

BMJ Open Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry

Xinji Zhang, Yuan Zhang, Xiaofei Ye, Xiaojing Guo, Tianyi Zhang, Jia He

To cite: Zhang X, Zhang Y, Ye X, *et al.* Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry. *BMJ Open* 2016;**6**: e010643. doi:10.1136/bmjopen-2015-010643

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-010643>).

XZ and YZ contributed equally.

Received 23 November 2015
Revised 3 August 2016
Accepted 30 August 2016



CrossMark

Department of Health Statistics, Second Military Medical University, Shanghai, China

Correspondence to
Professor Jia He;
hejia63@yeah.net

ABSTRACT

Objective: Phase IV trials are often used to investigate drug safety after approval. However, little is known about the characteristics of contemporary phase IV clinical trials and whether these studies are of sufficient quality to advance medical knowledge in pharmacovigilance. We aimed to determine the fundamental characteristics of phase IV clinical trials that evaluated drug safety using the ClinicalTrials.gov registry data.

Methods: A data set of 19 359 phase IV clinical studies registered in ClinicalTrials.gov was downloaded. The characteristics of the phase IV trials focusing on safety only were compared with those evaluating both safety and efficacy. We also compared the characteristics of the phase IV trials in three major therapeutic areas (cardiovascular diseases, mental health and oncology). Multivariable logistic regression was used to evaluate factors associated with the use of blinding and randomisation.

Results: A total of 4772 phase IV trials were identified, including 330 focusing on drug safety alone and 4392 evaluating both safety and efficacy. Most of the phase IV trials evaluating drug safety (75.9%) had enrolment <300 with 96.5% <3000. Among these trials, 8.2% were terminated or withdrawn. Factors associated with the use of blinding and randomisation included the intervention model, clinical specialty and lead sponsor.

Conclusions: Phase IV trials evaluating drug safety in the ClinicalTrials.gov registry were dominated by small trials that might not have sufficient power to detect less common adverse events. An adequate sample size should be emphasised for phase IV trials with safety surveillance as main task.

INTRODUCTION

Drug adverse reaction is a major global health concern accounting for more than 2 million injuries, hospitalisations, and deaths each year in the USA alone,¹ and associated with billions of US dollars in cost every year in the developed countries.² Although rigorous

Strengths and limitations of this study

- We provided a comprehensive descriptive assessment of the current portfolio of phase IV clinical trials evaluating drug safety in the ClinicalTrials.gov registry.
- We employed logistic regression models to determine the factors associated with the use of blinding and randomisation in phase IV clinical trials which evaluated drug safety.
- We followed a strict analysis process that was widely used in analysing the data from ClinicalTrials.gov to arrive at convincing results.
- Some clinical trials were not registered in ClinicalTrials.gov.
- There were some unavoidable missing data for certain data fields which might introduce some bias into the results.

premarketing studies are required for all new drugs,^{3 4} the safety profile of a drug at the time of regulatory approval is often incomplete due to some characteristics of phase I–III trials such as limited sample sizes, short duration and strict inclusion/exclusion criteria.⁵ Approximately 20% of drugs acquired new black box warnings postmarketing, and 4% of the drugs were ultimately withdrawn for safety reasons.⁶ In 2007, the Food and Drug Administration was authorised by the Food and Drug Administration Amendment Act (FDAAA)⁷ to require postmarketing clinical trials to address safety concerns regarding a given drug. Compared to premarketing phase I–III trials, phase IV studies evaluate drug safety in a real-world setting, which may provide evidence to ensure or further refine the safety of approved drugs.^{5 8 9} However, little is known about the characteristics of contemporary phase IV clinical trials and whether these studies are of sufficient quality to advance medical knowledge in pharmacovigilance.

ClinicalTrials.gov is a public trial registry established by the National Library of Medicine on behalf of the National Institutes of Health (NIH) and was first launched in February 2000.¹⁰ Since 2005, the International Committee of Medical Journal Editors has implemented a policy requiring the registration of clinical trials as a prerequisite for publication.¹¹ In addition, as of 2007, sponsors or their designees are obliged by FDAAA to register trials and report key data elements and basic trial results at ClinicalTrials.gov.¹² Hence, the ClinicalTrials.gov registry is considered to be the most comprehensive source for clinical trial information worldwide.^{13–15} Harnessing this expansive resource will enable us to gain a deeper understanding of postmarketing drug safety surveillance.

The objective of our study is to examine the characteristics of registered phase IV clinical trials regarding drug safety and identify areas which require greater attention. We focus on data elements that are desirable for generating reliable evidence from clinical trials, including sample size and factors associated with the use of randomisation and blinding.

