Original Research

Systematic review of adoption, reporting and impact of health-related quality of life in phase III non-inferiority trials of systemic oncology treatments

Sara Notarnicola [a,1], Lucrezia Zumstein [a,1], Jessica Paparo [a,1], Laura Marandino [b,2], Francesco Perrone [c], Massimo Di Maio [a,*,3,4]

[a] *Department of Oncology, University of Turin, Ordine Mauriziano Hospital, Turin, Italy*
[b] *Department of Medical Oncology, IRCCS Ospedale San Raffaele, Milan, Italy*
[c] *Clinical Trial Unit, Istituto Nazionale Tumori IRCCS Fondazione Pascale, Napoli, Italy*

ABSTRACT

Background: Quality of life (QoL) assessment and patient-reported outcomes appear to be crucial in the rationale and interpretation of non-inferiority (NI) trials. The aim of this study was to assess the inclusion of QoL among endpoints in phase III NI oncology trials and the relevance of QoL results in the reporting and interpretation of these studies.
Materials and methods: By PubMed search and hand-search of 11 selected journals, we identified phase III NI trials in adult patients affected by solid tumours, published between 2012 and 2021. Trials were classified according to 4 NI strategies: (1) different drugs; (2) alternative drug administration routes; (3) shorter treatment duration; (4) "deintensification" of treatment schedule. Three main endpoints were: (1) the proportion of publications including QoL among endpoints; (2) the proportion of primary publications reporting QoL results; (3) the proportion of trials with available QoL results actually favoring the experimental treatment out of trials declaring NI.
Results: 106 publications were eligible. QoL was included among endpoints in 59 studies (55.7%), and QoL results were available in 40 primary publications (37.7%). In the 73 trials testing the NI of different drugs, QoL was included in 43 trials (58.9%) and QoL results were present in 31 publications (42.5%). Among the 74 trials formally demonstrating NI, only 19 trials (25.7%) had QoL results actually supporting the experimental treatment.
Conclusions: In many NI trials in oncology, assessment and reporting of QoL are deficient. Furthermore, most trials formally claiming NI cannot count on QoL results actually supporting the experimental arm.

1. Introduction

A patient-reported outcome (PRO) is an outcome evaluated directly by the patient and based on the patient's perception of the disease and its treatment [1]. PROs include health-related quality of life (QoL), which is universally considered a measure of clinical benefit for patients with cancer [2]. Not only from a regulatory but also from a clinical point of view, the main objective of any anticancer treatment should be to allow patients to live longer and/or to live better, and the point of view

of patients is crucial to evaluate the value of new anticancer treatments [3].

Unlike superiority trials, conducted with the aim of demonstrating a better efficacy for the experimental treatment compared to the best therapy already available in clinical practice, non-inferiority (NI) trials are conducted accepting a potentially lower efficacy (within a clinically reasonable margin) of the experimental treatment [4,5]. The experimental treatment tested within a NI trial should have some clinically relevant advantages for the patient compared to the standard therapy so

---

that it is ethically acceptable to propose participation in these studies [6]. Namely, the experimental therapy should demonstrate to be associated with better tolerability, lower risk of adverse events, or a more comfortable way of administration. Consequently, PROs and QoL should have great relevance in the definition of the value of a candidate non-inferior treatment. European Medicines Agency (EMA) includes the need for "*differentiating two treatments in the NI trial setting*" among the good reasons to include PRO assessment in the clinical development programme for oncology medicinal products [7]. Definitely, QoL assessment and PROs appear to be crucial both in the very rationale and interpretation of NI trials.

In some cases, the benefit in terms of patients' QoL could be taken for granted by the investigators, without the need for an explicit measurement (for example, when patients assigned to the experimental arm receive a shorter duration of the same treatment, or when the route of administration is inherently considered more comfortable). This consideration could apparently justify the absence of QoL among the endpoints, or the absence of QoL results in the main presentation and interpretation of the study. However, it could be considered questionable, as patients' preferences, although *bona fide* presumed by the investigators, should be better demonstrated through the use of PROs. In other cases, when the study compares different drugs, the need to use PROs and to measure QoL should be intrinsically unavoidable. In this case, the demonstration of superiority of the experimental treatment in terms of QoL allows to support the rationale of the study and is essential for the proper interpretation of the result.

We previously showed that the proportion of randomised phase III trials including QoL among study endpoints, although increased in recent years, remains suboptimal and that, in several cases, QoL data are not presented in the primary publication, although collected according to study protocol [8,9]. The aim of this study was to specifically assess the inclusion of QoL among endpoints in phase III NI trials and the relevance of QoL results in the analysis, reporting, and interpretation of these studies.

