dopo l'intervento chirurgico, categoria di età e dimensione dell'istituto

da Pignata S. ASCO 2013

(Figura 3) un chiaro peggioramento della QdV con il trattamento tri-settimanale, una settimana dopo la somministrazione della chemioterapia. E questo si ripete in maniera coerente al primo ciclo, al secondo ciclo, al terzo ciclo. Mentre con il trattamento settimanale, se si esclude un modesto peggioramento dopo la prima settimana (la chemioterapia non fa piacere a nessuno...), ci si abitua e la QdV non cambia più nel tempo. La lezione da imparare qui è che ci possono essere delle differenze di QdV anche rilevanti e umanamente significative che però sono transitorie. E se non si usa uno strumento adeguato si corre il rischio di non identificarle e di trarre conclusioni erranee.

Riassumendo, le lezioni che ho appreso dall'esperienza fatta come coordinatore di studi non profit in cui la QdV era un endpoint rilevante sono:

- i dati missing esistono (non è un gioco di parole!), sono informativi, e rappresentano un notevole problema metodologico in particolar modo quando ci troviamo in una condizione di prognosi molto negativa, in cui lo scadimento delle condizioni del paziente può avvenire durante la terapia;
- la relazione inversa, apparentemente molto ragionevole, tra tossicità e QdV non è scontata e comunque potrebbe non essere colta dagli indicatori complessivi, ma solo dai singoli items;
- la molteplicità di indicatori e strumenti necessaria per avere buona qualità metodologica pone notevoli problemi di interpretazione a causa della possibile incoerenza interna dei risultati;
- alcuni effetti sulla QdV, pur importanti dal punto di vista del paziente, possono essere transitori.

E il messaggio che ho personalmente portato a casa è che si può, anzi si deve, utilizzare la QdV come obiettivo degli studi clinici, ma bisogna sapersi guardare dalle trappole e dalle distorsioni metodologiche.

Spulciando la letteratura, mi ha molto colpito leggere una revisione su quello che è accaduto nella registrazione di nuovi farmaci dalla FDA dal 2006 al 2010⁽⁹⁾. In questo periodo, la branca della medicina in cui la FDA ha registrato il più alto numero di farmaci è l'oncologia (16 farmaci) e per questi 16 farmaci non vi è alcun riferimento a un PRO nel foglietto illustrativo. Da un punto di vista concettuale, questo fatto rappresenta, a mio avviso, un tradimento di quello che l'American Society of Clinical Oncology (ASCO) ha definito nel 1996⁽¹⁰⁾, e cioè che la QdV rappresenta uno dei tre cardini di valutazione dell'efficacia (ovvero del beneficio per i pazienti) di un trattamento, insieme alla sopravvivenza e alla tossicità.

Ma la revisione mi ha colpito anche per il fatto che uno dei co-au-

tori è dipendente di una multinazionale del farmaco. Ho fatto qualche altra ricerca bibliografica e mi sono reso conto che molti articoli sui PRO vedono tra i co-autori dipendenti delle maggiori aziende farmaceutiche multinazionali. Mi sono anche reso conto che vi è una crescente attenzione, in ottica regolatoria, su problemi metodologici di sicura rilevanza. Per esempio, molti lavori affrontano il problema del definire che cos'è una differenza minima importante in un PRO. Scherzando si può dire che, mentre tutti noi comprendiamo “mi sento bene” e “non mi sento bene” e non abbiamo neanche bisogno di capire “vivo” e “morto”, può essere veramente complicato capire che cosa voglia dire che un trattamento peggiora la stanchezza di 1,5 o la nausea di 2,4, sebbene i numeri rappresentino il codice dotato di massima oggettività (dopo il decesso). Ed è un problema estremamente complicato, che va ben oltre una banale disputa di codici e di linguaggio. David Cella, ad esempio, rispettato metodologo in questo campo, ha dato già negli scorsi anni un notevole contributo su quali siano gli strumenti per valutare le differenze rilevanti, *item per item*, perché al di là di un *framework* di analisi che può essere generalizzato non esiste una regola magica che valga per tutti gli *items*⁽¹⁾. Inoltre, non è detto che quello che si definisce possa valere in contesti culturali differenti, il che crea un ulteriore rilevante livello di complessità e la necessità di considerare la qualità delle traduzioni tra gli elementi minimi di qualità per la valutazione dei PRO⁽¹²⁾. Eppure, tanta colta attenzione ai PROs ha fino ad oggi un impatto regolatorio pari praticamente a zero (pur con l'eccezione della registrazione di ruxolitinib nella mielofibrosi).

