Comparison of baseline patient characteristics in Italian oncology drug monitoring registries and clinical trials: a real-world cross-sectional study

Maria Lucia Iacovino,[a,j] Simone Celant,[b,j] Luca Tomassini,[b] Laura Arenare,[a] Andrea Caglio,[c] Andrea Canciello,[a] Flavio Salerno,[c] Pier Paolo Olimpieri,[b] Susanna Di Segni,[b] Antonella Sferrazza,[b] Maria Carmela Piccirillo,[a] Giordano Domenico Beretta,[d] Carmine Pinto,[e] Livio Blasi,[f] Saverio Cinieri,[g] Luigi Cavanna,[h] Massimo Di Maio,[i] Pierluigi Russo,[b,k] and Francesco Perrone[a,k,*]

[a]National Cancer Institute, IRCCS Fondazione G.Pascale, Naples, Italy
[b]Italian Medicines Agency, Rome, Italy
[c]Department of Oncology, University, Ordine Mauriziano Hospital Umberto I, Turin, Italy
[d]Medical Oncology, Hospital, Pescara, Italy
[e]Medical Oncology, Comprehensive Cancer Centre, AUSL-IRCCS di Reggio Emilia, Italy
[f]Medical Oncology, Civic Hospital Cristina Benfratelli, Palermo, Italy
[g]Medical Oncology and Breast Unit, Perrino Hospital, Brindisi, Italy
[h]Medical Oncology and Hematology, Civil Hospital, Piacenza, Italy
[i]Department of Oncology, University of Turin, AOU Città della Salute e della Scienza di Torino, Turin, Italy

The Lancet Regional Health - Europe 2024;41: 100912

Published Online xxx
https://doi.org/10.1016/j.lanepe.2024.100912

Summary

Background Generalizability of registrative clinical trials to real-world clinical practice is influenced by comparability of patients in the two settings. We compared characteristics of cancer patients in registrative trials with real-world clinical practice in Italy.

Methods Data on age, sex and performance status (PS) were derived from web-based monitoring registries developed by Italian Medicines Agency (AIFA) and corresponding registrative trials reported in the European Public Assessment Reports (EPAR) of European Medicines Agency (EMA). Weighted means were calculated in registries and trials and differences were described. Multivariate analysis was performed using Principal Component Analysis and Cluster Analysis.

Findings From January, 2013 to April, 2023, 419,461 unique pairs of patients and therapeutic indications were recorded in 129 AIFA registries. Within 140 related trials, 87,452 patients had been enrolled. Median age and rate of elderly (≥65 years old) patients were higher in monitoring registries than in clinical trials [mean difference of median age 5.3 years, $p < 0.001$; mean difference of elderly rate 17.17% (95% CI 1.06, 1.48)]. Overall, rate of female patients was not different between registries and trials [mean difference –0.55% (95% CI –1.06, –0.05)]. Mean rate of patients with deteriorated PS was low both in trials (3.1%) and in registries (4.3%) with a mean difference of 1.27% (95% CI 1.06, 1.48). Two clusters were identified with multivariate analysis: one including more registries (higher median age and elderly rate, lower female rate, higher rate of deteriorated patients), the other more trials (lower median age and elderly rate, higher female rate, lower rate of deteriorated patients).

Interpretation This study supports that cancer patients enrolled in trials do only partially represent those who have been treated in Italy in clinical practice. Inclusiveness of registrative trials should be increased to ensure generalizability of results to real-world population.

Funding Partially supported by Italian Ministry of Health.

Keywords: Real-world evidence; Anticancer treatments; Registrative clinical trials; Generalizability; Age; Gender; Performance status

*Corresponding author. National Cancer Institute, IRCCS Fondazione G.Pascale, Naples, Italy.
E-mail address: f.perrone@istitutotumori.na.it (F. Perrone).
[j]Co-first authors.
[k]Contributed equally.

Research in context

Evidence before this study

We searched Pubmed and Google Scholar from database inception to March 31, 2023, for studies comparing clinical trials (CTs) population to real-world (RW) patients. Precisely, we used the keywords "real-world", "real-world evidence", "clinical trials", "clinical trials population", "representativeness", "comparison", and their combination in our search criteria. Only English publications were considered. There were few studies that directly compared CTs and RW populations, mostly in cardiovascular, pulmonary and hematological fields. We did not identify studies in the oncological field that systematically evaluated RW population and its characteristics. Moreover, we did not identify oncological studies that compared RW patients to those enrolled in registrative CTs.

Added value of this study

This is the first cross-sectional RW study comparing oncological patients enrolled in CTs with RW population derived from a large size database of patients registered in clinical practice, thanks to web-based Monitoring Registries (wMRs) developed by the Italian Medicines Agency (AIFA). In Italy, the registration of patients in wMRs is mandatory for reimbursement by the National Health Service (NHS) of a number of oncological drugs, thus ensuring that patients included in our analysis accurately represent the entire population treated with the selected drug in the selected indication, throughout a specific period of time (off label treatments are excluded). Our findings illustrate the distribution of age, sex and performance status (PS) of oncological patients treated over 10 years with one or more drugs under monitoring, also having regard of different types of tumor and class of drugs. The differences emerged from the comparison with CTs population were described and the addition of a multivariate model to the statistical analysis added strength to the research.

Implications of all the available evidence

By describing differences in age, sex and PS, our work shows lower median age and lower rate of elderly and impaired PS patients in CTs than in RW. Sex distribution appears to be balanced overall although differences in female rate were found in some types of tumor. The use of wMRs allows to extremely reduce bias that might typically affect many real-world data sources. Such evidence highlights the risks of generalization of CTs results to RW population. Attention should be paid to expand knowledge about this topic and explore the causes of under-representation for specific categories, in order to improve CT inclusiveness.