2. Materials and methods

We included in this analysis the phase III NI trials testing different cancer treatment strategies in adult patients affected by solid tumours, published in a 10-year time-frame between 2012 and 2021.

A part of the eligible papers had already been included in a previous database, obtained by hand-searching all primary publications of phase III trials evaluating anticancer drugs published between 2012 and 2021 by 11 major journals [9]. In order to systematically complete the list of papers extracted from the original database, a PubMed search without journal restrictions was conducted in November 2022. The following keywords were used: *random* AND cancer AND (inferior* OR non-inferior*) AND ("2012/01/01"[Date - Publication]: "2021/12/31"[Date - Publication])*, with a filter for *Randomised Controlled Trials*. As a quality control of the sensitivity of PubMed search strategy adopted for the current analysis, we verified the proportion of papers that were part of the previous hand-search included in the PubMed output. This sensitivity was equal to 89.6% (60 out of 67 papers).

The data was collected with a modified version of the dedicated case report form used for the previous database, and the existing electronic dataset, including one record for each paper, was updated. For each trial, information about publication (journal, year, first author, date of publication, availability of online supplemental material, and/or study protocol) was collected. Information collected about the clinical trial included: the study sponsor (industry-sponsored versus academic), disease setting (early stages versus advanced/metastatic disease), type of malignancy, and details of treatment of both experimental and control arms.

We categorised eligible trials according to 4 NI strategies: (1) different drugs administered in the study arms *(e.g. paclitaxel plus cisplatin versus paclitaxel plus carboplatin)*; (2) alternative drug administration routes *(e.g. oral paclitaxel versus intravenous paclitaxel, or 3-monthly injection versus monthly injection of goserelin)*; (3) shorter duration of treatment, including intermittent treatment strategies *(e.g. 9-week trastuzumab versus 1-year trastuzumab, or intermittent versus continuous androgen deprivation therapy)*; (4) "de-intensified" treatment schedule, including the exclusion of one or more drugs from treatment schedule, or the sequential administration versus combination of the same drugs, or the administration of a reduced dose of the same drug *(e. g. maintenance bevacizumab alone versus bevacizumab plus fluoropyrimidine, or reduced dose versus full dose of cabazitaxel)*.

Details about study endpoints (both primary and secondary/exploratory) was retrieved from the Methods sections of the papers and from the study protocols (when available as supplementary material). When QoL was not listed among endpoints in the methods and the study protocol was not available, QoL was considered as apparently absent. For all records, secondary QoL publications were searched in PubMed, by using the name of the drug(s) and/or tumour type and/or the name of authors of the primary publication and/or the study acronym/code, when available.

The main endpoints of the analysis were three: (1) proportion of trials including QoL among study endpoints; (2) proportion of primary publications with available QoL results; (3) proportion of positive trials (NI declared) with published QoL results actually favoring the experimental treatment. For the latter endpoint, two different definitions of positive trial were applied: (a) trials with a formal demonstration of NI according to the study hypothesis and (b) sensitivity analysis, trials with positive conclusions by the authors about the experimental treatment, even in the absence of a formal demonstration of NI.

Contingency tables were used to describe the association among the outcomes of interest and the study characteristics (year of primary manuscript, journal impact factor, disease stage, study sponsor, type of malignancy, type of experimental therapy, NI strategy). Chi-square was applied to test for statistical significance. Given the exploratory nature of the analysis, no correction was applied for multiple testing. All analyses were performed with IBM SPSS Statistics for Windows, version 27.0.

3. Results

A flowchart of the study selection process is shown in Fig. 1. Out of the 834 papers included in the original database, 67 NI trials were selected. PubMed search retrieved 1316 publications, including 60 out of the 67 trials included in the original database (corresponding to a sensitivity of 89.6% for the string used), and further 40 eligible publications.

Overall, 106 eligible publications were included in the analysis, 53 published in the 2012–2016 period and 53 published between 2017 and 2021. The list of eligible publications is reported in Supplementary appendix, and the main characteristics of the eligible publications are summarised in Table 1. Most trials (n = 73, 68.9%) tested NI of different drugs (or different strategies) compared to the standard treatment.