D'altra parte, l'attenzione a definire la rilevanza clinica dei risultati dei *trials* clinici vede in questi mesi un fatto del tutto nuovo, che va anche molto oltre il problema dei PROs. Solo oggi, dopo circa 30 anni di studi clinici che hanno la sopravvivenza o il *progression free survival* come *endpoint* principale, l'ASCO affronta lo spinoso tema riassumibile nella domanda: “Quanta differenza di sopravvivenza ci deve essere perché ne valga la pena?”. Siamo per ora solo a una bozza di pubblicazione (Ellis LM, Bernstein D, Voest E et al. The ASCO Clinically Meaningful Outcomes Working Groups. Defining clinically meaningful outcomes: ASCO Recommendations to raise the bar for clinical trials) con delle indicazioni che personalmente condivido ma che spostano molto in alto la cosiddetta *barra dell'efficacia*, facendo dire a molti oncologi che se queste indicazioni dovessero diventare prescrittive a fini regolatori si ridurrebbe notevolmente l'arrivo di nuovi farmaci.

Volendo riassumere quello che mi ha colpito spulciando la letteratura potrei dire che:

- nei fatti, ai fini regolatori, il ruolo dei PROs in oncologia è ancora, a voler essere ottimisti, complementare (che non sarebbe in fondo sbagliato), ma molte volte del tutto secondario o ignorato (che il più delle volte è sicuramente sbagliato);
- l'interesse delle aziende farmaceutiche tuttavia è crescente e va in parallelo (o probabilmente agisce come propulsore) con l'attenzione, per ora solo teorica, delle agenzie regolatorie;
- la discussione metodologica è matura ma piena di trappole; si gioca a livelli molto tecnici, soprattutto per quanto riguarda il valore dell'outcome (si veda per esempio l'intenso dibattito sulla differenza minima importante).

Altri lavori recentemente pubblicati possono essere utili per alcune riflessioni. Nel 2012 viene pubblicato, sul *NEJM*, lo studio registrativo che dimostra come abiraterone prolunghi notevolmente (praticamente raddoppi in termini di mediana) la sopravvivenza libera da progressione radiologica dei pazienti con un tumore della prostata che è diventato resistente alla terapia endocrina⁽¹³⁾; anche la sopravvivenza in questo studio appare prolungata, sebbene tale analisi ancora immatura non abbia raggiunto la significatività statistica ai limiti richiesti dallo schema di analisi ad interim. Più recentemente è stato pubblicato su *Lancet Oncology* il lavoro che dimostra che i vantaggi evidenziati nello studio registrativo si accompagnano a un rinvio nel tempo del momento in cui si ha la progressione del dolore, o il peggioramento delle capacità funzionali, sulla base di questionari compilati dai pazienti⁽¹⁴⁾.

Devo dire, tutto sommato, che in presenza di vantaggi oggettivi molto chiari, quali quelli mostrati dallo studio registrativo, la qualità di vita potrebbe anche essere di ridotto interesse. In linea di massima, ritengo che laddove si riesca a prolungare in maniera rilevante un *endpoint* di efficacia come la sopravvivenza, la qualità di vita sia relativamente poco importante, tranne il caso improbabile di un farmaco che possa drammaticamente peggiorarla e che si debba immaginare di porre un paziente di fronte a una terribile domanda "Vuole vivere più a lungo ma peggio piuttosto che...?". Ma un caso del genere è più teorico che pratico, visto che è molto improbabile che un farmaco che peggiori drammaticamente la qualità di vita faccia vivere molto più a lungo.

Al contrario, la qualità di vita diventa cruciale nel caso, molto frequente, in cui confrontiamo due farmaci che tra di loro non sono particolarmente diversi in termini di effetto della sopravvivenza. È il caso dello studio COMPARZ⁽¹⁵⁾, in cui considero molto tranquillizzante sul piano della non inferiorità la curva di sopravvivenza globale, riportata nell'appendice. E la dimostra-