Introduction

Randomised clinical trials (RCTs) are the gold standard instruments to produce evidence for regulatory agencies to determine whether the benefit/risk ratio is positive enough to warrant introduction of innovative drugs into clinical practice.[1,2]

A cardinal principle of the drug development process is that patients enrolled in clinical trials (CTs) should be reasonably representative of the population that will be treated in clinical practice, the so-called real-world (RW). This principle is not guaranteed, especially for some categories that are historically excluded or under-represented when conducting a clinical trial (i.e. pediatric, geriatric, women, racial minorities, patients with impaired performance status or poor prognosis).[3,4] Moreover, strict eligibility criteria are frequently introduced into study protocols, both to improve patient safety (excluding those with pre-existing conditions that can represent risk factors for adverse events) and to reduce clinical heterogeneity that can dilute the effect of study treatments. However, the use of strict eligibility criteria is acknowledged as one of the barriers to inclusion of patients in CTs, and is one of the reasons why RCT population may differ from patients in RW in terms of age, sex and also clinical characteristics.[5] As a consequence, generalizability of results found within the clinical trial setting to patients in clinical practice represents a relevant matter of discussion.[6] For this reason, in the past decades, global regulatory authorities have developed specific guidelines to address the matter of scientific evidence in special populations.[7,8]

Collection and analysis of real-world evidence (RWE) is increasingly considered as a valuable tool to be used for further comprehension and description of safety and effectiveness of new treatments in clinical practice.[9,10] Intuitively, RWE can help identify potential differences between clinical trials and clinical practice, paving the way to interventions that may reduce the possible negative impact of such differences on the outcome of patients.

Since 2005, the Italian Medicines Agency (AIFA) developed a web platform of monitoring registries (wMRs) used as a tool for monitoring prescriptive appropriateness of drugs introduced into clinical practice, and for facilitating reimbursement strategies through the introduction of managed entry agreements (MEAs).[11] Such system requires the registration of each patient in order to confirm eligibility for each use of the drugs included in the monitoring strategy, and actually represents a large source of RW data.[12]

To quantify how much patients treated in clinical practice may differ from patients enrolled in trials,

baseline characteristics of patients recorded in wMRs for anticancer drugs were compared with those of patients enrolled in registrative CTs.

Methods

Study design and data sources

This is a cross-sectional real-world study, including all wMRs of oncological drugs, the monitoring of which was released online between January 16th, 2013 and May 19th, 2022. Corresponding CTs used as support for European Medicines Agency (EMA) registration were found in the European Public Assessment Reports (EPAR) available at EMA website.

AIFA wMRs are administrative tools designed to monitor the appropriateness of innovative and high-cost drugs in Italy in the context of real-world clinical practice. The role of AIFA wMRs within the Italian regulation as well as its technical characteristics have already been described elsewhere.[12,13] The use of wMRs is mandatory for the reimbursement by the National Health Service (NHS). As a consequence, patients included in the registries represent a census of the patients treated with the selected drugs in the selected indications within clinical practice, throughout the period in which the corresponding registries are active. For the purpose of the analysis conducted in this study, a database was created joining baseline characteristics extracted from both CTs and wMRs, including age (continuous [median/mean] and categorical [< or ≥65 years old]), sex (male, female) and Eastern Cooperative Oncology Group (ECOG) performance status (PS). The age threshold was chosen at 65 years because it was the value maximizing the possibility of comparison due to how data are reported in clinical trials.

Drugs were classified into the following categories: cytotoxic, hormonal, immunotherapy, target-based, immunotherapy + cytotoxic and target-based + cytotoxic.

Ethical statement

According to decree 196/2003 ("Italian Privacy Code") and decree 101/2018 ("Harmonization Decree" harmonizing the Italian data protection laws with the provision of the General Data Protection Regulation 679/2016—GDPR), the processing of anonymized data does not require authorization by patients if carried out in the performance of public interest or public powers based on a provision of law.

Statistical analysis

The median time of activity of registries was calculated using the Kaplan–Meier estimator, where the event is defined as the registry closure date, and registries still active were censored at the date of data extraction.

The aim of descriptive analysis was to describe differences between CTs and wMRs in terms of age (continuous or categorical), sex and PS distribution.

For each CT/wMR comparison, missing values were described and no substitution was done, except the use of mean age, if available, when median age was missing. In absence of ECOG PS information, Karnofsky categories were used, when transformation was feasible.

Each variable of interest (median age, rate of elderly patients [≥65 years old], rate of females and rate of patients with PS > 1) was graphically described in scatter plots reporting values for CT/wMR couples, size of symbols being proportional to the number of patients enrolled in CTs and color of symbols representing breast, lung, colorectal, prostate and other cancers. For each variable of interest, weighted means with standard deviations (SD), where the number of patients enrolled in CT was used as weight, were calculated in CTs and wMRs. The difference (Δ) between weighted means in wMRs and CTs was calculated (ΔwMRs-CTs) and described, in the whole dataset and according to tumor site and class of drug. For the overall comparison of median age, a Wilcoxon signed-rank test was applied to test the null hypothesis of no difference. For the comparison of the rate of elderly, female and PS > 1 patients 95% confidence intervals of the difference were calculated in order to provide information on both the statistical significance and the variability of the estimates.

Multivariate statistical analysis was performed using Principal Component Analysis (PCA) and Cluster Analysis (CA). PCA model was used to reduce the dimensionality of data through the use of orthogonal variables, obtained as linear combinations of the observed ones, that contain the underlying structure of the available data. CA was performed on the variables selected through PCA using k-means method to identify whether study variables tended to cluster based on euclidean distance as a proxy of their degree of association.[14] The k-means method provides with a centroid for each cluster, representing an average of the units contained in it. Clustering was validated through the maximization of Calinski-Harabasz's index.[15] Once individuated the clusters and the corresponding centroids, statistical significance of differences in the distribution of CTs/wMRs and of the other study variables was tested. A sensitivity multivariable analysis was performed to verify whether the exclusion of single-sex tumor types (prostate, cervix, ovary and breast, even acknowledging that the latter may rarely affect males) did modify the results of the main analyses.