Out of the 106 eligible publications, QoL was included among endpoints in 59 studies (55.7%), and QoL results were available in 40 primary publications (37.7%). (Fig. 2). The proportion of studies including QoL among endpoints according to study characteristics is reported in Table 2. No significant difference was found in the proportion of studies including QoL among endpoints neither according to the period of publication (52.8% between 2012 and 2016 versus 58.5% between 2017 and 2021, p = 0.56), nor according to study sponsor (56.4% in industry-sponsored trials versus 55.2% in academic trials, p = 0.91), nor according to type of malignancy (p = 0.63) nor type of experimental treatment (p = 0.29). The proportion of studies presenting QoL results in primary publications according to study characteristics is reported in Table 3. No significant difference was found in the proportion of studies including QoL results in primary publication neither according to the period of publication (35.8% between 2012 and 2016 versus 39.6%
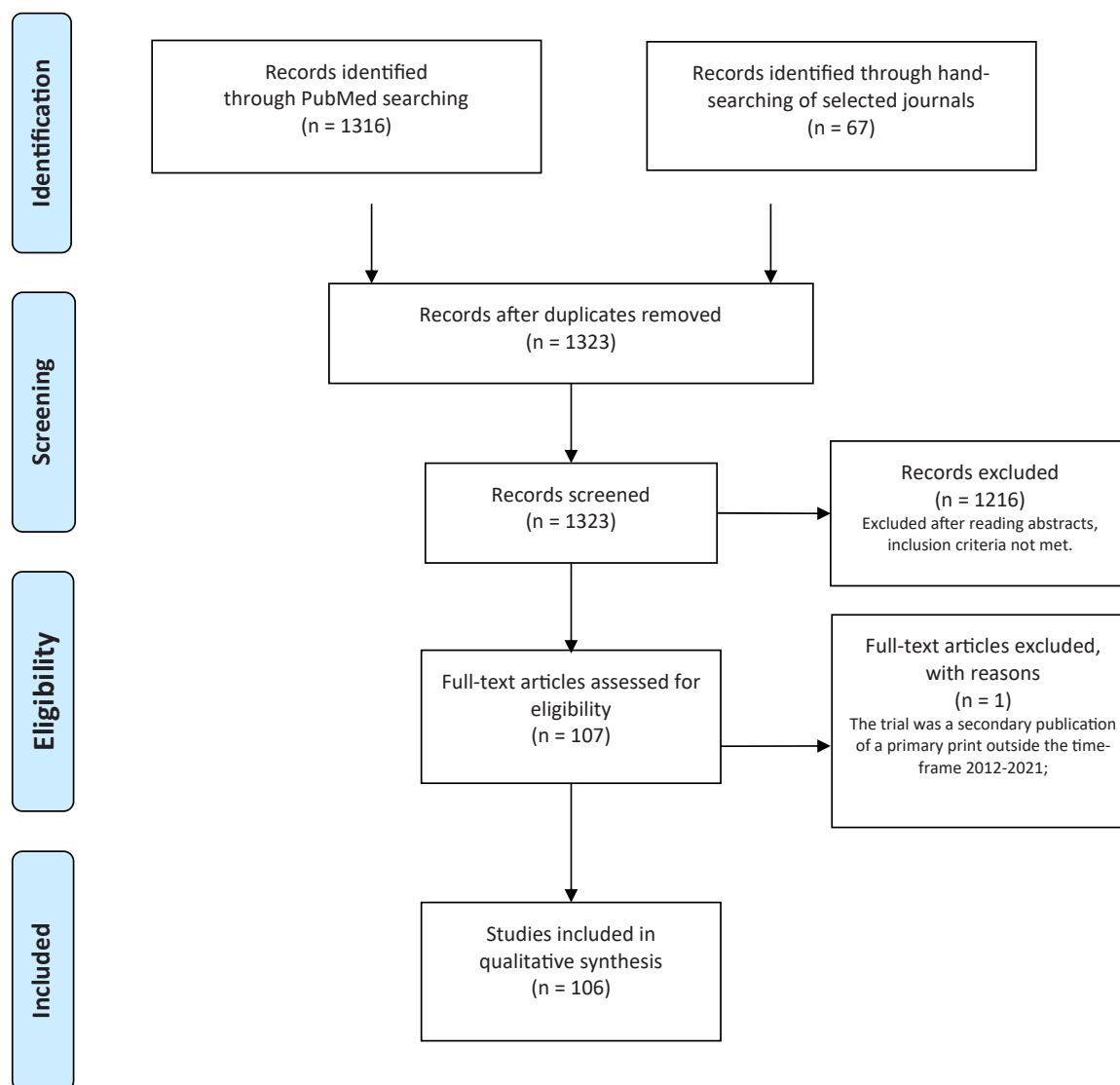
Fig. 1. PRISMA 2009 flow diagram, depicting the flow of information through the different phases of this systematic review.

between 2017 and 2021, p = 0.69), nor according to the study sponsor (33.3% in industry-sponsored trials versus 40.3% in academic trials, p = 0.48), nor according to type of malignancy (p = 0.18) nor type of experimental treatment (p = 0.80).

