EUROPEAN JOURNAL OF CANCER 47 (2011) 1468–1475

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Heterogeneity in the definition of dose-limiting toxicity in phase I cancer clinical trials of molecularly targeted agents: A review of the literature

Christophe Le Tourneau [a,*], *Albiruni R.A. Razak* [b], *Hui K. Gan* [c], *Simona Pop* [a], *Véronique Diéras* [a], *Patricia Tresca* [a], *Xavier Paoletti* [d]

[a] *Department of Medical Oncology, Institut Curie, Paris, France*
[b] *Department of Medical Oncology, Princess Margaret Hospital, Toronto, Canada*
[c] *Austin-Ludwig Cancer Unit, Austin Hospital, Melbourne, Australia*
[d] *Department of Biostatistics, INSERM U900, Institut Curie, Paris, France*

ARTICLE INFO

Article history:
Received 4 February 2011
Received in revised form 14 March 2011
Accepted 14 March 2011
Available online 7 April 2011

Keywords:
Phase I trials
Molecularly targeted agents
Dose-limiting toxicity
Definition
NCI CTCAE
Duration of toxicity
Reversibility

ABSTRACT

Aim: There is no consensus about what constitutes a dose-limiting toxicity (DLT) in phase I cancer clinical trials. We aimed to evaluate how DLTs are defined in phase I trials of molecularly targeted agents (MTA).

Methods: We retrieved all phase I trials testing monotherapy with an MTA published over the last decade. In each trial, all items used to define DLTs were recorded.

Results: Reports of 155 phase I trials evaluating 111 different MTAs were reviewed. The most frequent determinant of whether a toxicity was regarded as a DLT was severity, usually assessed using the NCI CTCAE classification. However, for any given toxicity, there was substantial variability in the degree of severity required for a toxicity to be considered a DLT. Specifications about minimum duration of toxicity, degree of reversibility, the need to delay treatment and to reduce dose-intensity because of toxicity were infrequently incorporated in the definition of DLT. The definition of DLT varied with administration schedule. Discrepancies between the initial and the final definition of DLT were reported in 25% of trials.

Conclusions: While our results do not support a standardisation of the definition of DLT, the inclusion of following specifications in its definition when relevant would reduce the heterogeneity observed across trials: (1) DLT assessment period, (2) absolute severity according to NCI CTCAE classification as well as severity relative to baseline status, (3) minimum duration of toxicity, (4) reversibility of toxicity within a certain period of time, and (5) necessity to delay treatment or to reduce dose-intensity.

1. Introduction

The assessment of dose-limiting toxicity (DLT) is critical in phase I cancer clinical trials, as it helps determine the maximum tolerated dose (MTD) and the recommended dose for phase II trials (RP2D), which usually represent the primary endpoints of these trials. While authorities such as the Consolidated Standards of Reporting Trials (CONSORT) committee as well as regulatory agencies such as the US Food and Drug Administration have provided guidelines for the use and

* *Corresponding author:* Address: Institut Curie, Department of Medical Oncology, 26, rue d'Ulm, 75248 Paris Cedex 05, France. Tel.: +33 1 4432 4086; fax: +33 1 5310 4006.
E-mail address: Christophe.LeTourneau@curie.net (C. Le Tourneau).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.03.016

reporting of endpoints in phase II/III randomised trials,[1–3] there is no consensus on how DLTs and MTD should be defined in phase I trials. The question about whether the definition of DLT should be standardised has previously been raised for cytotoxic drugs.[4] For example, there is often debate about the duration of grade 4 neutropenia that is required before this should be considered a study DLT.[4] Indeed, previous discussions about appropriate definitions of DLTs have focused on haematologic toxicity as most of the agents under evaluation at that time were cytotoxic agents. Despite these discussions, no specific guidelines have been published on this matter.