zione che la qualità di vita diventi importante è dimostrata dal fatto che il *NEJM* ha invece riservato spazio nella pubblicazione principale per una tabella che riassume l'outcome nei vari items di qualità di vita e per due figure che riportano una scala di misura della *fatigue*, e una di misura del dolore ai piedi. Questa scelta, in particolare quella delle figure che riguardano due indicatori che allo stesso tempo sono considerati effetti collaterali del trattamento e segnali di qualità della vita, sottolinea l'opportunità di integrare la valutazione del medico e del paziente per alcuni effetti collaterali, come accade nei PRO-CTCAE recentemente proposti. Nella tabella di tossicità i medici hanno riportato una percentuale di pazienti che va incontro a gradi elevati di sindrome mano-piede più alta con sunitinib (12%) rispetto a pazopanib (6%), e grazie alle risposte date dai pazienti su questo stesso problema sappiamo che questo effetto collaterale rappresenta un problema rilevante per la qualità di vita dei pazienti. Per inciso, una intera sessione dell'ASCO 2013 è stata dedicata ai PRO-CTCAE (<https://wiki.nci.nih.gov/pages/viewpage.action?pageId=10857328>).

L'ultima riflessione riguarda un'altra coppia di lavori pubblicati recentemente. La materia è il confronto tra la formulazione endovenosa di trastuzumab e quella somministrabile per via sottocutanea. Un primo lavoro, pubblicato nel 2012⁽¹⁶⁾, dimostra che l'efficacia del farmaco è uguale con le due vie di somministrazione, nel setting di terapia neoadiuvante del tumore della mammella. Il secondo lavoro, più recente⁽¹⁷⁾, indaga la preferenza delle pazienti, in un disegno di studio in cui la paziente, a random, faceva prima l'endovena e poi il sottocute o, viceversa, prima il sottocute e poi l'endovena. Semplificando, alla fine si chiedeva ad ogni paziente "Quale preferisce?". E oltre il 90% delle pazienti preferiva la somministrazione sottocutanea. Quanto possiamo considerare rilevante questo tipo di studi e questo tipo di risultati? Quanto corrispondono a reali bisogni non soddisfatti dei pazienti e quanto piuttosto a strategie di ricerca dettate da ragioni di mercato, seppur virtuosamente associate alla attenzione al punto di vista dei pazienti?

E veniamo alle mie riflessioni finali, riassumibili con la metafora del cadere dalla padella nella brace.

La padella è rappresentata dall'attenzione complessivamente ancora troppo scarsa alla qualità di vita e ai Patient-Reported Outcomes in oncologia. Sono molti i casi in cui l'impatto del trattamento o, più precisamente, il valore aggiunto dei nuovi farmaci è marginale; non prestare sufficiente attenzione e sensibilità alla qualità di vita dei pazienti in queste condizioni rappresenta a mio avviso un errore che si dovrebbe e potrebbe evitare. In contesti non regolatori è più facile che si presti attenzione a questi aspetti, ma sebbene la ricerca non a fini registrativi debba essere

orientata al miglioramento della pratica clinica, il suo impatto reale non è necessariamente garantito. E ancora, la padella è caratterizzata dalle notevoli difficoltà che si incontrano nel fare studi con indicatori di QdV e PRO; difficoltà intrinseche alla materia e agli strumenti di studio e che richiedono una enorme capacità critica e metodologica per trasformare opportunamente i risultati di studi di questo tipo in una migliore pratica clinica e una più onesta e completa informazione dei pazienti.

La brace, invece, è rappresentata a mio modo di vedere dal rischio che i Patient-Reported Outcomes diventino il nuovo campo di battaglia in cui si confronteranno le aziende farmaceutiche tra di loro o con gli enti pagatori, sotto lo sguardo pilatesco delle agenzie regolatorie che a mio modo di vedere sbagliano ad alimentare la convinzione che si possa raggiungere precisione e affidabilità, in una materia che per definizione tutto è tranne che precisa e affidabile. È del tutto giusto, come sottolineato da Arpinelli e Banfi alcuni anni fa ⁽¹⁸⁾, che le aziende farmaceutiche colgano i vantaggi e le opportunità derivanti dai Patient-Reported Outcomes nello sviluppo di nuovi farmaci. Sia per rendere i propri farmaci competitivi rispetto ad altri della stessa classe farmacologica, sia per puntare a condizioni di rimborsabilità più redditizie nel caso in cui si possano dimostrare significativi vantaggi per i pazienti. Tuttavia, vedo con preoccupazione l'ipotesi che un farmaco possa essere approvato o non approvato sulla base della considerazione che un cambiamento in un *item* di qualità di vita sia di un decimo al di sotto o al di sopra di quella che viene definita come differenza minima importante. Così come temo che ci sia un rilevante rischio di strumentalizzazione dei PRO ai fini della rimborsabilità e del convincimento della classe medica; e non posso non osservare che il rischio di strumentalizzazione è molto superiore di quanto possa essere con dati incontrovertibili come quelli di mortalità. È ovvio considerare che una più ampia cultura della misurazione soggettiva, e una maggiore considerazione per le opinioni dei pazienti rappresentino degli indiscutibili valori positivi, ma credo che non ci si possa limitare a questa lettura benevola e che la comunità scientifica debba vigilare affinché una materia di assoluta rilevanza umana ed etica non venga banalizzata e strumentalizzata.