Focusing on the subgroup of 73 trials examining the NI of different drugs, QoL was not included among endpoints in 43 trials (58.9%), and QoL results were available in 31 publications (42.5%).

Among the 74 trials with formal demonstration of NI of the experimental arm, QoL was evaluated in 38 (51.4%), QoL results were present in 31 primary publications (41.9%) and there were QoL results in favour of the experimental treatment in only 19 cases (25.7%) (Fig. 3A-C). In the subgroup of trials testing the NI of different drugs, out of 55 trials with formal demonstration of NI, QoL was included in 30 trials (55.5%), QoL results were available in 26 primary publications (43.3) and there were QoL results supporting the experimental treatment in only 15 trials (27.3%) (Fig. 3D-F). Similar results have been observed in the sensitivity analysis among the 76 trials with positive conclusions by the authors, independently of the formal demonstration of NI (Supplementary Fig. 2A-F).

4. Discussion

In this analysis, including 106 oncology trials designed to demonstrate the NI of the experimental treatment and published over a 10-year period, only a fraction (55.7%) included QoL among study endpoints, an even smaller fraction (37.7%) reported QoL results in the primary publication and as a consequence only 19 of the 74 trials where the experimental treatment was defined as non-inferior (and potentially suggested for adoption in clinical practice) had QoL result in favour of the experimental treatment and supporting its use.

NI trials have been accused of being unethical, in that the patient agrees to participate in a clinical "gamble" where - if things go well - the experimental treatment will produce efficacy not significantly worse than the standard [10]. Reasonably, this makes sense if the experimental treatment has clear benefits. If not, why would a patient prefer the experimental treatment? From this point of view, it was quite disappointing to observe that, among the 74 trials formally demonstrating NI, only 19 trials (25.7%) have actually produced QoL results supporting the experimental treatment.

Especially when sponsored by the pharmaceutical industry, NI trials could be accused of being a "fast lane" for introducing non-innovative and "*me too*" treatments into clinical practice. As a matter of fact, in our analysis, the majority of NI studies identified were academic trials. In terms of attention to QoL, industry-sponsored studies and academic studies were not significantly different: in detail, the percentage of studies that included QoL as an endpoint was 56.4% and 55.2% in

Table 1

Characteristics of the 106 publications included in the analysis.

	n	(%)
Year of primary manuscript		
2012–2016	53	50.0
2012	10	9.4
2013	13	12.3
2014	6	5.7
2015	12	11.3
2016	12	11.3
2017–2021	53	50.0
2017	9	8.5
2018	13	12.3
2019	12	11.3
2020	7	6.6
2021	12	11.3
Journal impact factor		
<15	55	51.9
15–30	27	25.5
>30	24	22.6
Disease stage		
Early (adjuvant / neoadjuvant)	36	34.0
Advanced / metastatic	70	66.0
Study sponsor		
Industry-sponsored	39	36.8
Academic	67	63.2
Type of malignancy		
Breast	29	27.4
Lung	15	14.2
Gastro-intestinal	39	36.8
Genito-urinary	15	14.2
Other	8	7.5
Type of experimental therapy		
Chemotherapy	70	66.0
Targeted therapy	18	17.0
Hormonal therapy	18	17.0
Non-inferiority strategy		
Different drugs	73	68.9
Different route of administration of the same drug	7	6.6
Shorter duration of treatment	17	16.0
Omission of one or more drugs	9	8.5
Results of the study		
Formal demonstration of non-inferiority according to study hypothesis	74	69.8
Positive conclusions by the authors	76	71.7

industry-sponsored and academic studies, respectively, while the percentage of primary publications that included QoL results were 33.3% and 40.3%. The methodology of design, analysis, and presentation of results of NI studies should be improved independently of the type of sponsor. Furthermore, no significant differences in the attention to QoL were found according to the period of publication, the journal impact factor, the study sponsor, the type of malignancy, and the type of experimental treatment.

Of course, there are different types of NI studies: in some cases, the preferability of the experimental treatment was probably taken for granted in the study design, choice of endpoints, and interpretation of results (for instance, due to shorter duration of the treatment or to the "deintensification" of therapy). For this reason, we performed subgroup analyses focusing specifically on studies comparing different drugs/combinations, where preferability of experimental arm should not be taken for granted. Unfortunately, results remained disappointing even when the analysis was limited to the latter studies. Namely, in the subgroup of 73 trials examining the NI of different drugs, QoL was included among endpoints in 58.9% of trials, and QoL results were available in 42.5% of the publications. In the end, due to the absence of QoL among the endpoints, to the absence of published QoL results, or to the negative QoL results, only a minority of treatments with formal demonstration of NI have QoL results supporting their preferability.