As PCA and CA are unsupervised classification techniques, their use is aimed at verifying whether or not the statistical properties of units allow to a meaningful classification. In the analysis presented in this paper, the scope is to verify whether patient baseline characteristics have relevance in discriminating between CTs and wMRs. This approach has been used following the consideration that, if the population of CTs and wMRs are the same, all the considered variables should be structurally equally distributed across them and,

subsequently, clusters should only separate indications: the centroids of CTs and wMRs in PCA should both be close to the origin of the axes and not be relevant in the characterization of clusters. Any significant difference in the distribution of CTs and wMRs across clusters suggests a structural difference between them.

Role of the funding source

This work was partially supported by the Italian Ministry of Health within the Ricerca Corrente Project L4/10 of the National Cancer Institute of Naples, IRCCS Fondazione G.Pascale. Neither AIFA nor AIFA coauthors received any fund for this work.

Results

From January 16th, 2013 to May 19th, 2023, 131 AIFA wMRs were activated (Supplementary Appendix pages 7–14). Two registries were excluded because dedicated to pediatric tumors (neuroblastoma, osteosarcoma). Across 129 registries, 355,563 patients were treated with one or more monitored drugs (419,461 unique pairs of patients and drug therapeutic indications). Median time of registry activity (namely open to the inclusion of new patients) was 7.2 years (IC 95% 6.6–7.8), with 80 out of 129 registries still active as of May 2023. The types of cancer with the highest number of registries were lung (31, 24%) and breast (20, 15%).

Starting from the drug therapeutic indications monitored by the wMRs, 144 supporting clinical trials (Supplementary Appendix pages 15–20) were identified and the corresponding data were obtained from EPARs. Overall, 121 indications were supported by one trial, 7 indications by two, and 3 indications by three. Four trials were excluded because related to pediatric tumors. Clinical trials used in subsequent analyses included 87,452 patients. Therefore, 129 indications, concerning 12 types of cancer and 61 drugs (Supplementary Table S1, Appendix page 2–3), were available with data coming from both trials and registries. In median age and percentage of elderly analyses, one and 16 CTs were excluded because of missing data, respectively. In PS analysis, 33 indications were excluded either because of incomplete or missing data in CT (n = 18) or because PS > 1 represented an ineligibility criterium for drug access in wMRs (n = 15). The flowchart of the study with details of exclusion criteria is reported in Fig. 1 and details of therapeutic indications by tumor site are reported in Table 1.

Distribution of median age (overall and by type of cancer) in wMRs/CTs is reported in Fig. 2a. Weighted mean (SD) of median age was 65.1 (5.3) years in wMRs and 59.8 (5.2) years in CTs, mean difference being 5.3 years ($p < 0.001$). In all subgroups the differences were positive (higher median age in wMRs), with skin, head & neck, melanoma and lung cancers showing the largest deviations, while no substantial heterogeneity was evident according to the adopted classification of drugs (Table 2 and Supplementary Figure S1, appendix page 4).

No clinical trial reported age of ≥65 as an excluding criterion. Distribution of the rate of elderly patients in wMR/CT is shown in Fig. 2b. Weighted mean (SD) rate of elderly patients was 53.9% (17.4) in wMRs and 36.8% (16.3) in CTs, mean difference being 17.17% (95% CI 16.67, 17.67). In all the subgroups the differences were positive (higher rates of elderly in MRs), higher values being evident in sarcoma, head & neck, lung and skin cancers, and treatments based on immunotherapy alone while a lower difference was found with treatments combining targeted with cytotoxic drugs (Table 2 and Supplementary Figure S1, appendix page 4).

Distribution of the rate of female patients is shown in Fig. 2c. Weighted mean (SD) rate of female patients was 50.6% (32.6) in and 50.7% (33.0) in wMRs and CTs respectively, mean difference being –0.55% (95% CI –1.06, –0.05). Differences according to cancer types were noted in both directions, the largest being observed in lung cancer (lower rate of females in wMRs) and in head & neck cancer (higher rate of females in wMRs) (Table 2 and Supplementary Figure S1, appendix page 4).

Distribution of the rate of PS > 1 patients in wMR/CT is shown in Fig. 2d. Weighted mean (SD) rate of PS > 1 patients was 4.3% (4.2) in wMRs and 3.1% (5.0) in CTs, mean difference being 1.27% (95% CI 1.06, 1.48). In most subgroups differences were in the direction of higher rate of PS > 1 in wMRs, with larger values seen in skin cancer, sarcoma, head & neck and genitourinary cancer and no major variation according to type of drug (Table 2 and Supplementary Figure S1, appendix page 4).

The multivariate analysis was performed on 80 indications, excluding those with missing values and those where the wMRs excluded patients with PS > 1. At PCA analysis, 67.3% of variance was explained by the first two principal components. The first principal component is strongly and positively correlated with age and negatively correlated with the rate of females, suggesting that the indications where women are prevalent have lower median age and lower rate of elderly. The second principal component is strongly correlated with ECOG performance status. PCA suggests no association between demographic features and worse performance status.