METHODS

Data source

Our analysis was restricted to phase IV clinical trials registered with ClinicalTrials.gov between 2004 and 2014. A data set of 19 359 phase IV clinical studies registered with ClinicalTrials.gov was downloaded and locked from the website on 18 March 2015. A database was designed to facilitate analysis.^{15 16}

Study selection

Two authors (XZ and YZ) selected the eligible studies and summarised their results independently. [Figure 1](#) shows the complete process of selection. Our analysis was restricted to phase IV clinical trials registered between 1 January 2004 and 31 December 2014 (n=18 642) according to the first date submitted to ClinicalTrials.gov. Interventional studies using drugs were identified by searching the sections of ‘study type’ and ‘intervention’ on ClinicalTrials.gov. Observational studies (n=981), expanded-access studies (n=10) and other studies that investigated ‘medical devices’, ‘vaccines’ or other products were removed (n=5878). On ClinicalTrials.gov, the ‘End point Classification’ section indicated the primary end point of the study, such as bio-equivalence, pharmacokinetics, safety and efficacy, and others. Additionally, based on the information in the ‘Primary Purpose’ section, studies could be divided into different groups: ‘Treatment’, ‘Prevention’, ‘Diagnostic’, ‘Supportive Care’, ‘Screening’, ‘Health Services Research’, ‘Basic Science’, ‘Educational/Counseling/Training’ and missing. We further identified studies whose purposes were ‘Treatment’ and primary end points were ‘Safety study’ or ‘Safety/efficacy study’ using ‘Primary Purpose’ and ‘End point Classification’

sections. Finally, 4722 eligible phase IV trials assessing drug safety alone or both safety and efficacy were included in our analysis.

The included trials were then categorised into three groups by different clinical specialties—mental health, oncology and cardiovascular diseases, using the information in the ‘Conditions’ section and the classification of studies both provided by ClinicalTrials.gov via matching the NCT number of each study.

Data collection

Trial data were reported by the trial sponsors or investigators, as required by the ClinicalTrials.gov registry.¹⁷ Each record contained a set of data elements describing the study’s conditions, enrolment, study design, eligibility criteria, location, sponsor and other protocol information.

The methods of defining derived variables have been described previously^{15 18} and are briefly summarised below. All trials were divided into six different groups by the funding sources according to the information in the ‘Sponsor_Collaborators’ and ‘Funded_By’ sections: NIH, industry, other, US federal (excluding NIH), university/college, hospital and other sources. The funding source was defined as the NIH if the lead sponsor or any collaborators were from the NIH, and the lead sponsor was not from industry. It was defined as industry if the lead sponsor was from industry or if any collaborators were from industry and none from the NIH. It was defined as from US federal sources if the sponsors were from US Federal only and none of the collaborators were from industry or NIH. The funding source was defined as ‘hospital’ if the lead sponsor was from a hospital or similar institutions and no collaborators were from industry, the NIH or a US federal. It was defined as ‘university/college’ if the lead sponsor was from a university, college or similar institutions and collaborator was not from industry, NIH, a US federal institution or hospitals. For the remaining studies, the funding source was defined as other sources. The start dates of trials could be obtained from the ‘Start_Date’ section. Information on the appointment of a data monitoring committee (DMC) became available since April 2007, and was not a required field.¹⁸ Thus, the DMC information was not considered in our study. The classifications of other variables were based on the information in the corresponding fields from ClinicalTrials.gov.

When a data field was incomplete, a web search (ClinicalTrials.gov) was conducted to find the missing information for the trial. If the information was not available on the website either, this field was identified as NA (not applicable) or missing. For studies reporting an interventional model of single group and the number of groups as 1, we inferred the value of allocation as non-randomised and the value of blinding as open if the information was missing.¹⁵ In addition, the allocation or blinding was reported as ‘Uncertain’ if single-arm trials were registered as randomised or blind.