Among the included studies, we found many trials with formally demonstrated NI of the experimental arm, where the clinical choice actually cannot be based on PROs and QoL results. In other studies, however, PROs and QoL results may contribute to the interpretation of the study, informing the clinical choices. For instance, the randomised trial comparing lenvatinib versus sorafenib as first-line treatment of patients with advanced hepatocellular carcinoma was designed to be formally positive not only in case of superiority but also in case of NI of lenvatinib, with primary endpoint overall survival (OS) and a NI margin set at 1.08 [11]. The trial was formally positive for NI, with hazard ratio for OS 0.92 (95% confidence interval 0.44 – 1.06), along with a higher objective response rate and better PFS with lenvatinib, with a different toxicity profile between the two tyrosine kinase inhibitors. PROs were among secondary and exploratory endpoints and were presented in a secondary publication [12]. Differences in overall mean changes from baseline generally favoured lenvatinib in most scales, although the differences were not statistically or clinically significant. Of note, patients treated with lenvatinib experienced statistically significant delays in

A. QoL included among endpoints
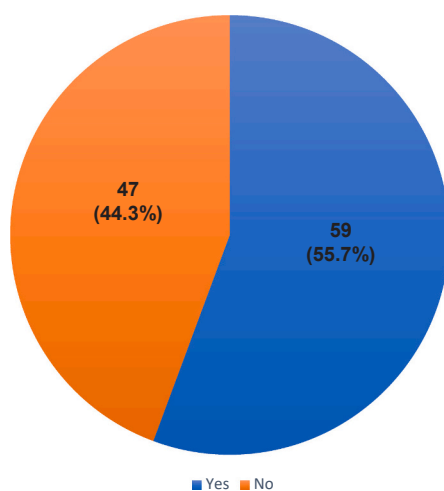
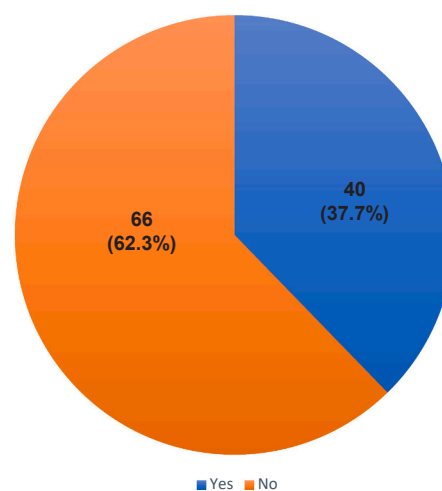**B. QoL results available in primary publication**

Fig. 2. Proportion of trials including quality of life (QoL) among endpoints (panel A) and proportion of publications with available QoL results (panel B), among the 106 publications of non-inferiority (NI) studies included in the analysis. QoL: quality of life.

Table 2
Inclusion of health-related quality of life (QoL) among trial endpoints.

	QoL not included	QoL included	P value
	n (%)	n (%)	
Whole series (n = 106)	47 (44.3%)	59 (55.7%)	
Year of primary manuscript			0.56
2012–2016 (n = 53)	25 (47.2%)	28 (52.8%)	
2017–2021 (n = 53)	22 (41.5%)	31 (58.5%)	
Journal impact factor			0.32
<15 (n = 55)	28 (50.9%)	27 (49.1%)	
15–30 (n = 27)	11 (40.7%)	16 (59.3%)	
>30 (n = 24)	8 (33.3%)	16 (66.7%)	
Disease stage			0.10
Early (adjuvant / neoadjuvant) (n = 36)	20 (55.6%)	16 (44.4%)	
Advanced / metastatic (n = 70)	27 (38.6%)	43 (61.4%)	
Study sponsor			0.91
Industry-sponsored (n = 39)	17 (43.6%)	22 (56.4%)	
Academic (n = 67)	30 (44.8%)	37 (55.2%)	
Type of malignancy			0.63
Breast (n = 29)	14 (48.3%)	15 (51.7%)	
Lung (n = 15)	6 (40.0%)	9 (60.0%)	
Gastro-intestinal (n = 39)	19 (48.7%)	20 (51.3%)	
Genito-urinary (n = 15)	4 (26.7%)	11 (73.3%)	
Other (n = 8)	4 (50.0%)	4 (50.0%)	
Type of experimental therapy			0.29
Chemotherapy (n = 70)	34 (48.6%)	36 (51.4%)	
Targeted therapy (n = 18)	5 (27.8%)	13 (72.2%)	
Hormonal therapy (n = 18)	8 (44.4%)	10 (55.6%)	
Non-inferiority strategy			0.012
Different drugs (n = 73)	30 (41.1%)	43 (58.9%)	
Different route of administration of the same drug (n = 7)	7 (100%)	0	
Shorter duration of treatment (n = 17)	8 (47.1%)	9 (52.9%)	
Omission of one or more drugs (n = 9)	2 (22.2%)	7 (77.8%)	

Table 3
Availability of health-related QoL results in the primary publication.