The coordinates of the centroid of CTs and wMRs (Fig. 3) suggest that CTs and wMRs are structurally different, especially in terms of the first principal component—hence, based on sex, age and PS > 1 distributions—as CTs tend to include younger patients, a higher prevalence of women and a lower percentage of patients with worse ECOG PS. Cluster analysis individuated two clusters as best partition, characterized by wMRs and CTs, as Cluster 1 included 66.7% of wMRs and 33.3% of CTs and Cluster 2 included 35.3% of wMRs and 64.7% of CTs ($p < 0.001$). Clusters were also

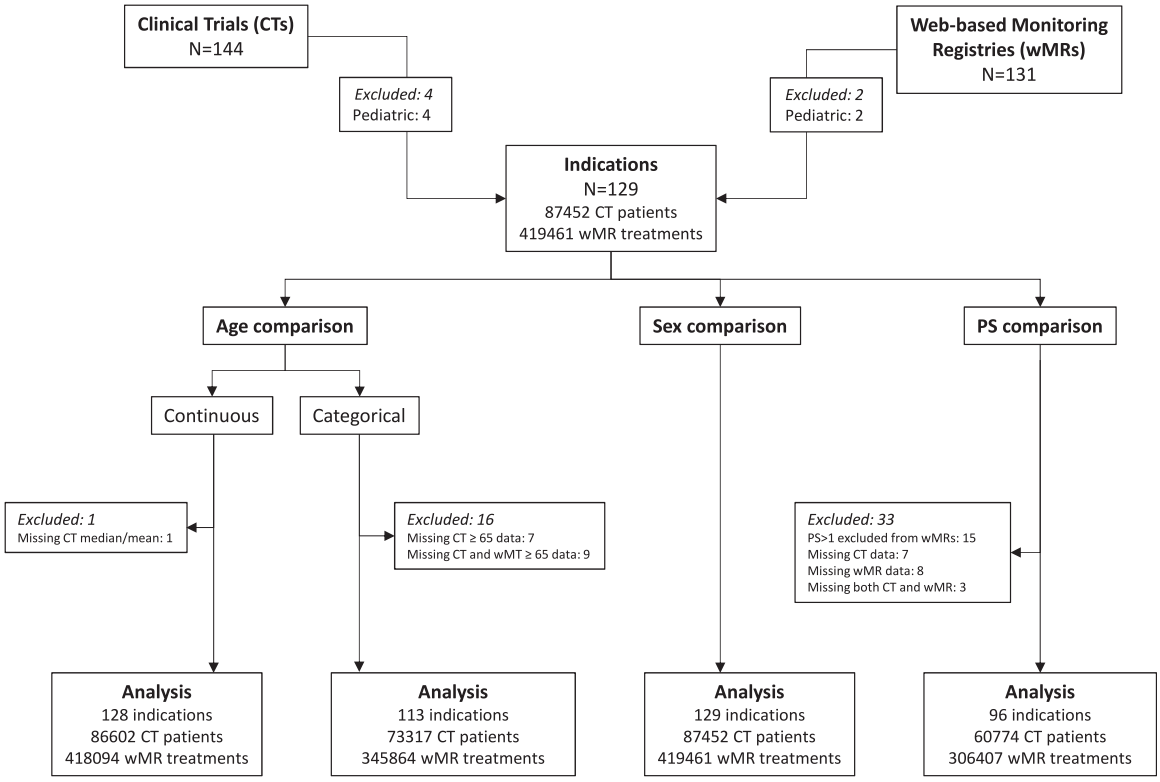

Fig. 1: Flow chart of the study.

	Indications n (%)	Patients in CTs n (%)	Treatments in wMRs n (%)
Tumor site			
Breast	20 (15.5)	17,154 (19.6)	74,020 (17.6)
Colorectal	11 (8.5)	10,954 (12.5)	56,688 (13.5)
Lung	31 (24.0)	16,854 (19.3)	116,560 (27.8)
Prostate	7 (5.4)	6455 (7.4)	25,090 (6.0)
Other	60 (46.5)	36,035 (41.2)	147,103 (35.1)
Genitourinary	14 (10.8)	9401 (10.7)	33,052 (7.9)
Gynecologic	10 (7.7)	6823 (7.8)	19,832 (4.7)
Head & neck	5 (3.9)	2624 (3.0)	8273 (2.0)
Melanoma	13 (10.1)	9355 (10.7)	28,959 (6.9)
Neuroendocrine	3 (2.3)	810 (0.9)	2647 (0.6)
Skin	4 (3.1)	667 (0.8)	5307 (1.3)
Sarcoma	2 (1.5)	635 (0.7)	5622 (1.3)
Upper gastrointestinal	9 (7.0)	5720 (6.5)	43,411 (10.3)
Type of drug			
Cytotoxic	10 (7.8)	7747 (8.9)	71,692 (17.1)
Hormonal	4 (3.1)	4534 (5.2)	19,454 (4.6)
Immunotherapy	22 (17.1)	14,250 (16.3)	71,227 (17.0)
Immunotherapy plus cytotoxic	6 (4.7)	3826 (4.4)	19,794 (4.7)
Target-based	69 (53.5)	36,098 (41.3)	139,502 (33.3)
Target-based plus cytotoxic	18 (14.0)	20,997 (24.0)	97,792 (23.3)

wMRs: web-based monitoring registries; CTs: clinical trials.