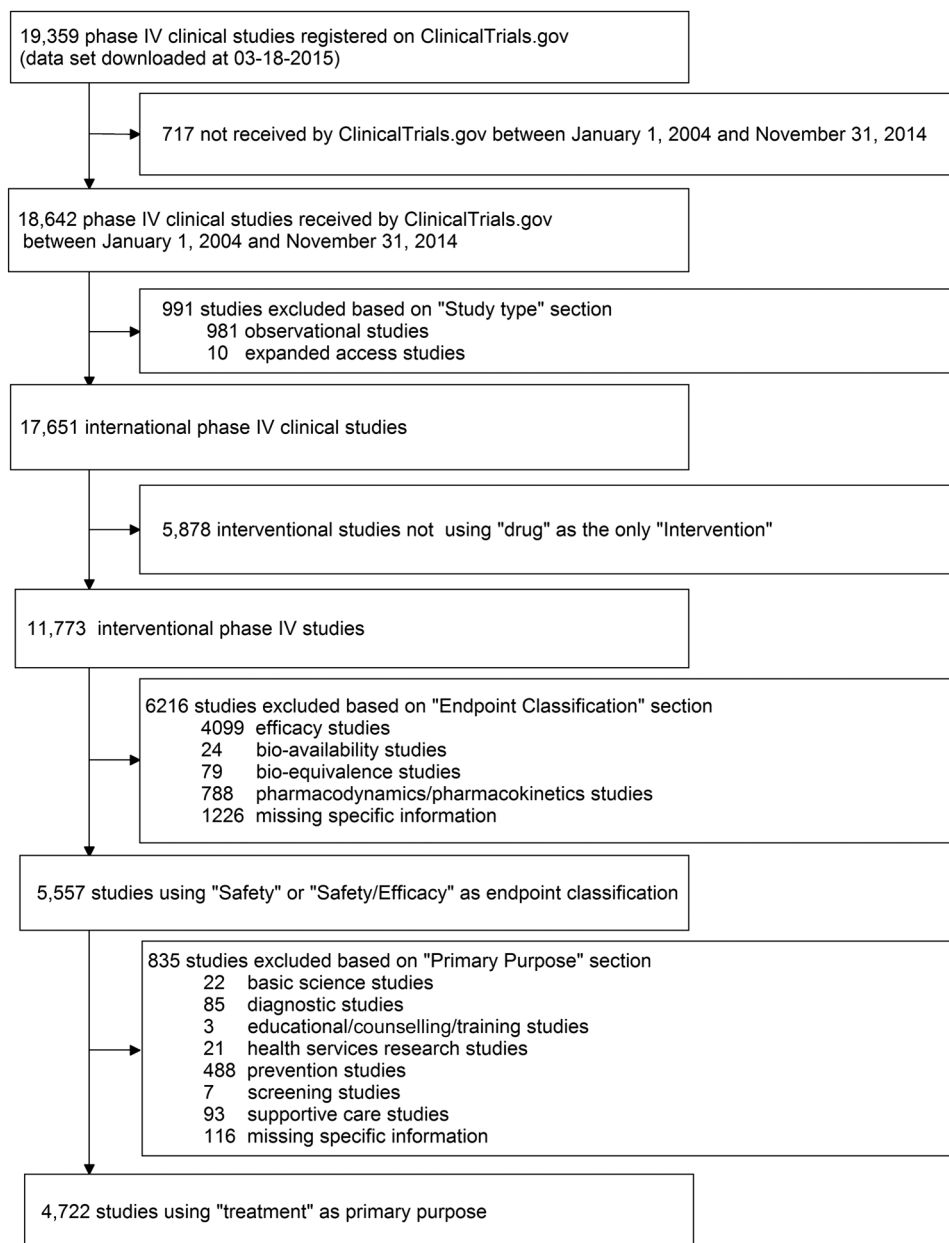


Figure 1 Flow chart of inclusion and exclusion.

Statistical analysis

The characteristics of the trials were assessed overall, by two end point classifications (safety only and safety/efficacy) and by three clinical specialties (mental health, oncology and cardiovascular diseases). The assessments included the study status, enrolment, intervention model, funding source and so on. The registration timeline of a trial was determined by comparing the date first received by ClinicalTrials.gov with the start date of the trial.

According to the binomial and Poisson distributions, if the adverse events (AEs) have a probability of occurrence 1%, 0.5% or 1%, the enrolment should be larger than 300, 600 or 3000, respectively (table 1), in order for the investigators to have a 95% chance to observe at least 1 case of AEs.¹⁹ Hence, we divided the included

Table 1 Numbers of patients necessary to enrol

Expected incidence of adverse reaction	Numbers of patients to enrol for detecting at least 1 event
1 in 100	300
1 in 200	600
1 in 1000	≥3000

trials into five types: trials with sample size <300, between 300 and 599, between 600 and 2999 and 3000 or above and missing. Frequencies and percentages were provided for categorical characteristics; medians and IQRs were provided for continuous characteristics.

Logistic regression analysis was used to evaluate factors associated with the use of randomisation and blinding.

Table 2 Characteristics of included trials by different end point classifications