	QoL results not included	QoL results included	P value
	n (%)	n (%)	
Whole series (n = 106)	66 (62.3%)	40 (37.7%)	
Year of primary manuscript			0.69
2012–2016 (n = 53)	34 (64.2%)	19 (35.8%)	
2017–2021 (n = 53)	32 (60.4%)	21 (39.6%)	
Journal impact factor			0.15
<15 (n = 55)	38 (69.1%)	17 (30.9%)	
15–30 (n = 27)	17 (63.0%)	10 (37.0%)	
>30 (n = 24)	11 (45.8%)	13 (54.2%)	
Disease stage			0.052
Early (adjuvant / neoadjuvant) (n = 36)	27 (75.0%)	9 (25.0%)	
Advanced / metastatic (n = 70)	39 (55.7%)	31 (44.3%)	
Study sponsor			0.48
Industry-sponsored (n = 39)	26 (66.7%)	13 (33.3%)	
Academic (n = 67)	40 (59.7%)	27 (40.3%)	
Type of malignancy			0.18
Breast (n = 29)	20 (69.0%)	9 (31.0%)	
Lung (n = 15)	7 (46.7%)	8 (53.3%)	
Gastro-intestinal (n = 39)	28 (71.8%)	11 (28.2%)	
Genito-urinary (n = 15)	8 (53.3%)	7 (46.7%)	
Other (n = 8)	3 (37.5%)	5 (52.5%)	
Type of experimental therapy			0.80
Chemotherapy (n = 70)	42 (60.0%)	28 (40.0%)	
Targeted therapy (n = 18)	12 (66.7%)	6 (33.3%)	
Hormonal therapy (n = 18)	12 (66.7%)	6 (33.3%)	
Non-inferiority strategy			0.17
Different drugs (n = 73)	42 (57.5%)	31 (42.5%)	
Different route of administration of the same drug (n = 7)	7 (100%)	0	
Shorter duration of treatment (n = 17)	11 (64.7%)	6 (36.3%)	
Omission of one or more drugs (n = 9)	6 (66.7%)	3 (33.3%)	

definitive, meaningful deterioration in fatigue, pain, and diarrhoea domains compared to sorafenib, although no significant differences in time to definitive deterioration were observed for other QoL domains or for global health status/QoL score. Based on these results, the authors concluded that the evidence of clinically relevant benefit in several QoL domains supported the use of lenvatinib compared to sorafenib. Of course, the multiplicity issue should be considered when planning and interpreting QoL analysis, given the number of domains and items tested.

In some cases, the non-inferior treatment could have been considered preferable based on the comparison of investigator-reported toxicity, both as part of the pre-existing evidence and based on the clinical trial results themselves. However, as repeatedly shown in previous studies, investigators can significantly under-report subjective symptoms and treatment-related toxicities [13–15]. The careful collection of investigator-assessed and reported adverse events is not a good reason for omitting the adoption of PROs and of QoL among study endpoints.

In recent years, the European Society for Medical Oncology (ESMO) developed a very important instrument for grading the magnitude of clinical benefit of treatments for patients with solid tumours, the Magnitude of Clinical Benefit Scale (MCBS) [16,17]. The ESMO-MCBS assigns categorical benefit scores to cancer drugs approved by EMA and the US Food and Drugs Administration, mostly based on positive results from superiority trials, but also considering NI trials reaching a conclusion of NI. The only form dedicated to results of NI studies is form 2C, which is for non-curative treatments (so it can be applied only to advanced settings), where the score ranges from 1 to 5 and scores ≥4 are associated with high value. According to evaluation form 2C, a treatment can receive a score of 4 "*in case of reduced toxicity or improved QoL (using validated scale), with evidence for statistical non inferiority or superiority in PFS / OS*", while it can receive a score of 3 "*in case of improvement in some symptoms (using a validated scale) but without evidence of improved overall QoL*". This means that (with the exception of those clearly showing reduced toxicity) according to the current version of ESMO-MCBS a NI study without PROs among endpoints cannot be properly evaluated, and those without QoL advantage are considered without evaluable benefit.