Table 1: **Details of indications, treatments and patients by tumor site.**

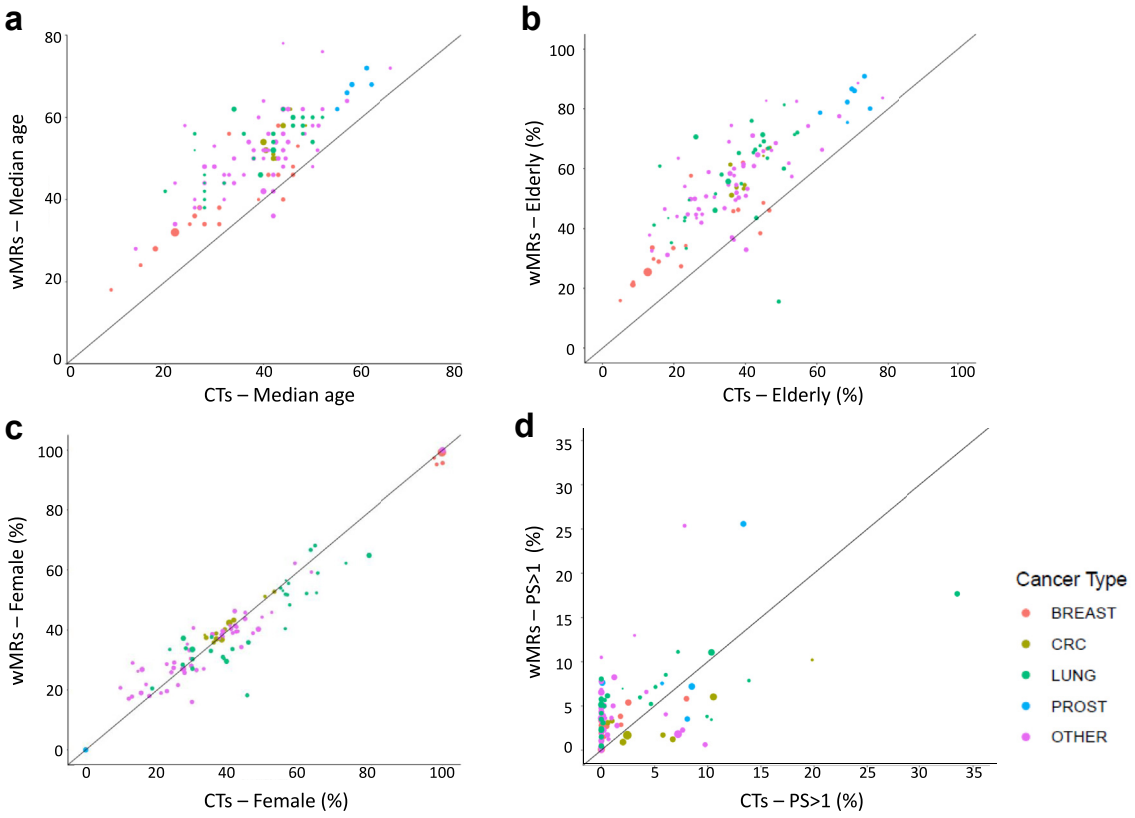

Fig. 2: Distribution by type of cancer of patients' characteristics in wMRs and CTs (2a: median age, 2b: elderly rate; 2c: female rate; 2d: PS > 1 rate).

characterized (Table 3) according to median age (68.2 vs 58.4, $p < 0.001$), elderly rate (62.2% vs 32.8%, $p < 0.001$), female rate (31.7% vs 66.4%, $p < 0.001$) and deteriorated PS (5.8% vs 2.7% rate of PS > 1, $p < 0.001$).

A sensitivity analysis excluding prostate, cervix, ovarian and breast cancer, shows similar results of PCA and CA (Supplementary Figure S2 and Supplementary Table S2, appendix page 5–6).

Discussion

It is commonly reported that patients enrolled in clinical trials of new anticancer drugs might not fully represent the population that will receive those drugs in clinical practice, once positive results lead to an authorization by regulatory agencies.[16] However, few data exist on quantification of differences between the two populations, that may ultimately produce divergent results between efficacy in experimental conditions and effectiveness in clinical practice, or in toxicity observed in the two settings.[17,18] Here, we report an analysis made possible by wMRs, established by AIFA, to warrant appropriate use of new expensive anticancer drugs and allow reimbursement agreements based on the collection of single-patient data. By definition, such registries include treatment indications that are appropriate and consistent with the approved labels, net of further restrictions enforced by AIFA Committees which might not exactly match with pivotal trials or the European marketing authorization. In addition, registries do not capture detailed clinical information on comorbidities that may represent a limitation for the interpretation of data on performance status.

Within a very large size RW population (355,563 patients receiving 419,461 treatments with different drugs and in different therapeutic indications), we found differences in age and performance status distribution as compared to the population of patients enrolled in registrative clinical trials. Such differences tended to vary across different tumor types, but were mostly similar across different drug classes. Even within a framework developed to monitor the appropriate use of drugs, patients in clinical practice were on average more than 5 years older than in clinical trials, with a more than 17% higher rate of elderly patients; and there was a higher rate of patients with worse performance status, although within a very low range of values. Sex

	Median age Δ wMRs–CTs yrs	Elderly rate Δ wMRs–CTs % (95% CI)	Female rate Δ wMRs–CTs % (95% CI)	PS > 1 rate Δ wMRs–CTs % (95% CI)
Overall	5.3 (p < 0.001)	17.17 (16.67, 17.67)	−0.55 (−1.06, −0.05)	1.27 (1.06, 1.48)
By tumor site				
Breast	4.1	12.16 (11.18, 13.13)	−0.62 (−0.78, −0.47)	1.08 (0.66, 1.5)
Colorectal	5.8	17.12 (15.05, 19.19)	1.08 (−0.22, 2.37)	−2.11 (−2.69, −1.53)
Lung	6.5	21.96 (20.89, 23.02)	−3.25 (−4.3, −2.2)	1.7 (1.15, 2.25)
Prostate	4.5	14.28 (12.85, 15.71)	NA	3.6 (2.38, 4.82)
Other	5.5	17.84 (17.06, 18.62)	0.43 (−0.32, 1.19)	1.85 (1.57, 2.13)
GenitoUrinary	5.2	17.99 (16.43, 19.55)	−1.31 (−2.56, −0.05)	3.97 (3.44, 4.49)
Gynecologic	3.1	12.03 (9.96, 14.11)	NA	−2.21 (−2.85, −1.58)
Head & neck	7.2	24.4 (21.82, 26.97)	6.27 (3.98, 8.56)	4.08 (3, 5.16)
Melanoma	6.6	17.23 (15.85, 18.61)	−2.2 (−3.6, −0.79)	1.92 (1.51, 2.34)
Neuroendocrine	3.3	18.67 (13.25, 24.09)	−2.16 (−7.01, 2.7)	NA
Skin	9.5	20.14 (15.56, 24.72)	4.31 (−0.71, 9.32)	10.35 (7.34, 13.36)
Sarcoma	6.1	24.63 (17.52, 31.73)	0.55 (−4.82, 5.93)	4.71 (3.01, 6.41)
Upper gastrointestinal	5.9	19.38 (17.5, 21.27)	4.8 (3.23, 6.38)	−2.33 (−3.31, −1.35)
By type of drug				
Cytotoxic	5.2	18.87 (17.14, 20.61)	1.65 (−0.08, 3.39)	0.65 (−0.19, 1.48)
Hormonal	5	15.94 (14.28, 17.6)	NA	2.83 (1.44, 4.21)
Immunotherapy	5.7	17.68 (16.54, 18.83)	−2.16 (−3.25, −1.06)	2.48 (2.12, 2.85)
Immunotherapy + cytotoxic	4.1	15.84 (13.63, 18.05)	2.75 (0.62, 4.88)	3.84 (3.21, 4.48)
Target-based	6.0	18.84 (18.09, 19.6)	−1.59 (−2.37, −0.82)	1.86 (1.5, 2.21)
Target-based + cytotoxic	4.4	12.51 (11.39, 13.62)	1.05 (0.04, 2.06)	−2.24 (−2.68, −1.81)