These complexities have increased further now that molecularly targeted agents (MTA) have become an important part of the therapeutic armamentarium. They present a challenge as they differ in significant ways from cytotoxic agents. MTAs usually have fewer haematologic toxicities, but this is balanced by a greater incidence and range of non-haematologic toxicities.[5] Determining DLTs by virtue of non-haematologic toxicities is more challenging, partly because they were usually eclipsed by haematological toxicities in earlier phase I trials with cytotoxic treatments and partly because lower grades of non-haematological toxicities may ultimately be intolerable when administered continuously, as is often the case for MTAs. Lastly, the common practice of determining the MTD and the RP2D based on toxicities may not be as relevant for MTAs as these may not need to be administered at their MTD to be active.[6] It remains to be determined whether standardised sets of DLTs should be used or whether their use should vary depending on drug characteristics such as schedule, administration route, putative target, etc.

In this study, we aimed to evaluate how DLTs have been defined in oncology phase I trials of MTA monotherapy published over the last decade.

2. Methods

2.1. Selection of trials

MTAs were defined in our study as anticancer agents that selectively target molecular pathways, as opposed to DNA, tubulin or cell division machinery.[5] Only trials administering MTAs orally or intravenously were included. Hormonal therapy and biological therapeutics such as immunotherapy, gene therapy and vaccines were excluded because of their unique mechanisms of action and toxicities (monoclonal antibodies were not excluded). Trials were also excluded if they were ongoing, testing drug combinations, reported only in abstract format, published before 2000 or in languages other than English.

2.2. Search strategy

To comprehensively identify phase I trials of MTAs, we searched SCOPUS database from January 1st, 2000 to April 18th, 2010 using appropriate search terms (Appendix, online only). To ensure that all relevant trials were identified, the bibliographies of selected papers were also manually searched for relevant publications.

2.3. Data extraction and analysis

The following data were extracted from each publication: agent's mechanism of action, administration route, dosing schedule including any non-treatment days between cycles (the recovery period), dose escalation method, DLT proportion threshold specified to define MTD, the toxicity grading scale used to assess toxicity, and DLTs encountered.

For each trial, all items used to define DLT in the Patients and Methods section were recorded (e.g. "neutropenia", "creatinine elevation", etc.). For each item, the following information was also extracted: severity of toxicity required for it to be considered a DLT, minimum duration required for it to be considered a DLT, and the degree of reversibility upon cessation of study drug. For each trial, it was also determined whether the necessity to delay treatment or to reduce dose-intensity due to toxicity was incorporated into the definition of DLT.

Each MTA was categorised as follows based on its presumed target: intracytoplasmic, intranuclear, vasculature, cellular receptor, and other. Dosing schedule was arbitrarily divided into four categories depending on the number of dosing days per cycle: up to 33%, between 33% and 66%, between 66% and 99%, and 100%. Recovery time before next cycle was dichotomised into $\leqslant$1 week *versus* >1 week.

To assess for changes over time, trials were divided into two sets based on publication date (2000–2005 *versus* 2006–2010).

To explore potential reasons for variations in the definition of DLT, characteristics of the trials and the MTAs were compared using χ^2 or exact Fischer tests at the 5% level. Means were compared using a Student *t* test.

3. Results

3.1. General results

Amongst 2, 151 references identified through the SCOPUS database, 131 fulfilled the pre-specified inclusion criteria. Review of these 131 references identified 24 additional trials that had been missed by electronic searching, usually because their title did not identify them as cancer trials. In total, 155 trials evaluating 111 MTAs were identified. Eighty-six trials were published during the 2000–2005 period and 69 trials during the 2006–2010 period.

Characteristics of the trials reviewed are presented in Table 1. The most commonly used toxicity grading scale was the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE). The DLT proportion threshold specified to define the MTD was 33% in 131 of the 144 trials where it was specified.

Characteristics of the MTAs are presented in Table 2. In the 50 studies in which the drug treatment was administered continually, the median cycle length was 28 d [range: 14–28].

3.2. Items used to define dose-limiting toxicity

DLT was explicitly defined in 148 of the 155 trials (95%) overall. It was defined in 99% of the trials published during the 2006–2010 timeframe *versus* 93% of the trials published during the 2000–2005 timeframe ($p = 0.80$). The DLT assessment period

Table 1 – Characteristics of the phase I cancer clinical trials reviewed.