	Number (%)		
	All, 2004–2014 N=4722	Safety alone N=330	Safety/efficacy N=4392
Overall status*			
Not yet recruiting	196 (4.2)	8 (2.4)	188 (4.3)
Recruiting	941 (19.9)	62 (18.8)	879 (20)
Completed	2858 (60.5)	193 (58.5)	2665 (60.7)
Suspended	17 (0.4)	4 (1.2)	13 (0.3)
Terminated	304 (6.4)	24 (7.3)	280 (6.4)
Withdrawn	87 (1.8)	8 (2.4)	79 (1.8)
Active, not recruiting	274 (5.8)	30 (9.1)	244 (5.6)
Enrolling by invitation	45 (1.0)	1 (0.3)	44 (1.0)
Enrolment, median (IQR),	104.0 (48.0–258.0)	120.0 (45.0–392.0)	103.0 (48.0–251.5)
1–299	3585 (75.9)	226 (68.5)	3359 (76.5)
300–599	629 (13.3)	43 (13.0)	586 (13.3)
600–2999	344 (7.3)	37 (11.2)	307 (7.0)
≥3000	57 (1.2)	13 (3.9)	44 (0.01)
Missing	107 (2.3)	11 (3.3)	96 (2.2)
Intervention model			
Crossover assignment	271 (5.7)	28 (8.5)	243 (5.5)
Single group assignment	1276 (27.0)	138 (41.8)	1138 (25.9)
Parallel assignment	3116 (66.0)	163 (49.4)	2953 (67.2)
Factorial assignment	52 (1.1)	1 (0.3)	51 (1.2)
Missing	7 (0.1)	0 (0.0)	7 (0.2)
Allocation			
Randomised	3310 (70.1)	179 (54.2)	3131 (71.3)
Non-randomised	1252 (26.5)	135 (40.9)	1117 (25.4)
Missing	35 (0.7)	8 (2.4)	27 (0.6)
Uncertain	125 (2.7)	8 (2.4)	117 (2.7)
Blinding			
Double-blind	1690 (35.8)	87 (26.4)	1603 (36.5)
Single blind	302 (6.4)	20 (6.1)	282 (6.4)
Open label	2620 (55.5)	218 (66.1)	2402 (54.7)
Missing	6 (0.1)	0 (0.0)	6 (0.1)
Uncertain	104 (2.2)	5 (1.5)	99 (2.3)
Sex,%			
Women only	337 (7.1)	19 (5.8)	318 (7.2)
Men only	159 (3.4)	20 (6.1)	139 (3.2)
Both	4224 (89.5)	291 (88.2)	3933 (89.5)
Missing	2 (0.0)	0 (0.0)	2 (0.0)
Included children (<18 years)	762 (16.1)	72 (21.8)	690 (15.7)
Excluded elderly (>65 years)	1362 (28.8)	89 (27.0)	1273 (29.0)
Lead sponsor			
Industry	2711 (57.4)	229 (69.4)	2482 (56.5)
NIH	97 (2.1)	4 (1.2)	93 (2.1)
US Federal	30 (0.6)	0 (0.0)	30 (0.7)
Hospital and similar institutions	682 (14.4)	38 (11.5)	644 (14.7)
Universities and similar institutions	758 (16.1)	37 (11.2)	721 (16.4)
Other	444 (9.4)	22 (6.7)	422 (9.6)
Region†			
Africa	168 (3.6)	14 (4.2)	154 (3.5)
Asia and Pacific	1332 (28.2)	82 (24.8)	1250 (28.5)
Central and South America	324 (6.9)	33 (10)	291 (6.6)
Europe	1250 (26.5)	106 (32.1)	1144 (26)
Middle East	239 (5.1)	21 (6.4)	218 (5)
North America	1626 (34.4)	146 (44.2)	1480 (33.7)
Missing	506 (10.7)	31 (9.4)	475 (10.8)

Continued

Table 2 Continued

	Number (%)		
	All, 2004–2014 N=4722	Safety alone N=330	Safety/efficacy N=4392
Study registration			
Start before submission	131 (2.8)	7 (2.1)	124 (2.8)
Start after submission	4591 (97.2)	323 (97.9)	4268 (97.2)

*‘Recruiting’, ‘Not yet recruiting’ refer to studies that are currently recruiting participants, or will be recruiting participants in the future, respectively. ‘Active, not recruiting’, ‘Completed’ refer to studies that are no longer recruiting participants because they have enough participants already or they are completed, respectively. ‘Terminated’, ‘Suspended’, ‘Withdrawn’ mean they studies that have been stopped for some reasons.

†Percentages may not sum to 100%, as categories are not mutually exclusive.

A full model containing eight characteristics was developed and adjusted ORs with Wald 95% CIs were calculated for these factors. The factors assessed included funding source, primary purpose, number of participants, trial specialty (yes/no), trial start year before or after the publication of FDAAA in 2007 and end point classification (safety/efficacy study or safety study). Single-arm trials or studies with any of the data elements missing were excluded from the regression analysis.

SAS V.9.2 (SAS Institute) was used for all statistical analyses.

RESULTS

From 1 January 2004, to 13 December 2014, 18 642 phase IV trials were registered at ClinicalTrials.gov. Of these trials, 4722 phase IV trials related to drug safety were included in our study. Figure 1 shows the search process. The number of trials evaluating safety alone was 330, which was lesser than the number of trials evaluating both safety and efficacy (n=4392). A total of 594 trials (12.6%) focused on mental health diseases, 251 trials (5.3%) focused on oncology and 601 trials (12.7%) on cardiovascular diseases.