wMRs: web-based monitoring registries; CTs: clinical trials; PS: performance status; NA: not applicable.

Table 2: **Weighted mean (SD) differences (Δ) of baseline patients' characteristics between Monitoring Registries (wMRs) and Clinical Trials (CTs) by tumor site and type of drug.**

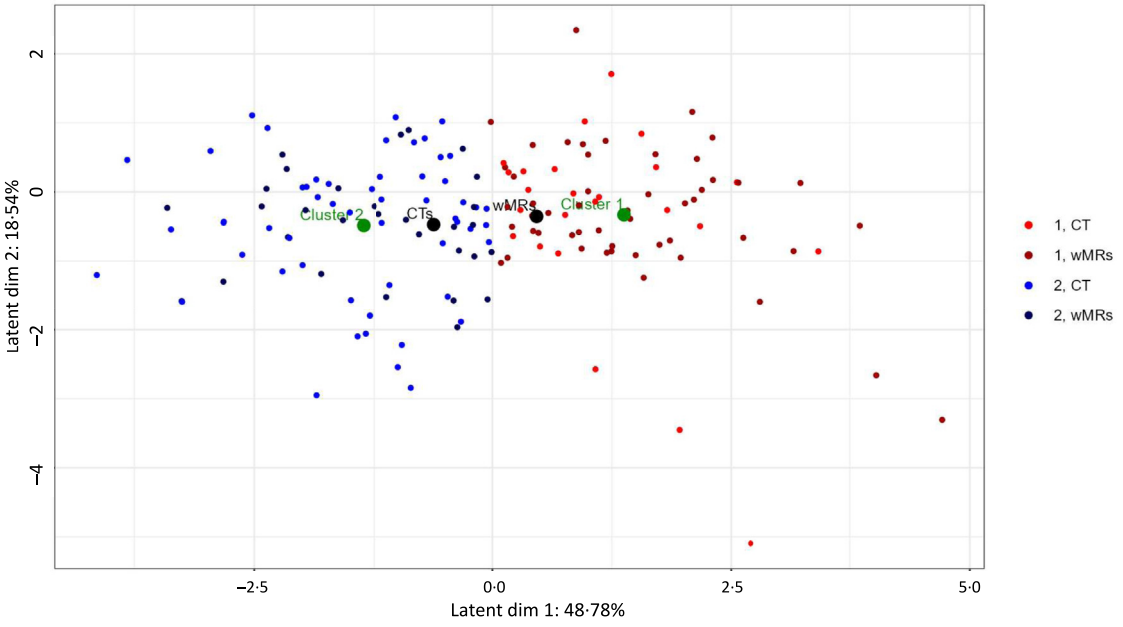

Fig. 3: Cluster analysis.

	Cluster 1	Cluster 2	p value
Median age (yrs)	68.2	58.4	<0.0001
Elderly rate (%)	62.2	32.8	<0.0001
Female rate (%)	31.7	66.4	<0.0001
PS > 1 rate (%)	5.8	2.7	<0.0001
wMRs, n (%)	50 (62.5)	30 (37.5)	<0.0001
CTs, n (%)	25 (31.3)	55 (68.8)	<0.0001

PS: performance status; wMR (web-based monitoring registry); CT (clinical trial).

Table 3: **Distribution of variables according to Cluster analysis.**

imbalance was, on average, less evident although female patients tended to be under-represented in clinical trials in some specific tumor types (like head & neck and upper gastrointestinal cancer). Moreover, our multivariate approach also confirmed that clusters characterized by differences in patient age, sex and ECOG PS are characterized as well by CTs/wMRs ratio, supporting the hypothesis of structural differences between the CTs and wMRs patients. These findings suggest that caution is required in the process of generalization of clinical trial data to clinical practice, due to the differences between the patients in the two settings.