	No.	%	Median	Range
Type of trial				
All tumours	137	88		
Tumour-specific	10	6		
Target-specific	4	3		
Tumour- and target-specific	1	1		
Hepatic and/or renal dysfunction	3	2		
Origin of patients				
Occidental countries				
USA	83	54		
Europe	40	26		
Canada	3	2		
Australia	1	1		
Mixed	16	10		
Non-occidental countries				
Japan	10	6		
Asia (excluding Japan)	2	1		
First-in-human trial				
Yes	111	72		
No	44	28		
Dose escalation method				
"3+3" design	96	62		
Accelerated titration design	46	30		
Continual reassessment method	8	5		
Pharmacologically guided dose escalation	1	1		
Not specified	4	3		
DLT proportion threshold specified to define MTD				
Yes	144	93	33%	[20–50]
No	11	7		
Toxicity grading scale				
NCI CTCAE v1.0	39	25		
NCI CTCAE v2.0	76	49		
NCI CTCAE v3.0	29	19		
not specified	4	3		
MTD reached				
Yes	120	77		
No	35	23		

Table 2 – Characteristics of the molecularly targeted agents evaluated in the trials reviewed.

	No.	%
Target(s)		
Vasculature	41	26
Intracytoplasmic	32	21
Cell surface receptor	30	19
Intranuclear	19	12
Other	33	21
FDA-approved drugs		
Yes	24	15
No	131	85
Route of administration		
Oral	77	50
Intravenous	78	50
Schedule		
Continuous schedule (100% of dose days)	50	32
Between 66% and 99% of dosing days per cycle e.g. drug given for 2 weeks out of every 3 weeks	15	10
Between 33% and 66% of dosing days per cycle e.g. drug given for 1 week out of every 3 weeks	7	5
Up to 33% of dosing days per cycle	77	50
Not applicable[a]	6	4
Number of non-treatment days between cycles		
⩽1 week	53	34
>1 week	96	62
Not applicable[a]	6	4

[a] Different schedules evaluated.

was defined for 142 of the 155 trials (92%) and did not vary significantly by publication timeframe. The median DLT assessment period was 28 d [range: 7–56].

A total of 825 items were used to define DLTs in these 155 trials, with 309 related to haematologic toxicities and 516 related to non-haematologic toxicities (median of five items per trial, range: 0–12).

Amongst the 155 trials, the most common category of haematologic toxicity was "haematologic toxicity, not otherwise specified (NOS)" (54%) (Table 3). Similarly, the most common category of non-haematologic toxicity was "non-haematologic toxicity, NOS". A total of eight other categories of haematological items and 46 categories of non-haematologic organ-specific items were employed to define DLTs in the 155 trials reviewed. Organ-specific items were reported in 111 trials (72%). Mean number of organ-specific items per trial increased from 2.03 in the 2000–2005 timeframe to 2.79 in the 2006–2010 timeframe ($p = 0.02$).

3.3. *Severity of toxicity by NCI CTCAE grade*

Of the 309 haematological DLT items that were defined, severity was usually defined using the NCI CTCAE grading scales: 57% of haematological DLT items were defined as those reaching NCI CTCAE grade 4 severity, 32% as those with grade 3 severity and only 1% as those reaching grade 2 severity (Table 3). In a few instances, DLT definition was based on severity criteria other than the NCI CTCAE classification (4%) or was absent entirely (4%). Of 69 studies where neutropenia was explicitly defined as being a toxicity of interest, only 6% considered grade 3 neutropenia to be a DLT whilst grade 4 neutropenia was considered a DLT in the remaining 94%. Only 46 studies explicitly defined febrile neutropenia as a DLT, although the severity of neutropenia required for this to be considered a DLT varied substantially (17 required grade 4 severity, 17 required grade 3 severity, and two required grade 2 severity). A similar pattern was seen in thrombocytopenia, with grade 3 thrombocytopenia being a DLT in 18 studies when specifically mentioned, whilst grade 4 thrombocytopenia was a DLT in 39 studies. Anaemia was less frequently considered a toxicity of interest, with only 16 studies considering grade 3/4 anaemia a DLT and three studies excluding it as a DLT regardless of severity.