The basic characteristics of all inclusive 4722 trials registered with ClinicalTrials.gov are shown in table 2. The median number of participants per trial was 104.0 (IQR: 48.0–258.0). About 72.7% of these phase IV trials conducted randomisation and 44.4% used blinding (including double-blind and single-blind). We also noted that 8.3% (n=391) of these phase IV trials were ‘terminated’ or ‘withdrawn’, which means these trials were stopped for some reasons. Most of the 4722 studies were small (median enrolment: 35.5; IQR: 11.0–104.3). The most common research sites in these phase IV trials were from North America, Asia and the Pacific and Europe, which accounted for 34.4%, 28.2% and 26.5%, respectively.

Of the total phase IV trials 68.5% evaluating drug safety alone had enrolment of <300 patients, and only 3.9% (n=13) of them enrolled more than 3000. The median number of participants per trial was 104.0 (IQR: 45.0–392.0). The average sample size of the phase IV trials assessing both safety and efficacy was similar, with a

median enrolment of 103.0 (IQR: 48.00–251.5). Compared with studies evaluating both safety and efficacy, phase IV trials focused on drug safety only showed larger proportion of studies using single group assignment (41.8% vs 25.9%) and a small proportion using randomisation (56.7% vs 74.0%). However, the difference in the proportion of studies using blinding was relatively small between trials focusing on safety only and those assessing safety/efficacy (34.0% vs 42.8%).

Table 3 showed the characteristics of the phase IV trials in three major therapeutic areas (cardiovascular, oncology and mental health). The cardiovascular diseases trials accounted for the most among these three categories (n=601, 12.7%). Also cardiovascular trials had more enrolment (median: 163; IQR: 70.0–400.0) than oncology trials (median: 100.0; IQR: 48.0–200.0) and mental health trials (median: 88.0; IQR: 40.0–226.0). Randomisation was less common in oncology trials than cardiovascular trials and mental health trials (43.0% vs 81.4% for cardiovascular and 67.5% for mental health). The difference in the use of blinding was similar (17.5% for oncology trials vs 46.2% for cardiovascular trials and 57.2% for mental health trials). As women-only trials, they accounted for the largest group for oncology trials at 13.5% compared to 1.3% for cardiovascular trials and 2.3% for mental health trials. It was noteworthy that nearly two-thirds of mental health trials (65.0%) excluded elderly patients. Geographical differences were also apparent. Mental health trials had the largest proportion of studies with at least one North American research site (52.9%), whereas, oncology trials showed the largest proportion of studies with at least one Asia and Pacific research site (42.2%). The NIH sponsored more mental health trials (8.9% vs 1.0% for cardiovascular trials and 0.4% for oncology trials).

Table 4 shows the results of the regression analyses. These analyses compared the trial characteristics that are related to the use of blinding and randomisation. A total of 1276 single-arm trials and 78 studies with any of the data elements missing were excluded from the regression analysis. Hence, there were 3361 trials which were considered in the regression model. Of these trials, 1950 (58.02%) studied were blind and 3234 (96.22%)

Table 3 Characteristics of included trials in different clinical specialties

	Number (%)		
	Cardiovascular diseases N=601	Oncology N=251	Mental Health N=594
Overall status*			
Not yet recruiting	39 (6.5)	7 (2.8)	16 (2.7)
Recruiting	129 (21.5)	84 (33.5)	106 (17.8)
Completed	331 (55.1)	106 (42.2)	404 (68)
Suspended	2 (0.3)	2 (0.8)	0 (0.0)
Terminated	39 (6.5)	13 (5.2)	30 (5.1)
Withdrawn	16 (2.7)	5 (2.0)	10 (1.7)
Active, not recruiting	42 (7)	28 (11.2)	26 (4.4)
Enrolling by invitation	3 (0.5)	6 (2.4)	2 (0.3)
Enrolment, median (IQR)	163.0 (70.0–400.0)	100.0 (48.0–200.0)	88.0 (40.0–226.0)
1–299	391 (65.1)	205 (81.7)	475 (80)
300–599	83 (13.8)	24 (9.6)	80 (13.5)
600–2999	90 (15)	14 (5.6)	27 (4.5)
≥3000	17 (2.8)	2 (0.8)	5 (0.8)
Missing	20 (3.3)	6 (2.4)	7 (1.2)
Intervention model			
Crossover assignment	23 (3.8)	5 (2.0)	35 (5.9)
Single group assignment	10 (1.7)	4 (1.6)	8 (1.3)
Parallel assignment	451 (75.0)	104 (41.4)	359 (60.4)
Factorial assignment	115 (19.1)	138 (55.0)	191 (32.2)
Missing	2 (0.3)	0 (0.0)	1 (0.2)
Allocation			
Randomised	469 (78.0)	101 (40.2)	390 (65.7)
Non-randomised	108 (18.0)	139 (55.4)	190 (32.0)
Missing	4 (0.7)	4 (1.6)	3 (0.5)
Uncertain	20 (3.3)	7 (2.8)	11 (1.9)
Blinding			
Double-Blind	213 (35.8)	33 (13.2)	292 (49.2)
Single blind	51 (8.5)	7 (2.8)	31 (5.2)
Open label	322 (53.6)	207 (82.5)	252 (42.4)
Missing	1 (0.2)	0 (0)	2 (0.3)
Uncertain	14 (2.3)	4 (1.6)	17 (2.9)
Sex, %			
Women only	8 (1.3)	34 (13.5)	13 (2.2)
Men only	9 (1.5)	21 (8.4)	31 (5.2)
Both	584 (97.2)	196 (78.1)	550 (92.6)
Missing	0 (0.0)	0 (0.0)	0 (0.0)
Included children (<18 years)	33 (5.5)	35 (13.9)	112 (18.9)
Excluded elder (>65 years)	47 (7.8)	24 (9.6)	386 (65.0)
Lead sponsor			
Industry	305 (50.7)	148 (59.0)	360 (60.6)
NIH	6 (1.0)	1 (0.4)	53 (8.9)
US Federal	2 (0.3)	0 (0.0)	3 (0.5)
Hospitals and similar institutions	119 (19.8)	39 (15.5)	55 (9.3)
Universities and similar institutions	108 (18.0)	32 (12.7)	80 (13.5)
Other	61 (10.1)	31 (12.4)	43 (7.2)
Region†			
Africa	20 (3.3)	13 (5.2)	13 (2.2)
Asia and Pacific	210 (34.9)	106 (42.2)	137 (23.1)
Central and South America	26 (4.3)	14 (5.6)	41 (6.9)
Europe	167 (27.8)	82 (32.7)	76 (12.8)
Middle East	24 (4.0)	19 (7.6)	25 (4.2)
North America	172 (28.6)	59 (23.5)	314 (52.9)
Missing	57 (9.5)	26 (10.4)	56 (9.4)