The reasons behind such differences do not seem simplistically related to inclusion/exclusion criteria, although these represent a relevant part of the phenomenon.[19] Indeed, upper age limits are generally absent in registrative trials, but nevertheless elderly patients are under-represented, possibly because of other criteria indirectly leading to their exclusion (eg. comorbidities, concomitant treatments for chronic conditions, difficulties in hospital access, etc); such exclusion might even occur in randomised trials of potentially less toxic drugs, when the experimental treatment has to be compared with a more toxic standard treatment (typically cytotoxic chemotherapy) that becomes the driver of patients' selection. Finally, ageism could induce doctors to prefer younger patients for experimental protocols in order to give them more therapeutic options and potentially more effective treatments.[20] In any case, we acknowledge that age differences might be slightly inflated in Italy, which is the European country with the highest median age of the citizens.[21] Contrasting this phenomenon is challenging, because a prudential approach in clinical trials justifies the exclusion of more vulnerable populations as a protective strategy, mostly in early phase studies where there is still poor knowledge on safety of experimental treatments. Hopefully, a wider inclusion of elderly patients in early phase (for example with dedicated cohorts in the dose-finding and/or expansion sections of phase I trials) might generate a higher confidence of protocol developer and collaborating clinicians in avoiding under-selection of elderly patients in later registrative trials.[22,23] The same might apply to patients with deteriorated performance status, particularly when such condition is a direct consequence of tumor dissemination rather than of other general conditions.[24,25] For both under-representation of elderly and PS > 1, planning dedicated trials might equilibrate the evidence.[26] This principle should be applied independently of the type of drug, considering that (in contrast with a diffuse perception) the under-representation of elderly and PS2 patients is similar among the classes of drugs considered in this analysis.

Overall, we found no significant evidence that female patients are under-represented in CTs as compared to clinical practice. However, a number of specific tumor types (lung, head & neck or upper gastrointestinal cancer and non-melanoma skin cancer) suffered an imbalance in sex; indeed, with the exception of lung cancer where females were under-represented in registries, they were prevalently under-represented in trials, consistently with previously reported data.[16] Even if the overall figure of no imbalance might be considered good news in terms of generalizability, specific attention has to be paid to single indications where imbalance exists and safety/effectiveness profiles of drugs might vary according to sex.[27]

For the future, further research is needed to quantify how much differences found in baseline characteristics do translate in outcome differences. In the worst-case scenario, all the differences we described might dilute the effectiveness of new treatments in clinical practice, and lead to more frequent and possibly severe side effects. Therefore, strategies to improve inclusiveness of clinical trials should be encouraged to alleviate selection biases; in addition, real-world data should be collected together with registration trial to integrate the evidence.[28–30]

In conclusion, the analysis of AIFA monitoring registries supports that cancer patients enrolled in clinical trials do only partially represent those who have been treated in Italy in clinical practice. Even if randomized clinical trials remain the gold-standard instruments to find out new effective drugs, their degree of inclusiveness has to be increased to ensure generalizability of results to the real-world population.

Contributors

Giordano Domenico Beretta, Carmine Pinto, Livio Blasi, Saverio Cinieri, Luigi Cavanna, Massimo Di Maio, Pierluigi Russo and Francesco Perrone planned the study. Maria Lucia Iacovino, Andrea Caglio, Andrea Canciello, and Flavio Salerno collected data on clinical trials. Simone Celant, Luca Tomassini, Pier Paolo Olimpieri, Susanna Di Segni, Antonella Sferrazza, and Pierluigi Russo collected data from AIFA registries. Simone Celant, Maria Lucia Iacovino, Luca Tomassini, Laura Arenare, Pier Paolo Olimpieri, Maria Carmela Piccirillo, Massimo Di Maio, Pierluigi Russo and Francesco Perrone performed analyses. Maria Lucia Iacovino, Simone Celant, Luca Tommasini, Laura Arenare, Pier Paolo Olimpieri, Maria Carmela Piccirillo, Massimo Di Maio, Pierluigi Russo and Francesco Perrone wrote the first draft of manuscript; all authors contributed to the interpretation of results, and approved the final version of the manuscript and accept responsibility to submit for publication.

Data sharing statement

A dataset with all information retrieved from the clinical trials included in this study is available online (https://doi.org/10.5281/zenodo.10491020).

Declaration of interests

SC declares: support for attending meetings and/or travel from Novartis and Roche; payment for participation on advisory board from EliLilly and AstraZeneca; leadership in scientific society: President of AIOM 2021–2023.

MCP declares: institutional grants or contracts from Bayer, AstraZeneca, Roche; payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Astellas, Pfizer, Ipsen, AstraZeneca; support for attending meetings and/or travel from Menarini.

MDM declares: institutional grants or contracts from Tesaro/GSK, Beigene, Exelixis, MSD, Pfizer and Roche; fees for consulting or participation in advisory board from AstraZeneca, Boehringer Ingelheim, Janssen, Merck Sharp & Dohme, Novartis, Pfizer, Roche, GlaxoSmithKline, Amgen, Merck, Takeda.

FP declares: institutional grants or contracts from Roche, Bayer, AstraZeneca, Pfizer, Incyte, Tesaro/GSK, Merck; consulting fees from Bayer, Pierre Fabre, Astra Zeneca, Incyte, Ipsen, Clovis, Astellas, Sanofi, Roche, Pfizer; leadership in scientific society: President of AIOM 2023–2025.

All the other Authors declare no potential conflict of interest.

Acknowledgements

The present work has been carried out through a collaboration between AIFA and AIOM (Associazione Italiana Oncologia Medica), CIPOMO (Collegio Italiano dei Primari di Oncologia Medica Ospedalieri) and FICOG (Federation of Italian Cooperative Oncology Groups), for the scientific development of the data extracted from AIFA monitoring registries.

Although the AIFA monitoring registries operate within a regulatory framework that provides for the mandatory collection of data from real clinical practice on specific treatments reimbursed by the Italian National Health Service (NHS), the authors are, however, grateful for the contribution of all the physicians and pharmacists who filled the registries.

We acknowledge all the efforts made by present and past components of AIFA Monitoring Registries Group in developing and maintaining the registries. Present composition of the AIFA's Monitoring Registries Group (in chronological order): Daniele Ricagni, Susanna Valentini, Alessia Cirilli, Valeria Angelini, Valentina Summa, Laura Veo, Annalisa Guarcello, Carlo Montecchiani, Annarita Saccia.

The views expressed in this work are personal and may not be understood or quoted as being made on behalf of or reflective of the position of the Italian Medicines Agency or of one of their committees or working parties.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanepe.2024.100912.