Although non-haematological toxicities were considered DLTs when only reaching a grade 4 for 26 of the 516 (5%) non-haematological DLT items defined in 20 of the 155 studies (13%), it was much more common for non-haematological toxicities of lesser grade to be considered DLTs. Occurrence

Table 3 – Minimum NCI CTCAE grade used to meet DLT definition for the 825 toxicity items described in the 155 trials.

Toxicity categories	Number of items							
	⩾G4	⩾G3	⩾G2	⩾G1	Other	Not specified	Not a DLT	All
Haematologic toxicity, nos	42	42	1	0	0	0	0	85
Neutropenia	65	4	0	0	0	0	0	69
Thrombocytopenia	39	18	0	0	11	0	0	68
Febrile neutropenia	17	17	2	0	0	10	0	46
anaemia	11	5	0	0	0	0	3	19
Thrombocytopenia with bleeding	2	11	0	0	0	3	0	16
Coagulation abnormality	0	2	0	0	2	0	0	4
Lymphopenia	0	0	0	0	0	0	1	1
Haemorrhage	0	0	1	0	0	0	0	1
Haematologic toxicity (all)	176	99	4	0	13	13	4	309
Non-haematologic toxicity, nos	2	135	10	0	1	0	0	148
Gastro-intestinal symptoms								
Nausea	3	51	4	0	7	0	12	77
Vomiting	6	53	4	0	7	0	7	77
Diarrhoea	2	27	4	0	3	0	0	36
Constipation	0	1	0	0	0	0	1	2
Hepatic toxicity								
AST elevation	1	8	0	0	5	0	0	14
ALT elevation	1	8	0	0	4	0	0	13
ALP elevation	1	0	0	0	1	0	4	6
Bilirubin elevation	2	0	1	0	3	0	0	6
GGT elevation	1	0	0	0	1	0	1	3
Hepatic, nos	0	1	1	0	0	0	0	2
Renal toxicity								
Creatinine elevation	0	2	2	0	6	0	0	10
Proteinuria	0	2	2	0	0	0	0	4
Haematuria	0	0	2	0	0	0	0	2
Renal, nos	0	0	2	0	0	0	0	2
Pancreatic enzymes elevation	1	0	0	0	0	0	0	1
Biochemical abnormalities								
Electrolytes abnormalities	0	1	0	0	0	0	2	3
Triglycerides	0	0	0	0	2	0	0	2
Hyperglycaemia	0	0	0	0	1	0	0	1
Cardiac toxicity								
High blood pressure	2	6	1	0	4	0	0	13
QTc prolongation	0	3	0	0	6	0	0	9
Cardiac, nos	0	0	6	0	0	0	0	6
PR prolongation	0	0	0	0	2	1	0	3
Elevation of cardiac enzymes	0	0	0	0	0	1	0	1
Heart conduction defect	0	0	0	0	0	1	0	1
Mitral valve regurgitation	0	0	0	0	1	0	0	1
Arrythmia	0	0	0	0	0	1	0	1
Pulmonary toxicity								
Pulmonary, nos	0	0	2	0	0	1	0	3
Interstitial pneumonitis	0	0	0	0	0	1	0	1
Hypersensitivity reactions								
Infusion/hypersensitivity reaction	1	3	3	0	0	0	4	11
Interruption of infusion	0	0	0	0	1	0	0	1
Bronchospasm	0	0	1	0	0	0	0	1
Local toxicity	0	0	1	0	0	0	0	1
Skin rash	0	4	1	0	3	0	0	8
Neurotoxicity	0	0	6	1	0	1	0	8
Ophthalmologic toxicity	0	0	0	0	4	0	0	4
Other								
Fatigue	1	6	0	0	0	0	3	10
Fever	2	1	0	0	0	0	6	9
Anorexia	0	1	0	0	0	0	3	4
Tumour pain	0	1	0	0	0	0	2	3