Continued

Table 3 Continued

	Number (%)		
	Cardiovascular diseases N=601	Oncology N=251	Mental Health N=594
Study registration			
Start before submission	27 (4.5)	10 (4.0)	12 (2.0)
Start after submission	574 (95.5)	241 (96.0)	582 (98.0)

*'Recruiting', 'Not yet recruiting' mean studies that are currently recruiting participants, or will be recruiting participants in the future, respectively. 'Active, not recruiting', 'Completed' mean studies that are no longer recruiting participants because they have enough participants already or they are completed, respectively. 'Terminated', 'Suspended', 'Withdrawn' mean they studies that have been stopped for some reasons.

†Percentages may not sum to 100%, as categories are not mutually exclusive.

were randomised. Different clinical specialties could affect the use of blinding and randomisation. Oncology trials were less likely to use both blinding (adjusted OR: 0.33; 95% CI 0.18 to 0.63) and randomisation (adjusted OR: 0.42; 95% CI 0.28 to 0.63). Mental health trials were more likely to implement blinding (adjusted OR: 3.35; 95% CI 2.56 to 4.38). Compared with the trials in which industry was the lead sponsor, the trials funded by universities or similar institutions were more likely to use blinding (adjusted OR: 1.32; 95% CI 1.08 to 1.60).

DISCUSSION

This study provided a descriptive assessment of the current portfolio of phase IV clinical trials evaluating drug safety. The characteristics of phase IV trials with different end point classifications and clinical specialties were compared. We also analysed the factors associated with trial quality. Thus, this study presented a unique opportunity to evaluate the landscape of phase IV trials

related to drug safety and to identify areas of relative strength or weakness.

Small sample size was the greatest concern in phase IV trials involving the safety surveillance of an approved drug. Small phase IV trials might be used to evaluate the effectiveness of a given drug in a special patient subgroup, or in special situations.⁵ However, our study included only phase IV trials with 'safety' as an end point and most of these trials (77.6%) had an enrolment of <300. In the phase IV trials with safety as the primary end point, the average sample size was only 104. Thus, these small trials might not have sufficient power to detect AEs, especially less common AEs.¹⁹ Paying greater attention to the quality of phase IV trials may facilitate postmarketing drug safety surveillance. For trials with safety assessment as their primary purpose, the sample size should be estimated according to the probability of occurrence expected for each AE. For example, to observe an AE with an occurrence probability of 1.5%, the China Food and Drug Administration

Table 4 Regression analyses of included trials and the reported use of blinding and randomisation