References

1 (CHMP) CfMPfHU. Guideline on the clinical evaluation of anticancer 5 medicinal products. https://www.ema.europa.eu/en/evaluation-anticancer-medicinal-products-man-scientific-guideline; 2020.
2 FDA. *Clinical trial endpoints for the approval of cancer drugs and biologics guidance for industry*; 2018. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics.
3 Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA*. 2004;291(22):2720–2726.
4 Scher KS, Hurria A. Under-representation of older adults in cancer registration trials: known problem, little progress. *J Clin Oncol*. 2012;30(17):2036–2038.
5 Kim ES, Bruinooge SS, Roberts S, et al. Broadening eligibility criteria to make clinical trials more representative: American society of clinical oncology and friends of cancer research joint research statement. *J Clin Oncol*. 2017;35(33):3737–3744.
6 Benbow JH, Rivera DR, Lund JL, Feldman JE, Kim ES. Increasing inclusiveness of patient-centric clinical evidence generation in oncology: real-world data and clinical trials. *Am Soc Clin Oncol Educ Book*. 2022;42:1–11.
7 EMA. *Gender considerations in the conduct of clinical trials*; 2005. http://www.emea.eu.int/pdfs/human/ich/391605en.pdf.
8 EMA. *ICH E7 Studies in support of special populations: geriatrics - scientific guideline*; 2010. https://www.ema.europa.eu/en/ich-e7-studies-support-special-populations-geriatrics-scientific-guideline#current-effective-version-section.
9 Concato J, Corrigan-Curay J. Real-world evidence - where are we now? *N Engl J Med*. 2022;386(18):1680–1682.
10 Di Maio M, Perrone F, Conte P. Real-world evidence in oncology: opportunities and limitations. *Oncologist*. 2020;25(5):e746–e752.
11 Montilla S, Xoxi E, Russo P, Cicchetti A, Pani L. Monitoring registries at Italian Medicines agency: fostering access, guaranteeing sustainability. *Int J Technol Assess Health Care*. 2015;31(4):210–213.
12 Vincenzi B, Napolitano A, Comandone A, et al. Trabectedin use in soft-tissue sarcoma patients in a real-world setting: data from an Italian national drug-access registry. *Int J Cancer*. 2023;152(4):761–768.
13 Torti C, Olimpieri PP, Bonfanti P, et al. Real-life comparison of mortality in patients with SARS-CoV-2 infection at risk for clinical progression treated with molnupiravir or nirmatrelvir plus ritonavir during the Omicron era in Italy: a nationwide, cohort study. *Lancet Reg Health Eur*. 2023;31:100684.
14 Hartigan JA, Wong MA. Algorithm as 136: a K-means clustering algorithm. *J R Stat Soc Ser C Appl Stat*. 1979;28(1):100–108.
15 Caliński T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat*. 1974;3(1):1–27.
16 Varma T, Wallach JD, Miller JE, et al. Reporting of study participant demographic characteristics and demographic representation in premarketing and postmarketing studies of novel cancer therapeutics. *JAMA Netw Open*. 2021;4(4):e217063.
17 Karim S, Booth CM. Effectiveness in the absence of efficacy: cautionary tales from real-world evidence. *J Clin Oncol*. 2019;37(13):1047–1050.
18 Tan YY, Papez V, Chang WH, Mueller SH, Denaxas S, Lai AG. Comparing clinical trial population representativeness to real-world populations: an external validity analysis encompassing 43 895 trials and 5 685 738 individuals across 989 unique drugs and 286 conditions in England. *Lancet Healthy Longev*. 2022;3(10):e674–e689.
19 Riner AN, Girma S, Vudatha V, et al. Eligibility criteria perpetuate disparities in enrollment and participation of black patients in pancreatic cancer clinical trials. *J Clin Oncol*. 2022;40(20):2193–2202.
20 Monfardini S, Perrone F, Balducci L. Pitfalls in oncogeriatrics. *Cancers*. 2023;15:2910 [accessed].
21 Population structure and ageing. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing&stable=1#Median_age_is_highest_in_Italy_and_lowest_in_Cyprus. Accessed December 15, 2023.
22 Charton E, Baldini C, Fayet Y, et al. Inequality factors in access to early-phase clinical trials in oncology in France: results of the EGALICAN-2 study. *ESMO Open*. 2023;8(4):101610.
23 Ho J, Pond GR, Newman C, et al. Barriers in phase I cancer clinical trials referrals and enrollment: five-year experience at the Princess Margaret Hospital. *BMC Cancer*. 2006;6:263.
24 Facchinetti F, Mazzaschi G, Barbieri F, et al. First-line pembrolizumab in advanced non-small cell lung cancer patients with poor performance status. *Eur J Cancer*. 2020;130:155–167.
25 Magnuson A, Bruinooge SS, Singh H, et al. Modernizing clinical trial eligibility criteria: recommendations of the ASCO-friends of cancer research performance status work Group. *Clin Cancer Res*. 2021;27(9):2424–2429.
26 Gridelli C, Aapro M, Ardizzoni A, et al. Treatment of advanced non-small-cell lung cancer in the elderly: results of an international expert panel. *J Clin Oncol*. 2005;23(13):3125–3137.
27 Conforti F, Pala L, Bagnardi V, et al. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. *Lancet Oncol*. 2018;19(6):737–746.
28 Harvey RD, Miller TM, Hurley PA, et al. A call to action to advance patient-focused and decentralized clinical trials. *Cancer*. 2024;130:1193.
29 Leary A, Besse B, Andre F. The need for pragmatic, affordable, and practice-changing real-life clinical trials in oncology. *Lancet*. 2023;403:406.
30 Perrone F, Di Liello R, Gargiulo P, et al. The opportunity of patient-journey studies for academic clinical research in oncology. *BMJ Open*. 2021;11(9):e052871.