In conclusion, in many NI trials in oncology, assessment, reporting and consideration of QoL are suboptimal. Furthermore, most trials formally claiming NI cannot count on QoL results actually supporting the experimental arm. Given the fact that NI studies are ethically sensitive, the scientific community should pay particular attention to patients' QoL in the design, analysis, and interpretation of this type of study.

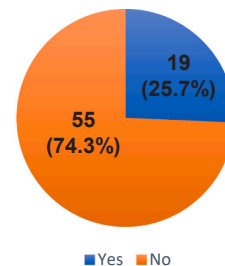## Trials demonstrating non-inferiority of the experimental arm (n=74)

A. QoL included among endpoints
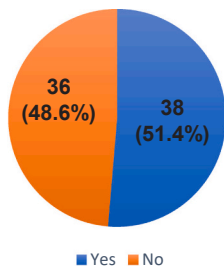
B. QoL results available in primary publication
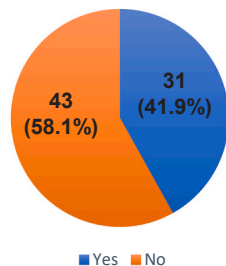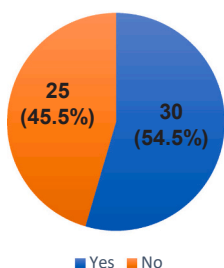
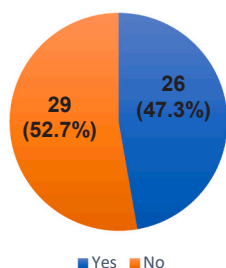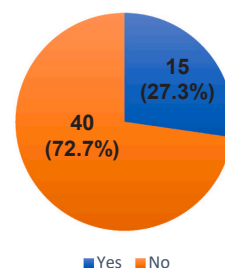### C. QoL results supporting experimental treatment

Trials testing the non-inferiority of different drugs, demonstrating non-inferiority of the experimental arm (n=55)

D. QoL included among endpoints

E. QoL results available in primary publication

F. QoL results supporting experimental treatment

Fig. 3. Proportion of trials including QoL among endpoints (panel A), proportion of publications with available QoL results (panel B), and proportion of trials with QoL results in favour of the experimental treatment (panel C) among the 74 trials formally demonstrating NI of the experimental arm. Proportion of trials including QoL among endpoints (panel D), proportion of publications with available QoL results (panel E), and proportion of trials with QoL results in favour of the experimental treatment (panel F) among the 55 trials testing the non-inferiority of different drugs, formally demonstrating NI of the experimental arm. QoL: health-related quality of life.

Ethics approval

Not applicable.

Funding

No specific funding was received for this work.

Declaration of Competing Interest

The authors declare no support from any organisation for the submitted work; LM reports research grant from AstraZeneca and travel expenses from Janssen; FP reports honoraria from Bayer, Pierre Fabre, AstraZeneca, Incyte, Ipsen, Clovis, Astellas, Sanofi, institutional funding for work in clinical trials/contracted research from Roche, Bayer, AstraZeneca, Pfizer, Incyte, Tesaro/GlaxoSmithKline and Merck; MDM reports honoraria from AstraZeneca, Boehringer Ingelheim, Janssen, Merck Sharp & Dohme (MSD), Novartis, Pfizer, Roche, Takeda for consultancy or participation to advisory boards and direct research funding from Tesaro/GlaxoSmithKline, institutional funding for work in clinical trials/contracted research from Beigene, Exelixis, MSD, Pfizer and Roche; no other relationships or activities that could appear to have influenced the submitted work.

Data availability

Data will be available upon reasonable request. Requests should be addressed to the corresponding author.

Author contributions

Sara Notarnicola, **Lucrezia Zumstein**, **Jessica Paparo**: Conceptualization, Methodology, Investigation, Writing – original draft. **Laura Marandino**, **Francesco Perrone**: Investigation, Writing – review & editing. **Massimo Di Maio**: Conceptualization, Methodology, Investigation, Writing – original draft, Supervision. All authors contributed to the data interpretation, critically revised the manuscript, and approved the final version for publication. Massimo Di Maio is the guarantor of this manuscript and accepts full responsibility for the work and the conduct of the study, has access to the data, and controls the decision to publish. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Transparency statement

The lead author (MDM) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported and that

no important aspects of the study have been omitted.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejca.2023.113374.

References

[1] Di Maio M, Basch E, Denis F, Fallowfield LJ, Ganz PA, Howell D, et al. The role of patient-reported outcome measures in the continuum of cancer clinical care: ESMO Clinical Practice Guideline. Ann Oncol 2022;33(9):878–92. https://doi.org/10.1016/j.annonc.2022.04.007.