Variable	Blinding*		Randomisation†	
	Adjusted OR (95% CI)	p Value	Adjusted OR (95% CI)	p Value
Lead sponsor (vs industry)				
NIH	0.92 (0.56 to 1.51)	0.746	0.91 (0.27 to 3.08)	0.884
Other	1.24 (0.96 to 1.59)	0.094	1.19 (0.58 to 2.42)	0.638
US federal	0.80 (0.34 to 1.84)	0.594	0.35 (0.08 to 1.53)	0.162
Hospital or similar institutions	1.02 (0.84 to 1.23)	0.877	0.69 (0.43 to 1.09)	0.111
University or similar institutions	1.32 (1.08 to 1.60)	0.006	0.93 (0.57 to 1.53)	0.781
Study size (vs <300)				
300–599	0.93 (0.77 to 1.13)	0.472	1.11 (0.65 to 1.89)	0.706
≥600	0.83 (0.65 to 1.06)	0.132	0.87 (0.47 to 1.59)	0.639
Intervention model (vs parallel assignment)				
Crossover assignment	1.40 (1.06 to 1.84)	0.016	0.95 (0.26 to 3.55)	0.941
Factorial assignment	1.10 (0.61 to 1.98)	0.764	1.54 (0.86 to 2.76)	0.148
Cardiovascular (yes vs no)	1.02 (0.83 to 1.24)	0.876	1.41 (0.78 to 2.57)	0.256
Oncology (yes vs no)	0.42 (0.28 to 0.63)	<0.001	0.33 (0.18 to 0.63)	0.001
Mental health (yes vs no)	3.35 (2.56 to 4.38)	<0.001	1.23 (0.66 to 2.3)	0.518
Start year (after FDAAA 2007 vs before)	1.08 (0.7 to 1.66)	0.733	1.46 (0.58 to 3.71)	0.422
End point Classification (safety/efficacy vs safety only)	1.07 (0.79 to 1.45)	0.661	1.68 (0.88 to 3.19)	0.117

*1950 (58.02%) of the included 3361 studied were blind in this regression model.

†3234 (96.22%) of the included 3361 studied were randomised in this regression model.

NIH, National Institutes of Health.

requires that the enrolment of phase IV trials focusing on drug safety should be more than 2000.²⁰ For phase IV trials evaluating both efficacy and safety, the sample size should be calculated based on the effect sizes of efficacy and safety, respectively, and the study size should be determined by the larger one.

Phase IV clinical trials can have various designs and single-arm, non-randomised or open-label studies are accepted. If randomisation and blinding are feasible in the studies with controls arm, they can reduce bias and make evidence more reliable. Among the phase IV clinical trials with control, trials sponsored by a university or college were more likely to use blinding as compared to the phase IV clinical trials sponsored by industry. The methodological differences in trials were also evident among therapeutic areas. Oncology trials were less likely to use randomisation and blinding, which was consistent with the results of previous studies.¹⁵ One possible reason is that some of the oncology trials are conducted to investigate individualised or personalised treatment and randomisation or blinding is not feasible. Owing to the limitation of information on ClinicalTrials.gov, it is difficult to check whether all the phase IV trials with control are appropriately designed. However, the researcher should adopt randomisation and blinding when they are feasible.

Compared to prior analyses assessing the overall quality of the clinical trials landscape,¹⁵ our results showed some interesting findings. First, the Asia and Pacific area played a more important role in phase IV trials. Of the phase IV trials, 30.5% including the Asia and Pacific area, were a significant improvement over prior analyses of all clinical trials (13.5%).¹⁵ Including diverse populations could provide more information and help clinicians to ensure or refine the safety of approved drugs. Second, it was noted that the percentage of terminated or withdrawn phase IV trials was relatively high (8.6%). Califf's *et al*¹⁵ research revealed that 3.3% of all interventional clinical trials registered from October 2007 through September 2010 were terminated or withdrawn. We further analysed the conditions, end points and locations of the terminated or withdrawn phase IV trials but did not find any special characteristics other than small size (median: 38.0; IQR: 12.0–116.5). Third, the largest proportion of phase IV trials was funded by industry. Industry could use phase IV trials to expand the label of an approved drug or look for a completely new indication, which might be a potential explanation for the numerous small phase IV trials. However, the identification and characterisation of the risks associated with the prescription and use of medications are also essential and should be based on appropriate designs and sufficiently large sample sizes.

There are some inevitable limitations in this study. First, some clinical trials were not registered in the ClinicalTrials.gov registry, and these studies were not included in our analysis. However, ClinicalTrials.gov still accounts for more than 80% of all clinical studies in the

WHO portal,¹⁵ so our analysis is broadly representative. Second, there were some missing data for certain data fields, which may introduce some bias into the results. Third, as described in the 'Methods' section, we used the end point classification field from the ClinicalTrials.gov registry to identify phase IV trials related to drug safety; however, we did not perform additional manual screening to specify the primary end point for trials evaluating both safety and efficacy.

CONCLUSION

We found that the phase IV trials enterprise related to drug safety in ClinicalTrials.gov were dominated by small trials with significant heterogeneity in quality. These findings raise questions about the capacity of the phase IV trials to supply sufficient amounts of high quality evidence for safe medication. Adequate sample size should be emphasised for phase IV trials with safety as the primary end point.