Bibliografia

- 1) The Elderly Lung Cancer Vinorelbine Italian Study Group. Effects of vinorelbine on quality of life and survival of elderly patients with advanced non-small-cell lung cancer. *J Natl Cancer Inst* 1999; 91(1): 66-72.
- 2) Gridelli C, Gallo C, Shepherd FA et al. Gemcitabine plus vinorelbine compared with

- cisplatin plus vinorelbine or cisplatin plus gemcitabine for advanced non-small-cell lung cancer: a phase III trial of the Italian GEMVIN Investigators and the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 2003; 21(16): 3025-34.
- 3) Gridelli C, Gallo C, Di Maio M et al. A randomised clinical trial of two docetaxel regimens (weekly vs 3 week) in the second-line treatment of non-small-cell lung cancer. The DISTAL 01 study. *Br J Cancer* 2004; 91(12): 1996-2004.
 - 4) Nuzzo F, Morabito A, Gravina A et al. Effects on quality of life of weekly docetaxel-based chemotherapy in patients with locally advanced or metastatic breast cancer: results of a single-centre randomized phase 3 trial. *BMC Cancer* 2011; 11:75.
 - 5) Pignata S, Scambia G, Katsaros D et al. Carboplatin plus paclitaxel once a week versus every 3 weeks in patients with advanced ovarian cancer (MITO-7): a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncology* 2014; 15(4): 396-405.
 - 6) Aaronson NK, Ahmedzai S, Bergman B et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993; 85(5): 365-76.
 - 7) Fayers PM, Bleehen NM, Girling DJ, Stephens RJ. Assessment of quality of life in small-cell lung cancer using a Daily Diary Card developed by the Medical Research Council Lung Cancer Working Party. *Br J Cancer* 1991; 64(2): 299-306.
 - 8) Basen-Engquist K, Bodurka-Bevers D, Fitzgerald MA et al. Reliability and validity of the functional assessment of cancer therapy-ovarian. *J Clin Oncol* 2001; 19(6): 1809-17.
 - 9) Gnanasakthy A, Mordin M, Clark M et al. A review of patient-reported outcome labels in the United States: 2006 to 2010. *Value Health* 2012; 15(3): 437-42.
 - 10) Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. American Society of Clinical Oncology. *J Clin Oncol* 1996; 14(2): 671-9.
 - 11) Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61(2): 102-9.
 - 12) Reeve BB, Wyrwich KW, Wu AW et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013; 22(8): 1889-905.
 - 13) Ryan CJ, Smith MR, de Bono JS et al. Abiraterone in metastatic prostate cancer without previous chemotherapy. *N Engl J Med* 2013; 368(2): 138-48.
 - 14) Basch E, Autio K, Ryan CJ et al. Abiraterone acetate plus prednisone versus prednisone alone in chemotherapy-naïve men with metastatic castration-resistant prostate cancer: patient-reported outcome results of a randomised phase 3 trial. *Lancet Oncology* 2013; 14(12): 1193-9.
 - 15) Motzer RJ, Hutson TE, Cella D et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med* 2013; 369(8): 722-31.
 - 16) Ismael G, Hegg R, Muehlbauer S et al. Subcutaneous versus intravenous administration of (neo)adjuvant trastuzumab in patients with HER2-positive, clinical stage I-III breast cancer (HannaH study): a phase 3, open-label, multicentre, randomised trial. *Lancet Oncology* 2012; 13(9): 869-78.
 - 17) Pivot X, Gligorov J, Muller V et al. Preference for subcutaneous or intravenous administration of trastuzumab in patients with HER2-positive early breast cancer (PrefHer): an open-label randomised study. *Lancet Oncology* 2013; 14(10): 962-70.
 - 18) Arpinelli F, Banfi F. The FDA guidance for industry on PROs: the point of view of a pharmaceutical company. *Health Qual Life Outcomes* 2006; 4: 85.