(*Continued on next page*)

Table 3 – (*continued*)

Toxicity categories	Number of items							
	⩾G4	⩾G3	⩾G2	⩾G1	Other	Not specified	Not a DLT	All
Arthramyalgia	0	1	0	0	0	0	1	2
Myalgia	0	1	0	0	0	0	0	1
Headaches	0	1	0	0	0	0	0	1
Rhinitis	0	0	1	0	0	0	0	1
Subjectively intolerable toxicity	0	0	0	0	0	1	0	1
Hypotension	0	1	0	0	0	0	0	1
Malaise	0	0	0	0	0	0	1	1
Non-haematologic toxicity (all)	26	318	54	1	62	8	47	516

G, NCI CTCAE grade; NOS, not otherwise specified; ALT, alanine amino transferase; AST, aspartate amino transferase; GGT, gamma-glutamyltransferase; ALP, alkaline phosphatase.

of NCI CTCAE grade 3 toxicities was sufficient to constitute a DLT in 318 of the 516 non-haematologic items (62%) (Table 3). Even grade 2 non-haematological toxicities were considered a DLT in a significant minority of trials (54 of the 516 [10%] DLTs were grade 2 toxicities, being reported in 31 of the 155 trials [20%]). For 62 of the 516 items (12%), toxicities were not graded using the NCI CTCAE classification. In one trial, a 2-point decrease in ECOG performance status from baseline was considered a DLT. A substantial number of non-haematologic toxicities were also specifically excluded from being DLTs (47 of the 516 items [9%] reported in 27 of the 155 trials [17%]) (Table 3).

3.4. *Duration and reversibility of toxicity*

Regarding haematologic toxicities, persistence for a minimum duration of time was a prerequisite for neutropenia and thrombocytopenia to be considered a DLT in 42 of 69 studies (61%) and in four of 68 studies (6%), respectively. Median minimum durations to meet DLT definition were 5 d [range: 3–7], and 7 d [range: 4–14] for neutropenia and thrombocytopenia, respectively. There was a trend for minimum duration of toxicity to be incorporated more frequently into the definition of haematologic toxicity in papers published between 2006 and 2010 as compared to those published between 2000 and 2005 ($p = 0.11$).

A minimum duration of toxicity to meet non-haematologic DLT definition was defined for 35 of the 516 non-haematologic items (7%) as reported in 20 of the 155 trials (13%) (Table 4). A specification on minimum duration of toxicity was grade-specific for 21 of these 35 items (60%). Fifteen of these 21 items (71%) were specific for grade 2 toxicity. The requirements for duration of toxicity to be used in DLT definition for non-haematological toxicities did not appear to change between the two publication timeframes ($p = 0.86$).

Table 4 – Specifications of DLT items used in the 155 trials according to the type of DLT (haematologic *versus* non-haematologic) and the publication timeframe (2000–2005 *versus* 2006–2010).

	Haematologic DLT						Non-haematologic DLT					
	All trials		2000–2005		2006–2010		All trials		2000–2005		2006–2010	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
No. of items defined	309		163		146		516		255		261	
Minimum duration specified												
Yes	46	15	19	12	27	18	35	7	18	7	17	7
Duration												
<5 d	8	3	7	4	1	1	8	2	6	2	2	1
5 d	22	7	11	7	11	8	0	0	0	0	0	0
7 d	15	5	1	1	14	10	18	3	9	4	9	3
8 d	0	0	0	0	0	0	1	0.20	0	0	1	0.40
14 d	1	0.30	0	0	1	1	8	2	3	1	5	2
Grade-specific	4	1	1	1	3	2	21	4	9	4	12	5
No	263	85	144	88	119	82	481	93	237	93	244	93
Reversibility specified												
Yes	6	2	5	3	1	1	31	6	13	5	18	7
Irreversible	2	0.60	1	1	1	1	14	3	7	3	7	3
Failure to recover[a]	4	1	4	2	0	0	17	3	6	2	11	4
Grade-specific	1	0.30	1	1	0	0	8	2	7	3	1	0.40
No	303	98	158	97	145	99	485	94	242	95	243	93

DLT, dose-limiting toxicity.
[a] Within a certain period of time or by next cycle.