Acknowledgements The authors gratefully acknowledge the valuable advice on revision from the reviewers. The authors also thank Jian Lu, PhD, for his assistance in designing the study. The authors acknowledge Meijing Wu and American Journal Experts, LLC for their professional copyediting service.

Contributors XZ and YZ contributed equally in conceiving this project, facilitating protocol, analysing data and drafting this manuscript. XY led the development of performance-based incentives and revised the manuscript critically. TZ and XG gave their time and effort to modify the programmes. JH provided expertise for the overall design of the study, and revised and approved the manuscript.

Funding This study was sponsored by the National Nature Science Foundation of China (number 81502895, 81373105), a grant from the key discipline for construction of evidence-based public health in Shanghai (number 12GWZX0602) and the Fourth Round of Three-year Action Plan on Public Health Discipline and Talent Programme: Evidence-based Public Health and Health Economics (number 15GWZK0901).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The analysed data set was upgraded on the Datadryad.org website. The title of the data set used in this revision is "phase IV clinical studies received by ClinicalTrials.gov between 2004 and 2014". URL: <http://datadryad.org/review?doi=doi:10.5061/dryad.3t6sc>.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. Harpaz R, Dumouchel W, Shah NH, *et al*. Novel data-mining methodologies for adverse drug event discovery and analysis. *Clin Pharmacol Ther* 2012;91:1010–21.
2. Bakke OM, Manocchia M, Deabajo F, *et al*. Drug safety discontinuations in the United Kingdom, the United States, and Spain from 1974 through 1993—a regulatory perspective. *Clin Pharmacol Ther* 1995;58:108–17.
3. Englev E, Petersen K. ICH-GCP guideline: quality assurance of clinical trials. Status and perspectives. *Ugeskr Laeger* 2003;165:1659–62.

4. Dixon JR. The international conference on harmonization good clinical practice guideline. *Qual Assur* 1999;6:65–74.
5. Gale EAM. Post-marketing studies of new insulins: sales or science? *BMJ* 2012;344:e3974.
6. Lasser KE, Allen PD, Woolhandler SJ, *et al.* Timing of new black box warnings and withdrawals for prescription medications. *JAMA* 2002;287:2215–20.
7. Food and Drug Administration. Food and Drug Administration Amendments Act (FDAAA) of 2007. <http://www.fda.gov/downloads/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticAct/FDCAAct/SignificantAmendmentstotheFDCAAct/FoodandDrugAdministrationAmendmentsActof2007/FDAAAImplementationChart/UCM213016.pdf> (accessed 4 Aug 2015).
8. Glasser SP, Salas M, Delzell E. Importance and challenges of studying marketed drugs: what is a phase IV study? Common clinical research designs, registries, and self-reporting systems. *J Clin Pharmacol* 2007;47:1074–86.
9. Schmidt LG, Grohmann R, Helmchen H, *et al.* Adverse drug-reactions—an epidemiological study at psychiatric hospitals. *Acta Psychiatr Scand* 1984;70:77–89.
10. Gillen JE, Tse T, Ide NC, *et al.* Design, implementation and management of a web-based data entry system for ClinicalTrials.gov. *Stud Health Technol Inform* 2004;107(Pt 2):1466–70.
11. Laine C, Horton R, DeAngelis CD, *et al.* Clinical trial registration: looking back and moving ahead. *Ann Intern Med* 2007;147:275–7.
12. Zarin DA, Tse T. Moving towards transparency of clinical trials. *Science* 2008;319:1340–2.
13. Cihoric N, Tsikkinis A, van Rhoon G, *et al.* Hyperthermia-related clinical trials on cancer treatment within the ClinicalTrials.gov registry. *Int J Hyperthermia* 2015;31:609–14.
14. Shields KE, Lyster AD. Exclusion of pregnant women from industry-sponsored clinical trials. *Obstet Gynecol* 2013;122:1077–81.
15. Califf RM, Zarin DA, Kramer JM, *et al.* Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007–2010. *JAMA* 2012;307:1838–47.
16. Tasneem A, Aberle L, Ananth H, *et al.* The database for Aggregate Analysis of ClinicalTrials.gov (AACT) and subsequent regrouping by clinical specialty. *PLoS ONE* 2012;7.
17. Zarin DA, Tse T, Williams RJ, *et al.* The clinicaltrials.gov results database—update and key issues. *N Engl J Med* 2011;364:852–60.
18. Hill KD, Chiswell K, Califf RM, *et al.* Characteristics of pediatric cardiovascular clinical trials registered on ClinicalTrials.gov. *Am Heart J* 2014;167:921–9.e2.
19. Suvama V. Phase IV of drug development. *Perspect Clin Res* 2010;1:57–60.
20. China Food and Drug Administration (CFDA). *Drug registration management measures*. 2007. <http://www.sfda.gov.cn/WS01/CL0053/24529.html> (accessed 4 Aug 2015).