[2] Oosting SF, Barriuso J, Bottomley A, Galotti M, Gyawali B, Kiesewetter B, et al. Methodological and reporting standards for quality-of-life data eligible for European Society for Medical Oncology-Magnitude of Clinical Benefit Scale (ESMO-MCBS) credit. Ann Oncol 2023;34(4):431–9. https://doi.org/10.1016/j.annonc.2022.12.004.

[3] Arciero V, Delos Santos S, Koshy L, Rahmadian A, Saluja R, Everest L, et al. Assessment of Food and Drug Administration- and European Medicines agency-approved systemic oncology therapies and clinically meaningful improvements in quality of life: a systematic review. JAMA Netw Open 2021;4(2):e2033004. https://doi.org/10.1001/jamanetworkopen.2020.33004. PMID: 33570573; PMCID: PMC7879236.

[4] Pocock SJ. The pros and cons of noninferiority trials. Fundam Clin Pharmacol 2003;17(4):483–90. https://doi.org/10.1046/j.1472-8206.2003.00162.x.

[5] Cuzick J, Sasieni P. Interpreting the results of noninferiority trials—a review. Br J Cancer 2022;127:1755–9. https://doi.org/10.1038/s41416-022-01937-w.

[6] Riechelmann RP, Alex A, Cruz L, Bariani GM, Hoff PM. Non-inferiority cancer clinical trials: scope and purposes underlying their design. Ann Oncol 2013;24(7):1942–7. https://doi.org/10.1093/annonc/mdt073. Epub 2013 Mar 13.

[7] European Medicines Agency. Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies; 2016. ⟨https://www.ema.europa.eu/en/news/integrating-patients-views-clinical-studies-anticancer-medicines⟩. [Last accessed 11th October 2022].

[8] Marandino L, La Salvia A, Sonetto C, De Luca E, Pignataro D, Zichi C, et al. Deficiencies in health-related quality-of-life assessment and reporting: a systematic review of oncology randomized phase III trials published between 2012 and 2016. Ann Oncol 2018;29(12):2288–95. https://doi.org/10.1093/annonc/mdy449.

[9] Marandino L, Trastu F, Ghisoni E, Lombardi P, Mariniello A, Reale ML, et al. Time trends in health-related quality of life assessment and reporting within publications of oncology randomised phase III trials: a meta-research study. BMJ Oncol 2023;2:e000021. https://doi.org/10.1136/bmjonc-2022-000021.

[10] Garattini S, Bertele' V. Non-inferiority trials are unethical because they disregard patients' interests. Lancet 2007;370(9602):1875–7. https://doi.org/10.1016/S0140-6736(07)61604-3. Epub 2007 Oct 23.

[11] Kudo M, Finn RS, Qin S, Han KH, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 2018;391(10126):1163–73. https://doi.org/10.1016/S0140-6736(18)30207-1.

[12] Vogel A, Qin S, Kudo M, Su Y, Hudgens S, Yamashita T, et al. Lenvatinib versus sorafenib for first-line treatment of unresectable hepatocellular carcinoma: patient-reported outcomes from a randomised, open-label, non-inferiority, phase 3 trial. Lancet Gastroenterol Hepatol 2021;6(8):649–58. https://doi.org/10.1016/S2468-1253(21)00110-2. Epub 2021 Jun 2. PMID: 34087115.

[13] Basch E. The missing voice of patients in drug-safety reporting. N Engl J Med 2010;362(10):865–9. https://doi.org/10.1056/NEJMp0911494.

[14] Di Maio M, Gallo C, Leighl NB, Piccirillo MC, Daniele G, Nuzzo F, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol 2015;33(8):910–5. https://doi.org/10.1200/JCO.2014.57.9334.

[15] Di Maio M, Basch E, Bryce J, Perrone F. Patient-reported outcomes in the evaluation of toxicity of anticancer treatments. Nat Rev Clin Oncol 2016;13(5):319–25. https://doi.org/10.1038/nrclinonc.2015.222.

[16] Cherny NI, Sullivan R, Dafni U, Kerst JM, Sobrero A, Zielinski C, et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). Ann Oncol 2015;26(8):1547–73. https://doi.org/10.1093/annonc/mdv249. Epub 2015 May 30. Erratum in: Ann Oncol. 2017 Nov 1;28(11):2901-2905. PMID: 26026162.

[17] Cherny NI, Dafni U, Bogaerts J, Latino NJ, Pentheroudakis G, Douillard JY, et al. ESMO-magnitude of clinical benefit scale version 1.1. Ann Oncol 2017;28(10):2340–66. https://doi.org/10.1093/annonc/mdx310.