A specification about the reversibility of toxicity was included in the definition of DLT for six of the 309 haematologic toxicity items (2%) reported in four of the 155 trials (3%), and for 27 of the 516 non-haematologic items (5%) reported in 18 of the 155 trials (12%) (Table 4). A specification on reversibility was specific for grade 2 non-haematological toxicity in seven of the 21 items (33%).

3.5. *Treatment delay and dose-intensity reduction*

The necessity to delay treatment because of toxicity was included in the definition of DLT in 29 trials (19%) (Table 5). This characteristic was mentioned in 23% of trials published during the 2006–2010 timeframe as opposed to 15% of trials published during the 2000–2005 timeframe ($p = 0.07$). Where specified, a median delay of 14 d [range: 0–28] was the threshold to consider that toxicity a DLT.

Similarly, dose intensity was included in the definition of DLT in only 11 trials amongst the 142 trials (8%) evaluable for dose-intensity (e.g. involving more than one administration per cycle) (Table 5). This characteristic was mentioned in 10% of trials published during the 2006–2010 timeframe as opposed to 5% of trials published during the 2000–2005 timeframe ($p = 0.10$).

3.6. *Comparison of the initial and the final definition of DLT*

A total of 430 non-haematologic DLT events were reported in the 155 trials. These organ-specific DLTs were specifically defined in the "Patients and Methods section" of the paper for 69 of them (16%), corresponding to 23% of all 296 organ-specific DLTs defined in the 129 trials where non-haematologic DLTs occurred.

In 28 of the 155 trials (18%), toxicities were eventually considered to be DLTs even though they were not initially defined as such. The two main scenarios where this occurred were where study defined DLTs occurred after the specified DLT assessment period (12 trials, 43%) or where toxicities not severe enough to be considered a DLT (usually a grade 2 toxicity) were ultimately found to be intolerable (12 trials, 43%). Conversely, in 13 of the 155 trials (8%), toxicities were not reported as DLTs even though they were initially defined as such (Table A1). Overall, discrepancies between the initial and the final definition of DLTs were present in 25% of trials. There was no significant difference between the two publication timeframes for this phenomenon (data not shown).

3.7. *Influence of drugs characteristics on the definition of DLT*

Given the heterogeneity observed regarding the minimum NCI CTCAE grade used to define non-haematologic DLT, we investigated whether specific characteristics of the drugs or their schedules influenced the definition of DLT.

NCI CTCAE grade 2 toxicities were significantly more likely to be included in the definition of non-haematologic toxicities where recovery time before the next cycle was $\leqslant$1 week ($p = 0.049$) (Table 6). Conversely, NCI CTCAE grade 1–3 toxicities were statistically less likely to be included in the definition of non-haematologic DLT in trials with discontinuous schedules ($p = 0.01$) or where recovery time before the next cycle was >1 week ($p = 0.04$).

Duration of toxicity required to meet DLT definition was significantly more often specified in trials involving oral drugs ($p = 0.03$) and with recovery time before next cycle of $\leqslant$1 week ($p = 0.01$). Specification of reversibility in the DLT definition was significantly more common in trials involving discontinuous schedules ($p = 0.01$) or where recovery time before the next cycle was >1 week ($p = 0.04$).

Definition of DLT did not statistically differ according to classes of agents as described in Table 2, first-in-human *versus* not first-in-human studies, FDA-approved MTAs *versus* not, and dose escalation method used (data not shown).

4. Discussion

Although the definition of DLT was almost always described in the publications reviewed, the actual definitions of DLT used were heterogeneous across the 155 phase I trials studied. There was a significant trend towards defining more organ-specific DLT items over time. DLT definitions differed not only in terms of the toxicity items used but also in the way these items were described. While DLT definitions were based on NCI CTCAE classification for most toxicity items, the minimum grade considered to be dose-limiting differed substantially for any given item. Baseline status was taken into consideration for only one item reported in one trial. While grade 3/4 non-haematologic toxicities were considered as

Table 5 – Definition of DLT reported in the trials reviewed according to the publication timeframe (2000–2005 *versus* 2006–2010).

