BMJ Open

Characteristics of phase IV clinical trials for post-market drug safety surveillance

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010643
Article Type:	Research
Date Submitted by the Author:	23-Nov-2015
Complete List of Authors:	Zhang, Xinji; Second Mil Med Univ Zhang, Yuan; Second Mil Med Univ Ye, Xiaofei; Second Mil Med Univ Zhang, Tianyi; Second Mil Med Univ Guo, Xiaojing; Second Mil Med Univ He, Jia; Second Mil Med Univ,
Primary Subject Heading:	Medical management
Secondary Subject Heading:	Health informatics, Pharmacology and therapeutics
Keywords:	Clinical trials < THERAPEUTICS, Adverse events < THERAPEUTICS, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

Characteristics of phase IV clinical trials for post-market drug safety surveillance

Xinji Zhang [1*], Yuan Zhang [1*], Xiaofei Ye[1], Tianyi Zhang[1], Xiaojing Guo[1], Jia He [1△]

Author Affiliations and Address:

[1] Department of Health Statistics, Second Military Medical University, Shanghai, 200433, China;

[*] These authors contributed equally and are co-first authors of this article.

[△]Corresponding author:

Prof Jia He, Department of Health Statistics, Second Military Medical University, No. 800 Xiangyin Road, Shanghai 200433, China

Tel: +86-021-81871441, Fax: +86-021-81871441, E-mail: hejia63@yeah.net

Key words or phrases for search: clinical trials; health and safety; adverse events; pharmacology

Number of words: 2,636

Abstract

Objective: Phase IV trial is an important contributor to the postmarketing surveillance of drug safety. We aimed to determine fundamental characteristics of phase IV clinical trials which evaluate drug safety using ClinicalTrials.gov data

Methods: A data set of 19,359 phase IV clinical studies registered with ClinicalTrials.gov (between January 1, 2004 and November 31, 2014) was downloaded. Characteristics of the phase IV trials focus on safety only were compared with those evaluating both safety and efficacy. We also compared the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular, mental health, and oncology). Multivariable logistic regression was used to evaluate factors associated with usage of blinding and randomization.

Results: 5847 Phase IV trial were identified, including 497 focused on drug safety only and 5350 evaluating both safety and efficacy. Most of the phase IV trials evaluating drug safety (77.6%) had the enrollment less than 300 and 98.4% less than 3000. 8.6% of trials were terminated or withdrawn. Factors associated with use of blinding and randomization included endpoint classification, clinical specialty and sponsor type.

Conclusions: Phase IV trials enterprise evaluating drug safety in ClinicalTrials.gov is dominated by small trials which might not have sufficient power to detect less common events. Adequate sample size should be emphasized for phase IV trials with safety surveillance as main task.

Article summary

Strengths and limitations of this study

Strengths

1. We provided a comprehensive descriptive assessment of the current portfolio of phase IV clinical

trials which evaluated drug safety.

2. We employed logistic regression models to find out the factors which influence the usage of blinding and randomization of phase IV clinical trials which evaluated drug safety.

3. We followed the strict analysis process which was widely used in analyzing the data from ClinicalTrials.gov and made the results convincing.

Limitations

1. Some clinical trials are not registered in ClinicalTrials.gov.

2. There was some unavoidable missing data for certain data fields which may induce some bias to the results.

INTRODUCTION

Drug adverse reaction is a major global health concern accounting for more than 2 million injuries, hospitalizations, and deaths each year in the US alone,[1] and associated billions of US dollars in costs every year in all developed countries.[2] Although rigorous premarketing studies are required for all new drugs,[3 4] a drug's safety profile at the time of regulatory approval is often incomplete due to the 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 post-marketing, and 4% of drugs were ultimately withdrawn for safety reasons[6]. In 2007, the Food and Drug Administration was authorized by the Food and Drug Administration Amendment Act (FDAAA)[7] to require postmarketing clinical trials to address safety concern of a given drug. Compare to premarket phase I-III trials, phase IV studies can evaluate drugs' safety in a related real-world setting, which make them become important and increasing contributors to the postmarketing active surveillance of drug safety.[5 8 9] In order to provide reliable evidence to ensure or further refine the safety of approved drugs, phase IV trials should have appropriate designs and sufficiently large sample sizes. Until recently, however, we have lacked tools for comprehensively assessing phase IV clinical trials involving drug safety.