	All trial				2000–2005				2006–2010			
	No.	%	Median	Range	No.	%	Median	Range	No.	%	Median	Range
Treatment delay incorporated in the definition of DLT												
Yes	29	19	14 d	[0–28]	13	15	14 d	[7–21]	16	23	14 d	[0–28]
No	126	81			73	85			53	77		
Dose-intensity reduction incorporated in the definition of DLT												
Yes	11	8			4	5			7	10		
No	131	92			75	95			56	90		

DLT, dose-limiting toxicity.

Table 6 – Influence of route, schedule, and recovery time before next cycle on minimum NCI CTCAE grade used to define non-haematologic DLT, specification of a minimum duration, and specification of reversibility in the 155 trials.

	Number of trials (%)							
	⩾G4 to meet DLT definition		⩾G2 to meet DLT definition		Minimum duration mentioned		Reversibility mentioned	
	Yes	No	Yes	No	Yes	No	Yes	No
Route								
PO	8 (10%)	69 (90%)	17 (22%)	60 (78%)	14 (18%)	63 (82%)	6 (8%)	71 (92%)
IV	12 (15%)	66 (85%)	14 (20%)	64 (80%)	6 (8%)	72 (92%)	12 (15%)	66 (85%)
	$P = 0.11$		$P = 0.13$		$P = 0.03$		$P = 0.07$	
Schedule								
Between 33% and 100% of dosing days per cycle	4 (6%)	68 (94%)	14 (19%)	58 (81%)	12 (17%)	60 (83%)	4 (6%)	68 (94%)
Up to 33% of dosing days per cycle	14 (18%)	63 (82%)	15 (19%)	62 (81%)	8 (10%)	69 (90%)	14 (18%)	63 (82%)
	$P = 0.01$		$P = 0.16$		$P = 0.10$		$P = 0.01$	
Recovery time before next cycle								
⩽1 week	3 (6%)	50 (94%)	14 (26%)	39 (74%)	12 (23%)	41 (77%)	3 (6%)	50 (94%)
>1 week	15 (16%)	81 (84%)	15 (16%)	81 (84%)	8 (8%)	88 (92%)	15 (16%)	81 (84%)
	$P = 0.04$		$P = 0.049$		$P = 0.01$		$P = 0.04$	

dose-limiting most of the time, it is worth asking whether a decline in NCI CTCAE grade from 0 to 2 should also be considered a DLT for specific items. In some instances, other classifications than the NCI CTCAE definitions were used, suggesting that the latter classification is not always optimal to grade certain clinical toxicities.

With the exception of neutropenia, duration of toxicity was infrequently included in the definition of DLTs. Where specified, a minimum duration of toxicity was included in the definition of DLT mainly for grade 2 non-haematologic toxicities. For severe or potentially life-threatening toxicities, such as febrile neutropenia, duration was logically not taken into consideration. Specification on reversibility of toxicity was also rarely included in the definition of DLT. Whereas every irreversible toxicity should probably always be considered DLT, failure to recover from toxicity within a certain time might also be relevant in defining DLT.

Although still infrequent, the necessity to delay treatment or to reduce dose-intensity because of toxicity has been increasingly included in the definition of DLTs over time. The clinical relevance of these DLTs is easily seen as an inability to maintain dose intensity over sufficient period of time may hamper the potential efficacy of a cytostatic MTA.

The proportion of dosing days per cycle also appeared to influence the definition of DLT. Milder (NCI CTCAE grade 2) non-haematologic toxicities were more likely to be considered DLTs where the recovery time before the next cycle was short, presumably because a short recovery time would not allow even mild toxicities to resolve before treatment resumption. Alternately, discontinuous schedules were more likely to consider only severe (grade 4) non-haematological toxicities as DLTs, presumably because the longer recovery time allowed less severe toxicities to recover before treatment re-challenge. Consistently, minimum duration was included more frequently in definitions of non-haematologic DLTs where trials utilised schedules with short recovery time before treatment re-challenge, while reversibility was more often included in DLT definitions where trials utilised discontinuous schedules with long recovery time before treatment resumption.

Our study revealed that 23% of the organ-specific DLTs defined in the trials in which non-haematologic DLTs occurred were actually encountered, probably adding value to the generic DLT definitions. However, discrepancies between the initial and the final definition of DLT were reported in 25% proportion of trials. This observation may not be inappropriate, as flexibility is essential in phase I trials. What is considered a DLT may need to evolve as more experience is obtained with a drug, especially if such changes are required for patient safety.

Our paper has several limitations. Other factors not recorded in our study may have partly been able to explain the heterogeneity in the observed definition of DLT, including: (1) data obtained from preclinical toxicology, (2) differences in inclusion/exclusion criteria, (3) the input of regulatory agencies, and (4) space limitation in manuscripts leading to present less information on the definition of DLT than in the protocols. In addition, the use of different versions of the NCI CTCAE classification might have biased the results, although alterations in subsequent versions were more of additions than absolute changes.

It would also have been interesting to examine whether the study of a first-in-class agent impacted on the definition of DLT. One would expect that DLT definition may differ for the first-in-class *versus* subsequent studies as there are more human toxicity data known in the latter studies. However, we found it very difficult to retrieve this information thus limiting our analyses.

Finally, we acknowledge that the MTAs classification used in this study may not be optimal to evaluate the influence of the different classes of agents on the DLT definition. We initially intended to categorise the MTAs according to therapeutic targets or pathways, which led to a classification comprising of 35 different items. As this was overly complicated and insensible, we elected to use the classification described in this paper. Although imperfect, we believe this

classification provides a balanced approach between practicality and useful dissemination of information. Furthermore, as this classification comprises only of a few items, it made direct comparisons not only possible, but decipherable.

This study is, to our knowledge, the first to describe the heterogeneity of DLT definition in phase I cancer clinical trials in the era of MTAs. The fact that the definition of DLT was largely dependent on drug schedule is an essential finding, as it negates the ability to standardise the DLT definition in phase I trials. However, as there was a high rate of discrepancy between the initial and final definitions of DLT, it is pivotal that an expert panel be brought together in an attempt to improve the definition of DLT in phase I trials. In this regard, we highly commend the support given by the European Organisation for Research and Treatment of Cancer (EORTC) on the creation of the European Task Force on Phase I Methodology.

As a first step, we recommend the inclusion of the following points in the definition of DLT in order to reduce the heterogeneity observed across trials: (1) the DLT assessment period should always be clearly stated in the manuscript, (2) for toxicity events that can be graded according to NCI CTCAE classification, the definition of some DLTs should incorporate both absolute severity as well as severity relative to baseline status, (3) a minimum duration of toxicity to meet DLT definition should be specified when relevant, (4) reversibility of toxicity within a certain period of time should also be incorporated in the definition of DLT when relevant, and (5) treatment delay and/or dose-intensity reduction should be incorporated in the definition of DLT.

Conflict of interest statement

None declared.

Acknowledgement

We would like to thank Lillian L. Siu for her review and constructive suggestions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2011.03.016.

REFERENCES

1. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;**276**:637–9.
2. Moher D, Schulz KF, Altman D. CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**285**:1987–91.
3. Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;**13**:S19–21.
4. Arbuck SG. Workshop on phase I study design. Ninth NCI/EORTC New Drug Development Symposium, Amsterdam, March 12, 1996. *Ann Oncol* 1996;**7**:567–73.
5. Le Tourneau C, Stathis A, Vidal L, et al. Choice of starting dose for molecularly targeted agents evaluated in first-in-human phase I cancer clinical trials. *J Clin Oncol* 2010;**28**:1401–7.
6. Postel-Vinay S, Arkenau HT, Olmos D, et al. Clinical benefit in Phase-I trials of novel molecularly targeted agents: does dose matter? *Br J Cancer* 2009;**100**:1373–8.