ClinicalTrials.gov is a public trial registry which was established by the National Library of Medicine on behalf of the National Institutes of Health (NIH) and first lunched in February 2000.[10] Since 2005, the International Committee of Medical Journal Editors (ICMJE) has implemented the policy requiring registration of clinical trials as a prerequisite for publication.[11] In addition, in 2007, sponsors or their designees were obliged by FDAAA to register trials and report key data elements and basic trial results at ClinicalTrials.gov.[12] The data from ClinicalTrials.gov had been studied from different angles and provided lots information.[13-15]

In this article, we examine fundamental characteristics of phase IV clinical trials which referred to drug safety by evaluating the ClinicalTrials.gov data. We focus on data elements that are desirable for generating reliable evidence from clinical trials, including factors associated with use of randomization 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 (between January 1, 2004 and November 31, 2014) was downloaded from the website. The data set was locked, and a relational database was subsequently designed to facilitate analysis.[15 16]

Study selection

Our analysis was restricted to interventional studies. To identify interventional studies, we used the "study types" field from the ClinicalTrials.gov registry.[15] Besides, we restricted to intervention type to "drug" only, based on the "Interventions" field from the ClinicalTrials.gov, which included many other choices, including: Biological, Device, Procedure and so on. Finally, we singled out with the endpoint classification of "safety study" or "safety/efficacy study", which means these trials were referred to drug safety.

The included trials were regrouped to three derivative databases according to the 3 clinical specialties—mental health, oncology and cardiovascular. For this regrouping, we used the information of "Conditions" fields from the ClinicalTrials.gov and the information of the classification of studies based on category by ClinicalTrials.gov. We used the studies groups classified by the ClinicalTrials.gov as the matching database and regrouped the original database by

matching the NCT number of each study between the original database and the matching database.

Data collection

Trial's data is reported by trial sponsors or investigators as ClinicalTrials.gov required.[17] Each record contains a set of data elements describing the study's conditions, enrollment, study design, eligibility criteria, location, sponsor, and other protocol information. When the data field was incomplete, a web search (ClinicalTrials.gov) was conducted to identify the missing information of the trial. If the information was not available on the website either, this field will be identified as NA (not applicable) or missing.

Derived variables

The methods of defining derived variables have been described previously [15 18] and are briefly summarized below. Funding source was derived using the information of the lead sponsor and collaborators and all trials were divided into four groups: industry, NIH, other, US Federal (excluding NIH). Funding source was defined as NIH if the lead sponsor or any collaborators were from NIH, and the lead sponsor was not from industry. Funding source was defined as industry if the lead sponsor was from industry or if any collaborators were from industry and there was no NIH involvement. Funding source was defined as from US Federal sources, if the sponsor were from US Federal only. For the remaining studies, funding was defined as from other sources. The primary purpose of the trial was divided into nine groups based on the information in "Primary purpose" field: Treatment, Prevention, Diagnostic, Supportive Care, Screening, Health Services Research, Basic Science grouped, Educational/Counseling/Training and missing. The information about appointment of a data monitoring committee (DMC) became available in April 2007, and is not a required field[18], so in our study, we did not take this into consideration. The other variable' classification

was based on the information of related fields from the ClinicalTrials.gov.

Analysis Methods

Trials' characteristics were evaluated overall and by different endpoint classifications and clinical specialties. Assessments included overall status, enrollment, intervention model, funding source, et al. The percentage of trials registered before and after enrollment of the first participant was also determined by comparing the date of registration with the trial's start date. According to the principles of Binomial and Poisson distribution, if investigators plan to observe at least 1 case of adverse events with occurrence probability less than 1%, 0.5% and 1‰, the enrollment should be larger than 300, 600 and 3000,respectively (Table1).[19] Hence, we divided the included trials into five kinds: trials with sample size less than 300, between 300 and 599, between 600 and 2999, no less than 3000 and missing. Frequencies and percentages are provided for categorical characteristics; medians and interquartile ranges (IQRs) are provided for continuous characteristics.

Table 1 the numbers of patients needed to be observed

Expected incidence adverse reaction	Numbers of patients to observed to of detect at least 1 event
1 in 100	300
1 in 200	600
1 in 1000	3000s

Logistic regression analysis was performed to calculate adjusted odds ratios (ORs) with wald 95% confidence intervals for factors associated with trials that report usage of randomization and blinding. 8 prespecified variables were included in the model: funding source; primary purposes; number of participants; trial specialty (yes/no); trial's start year and endpoint classification (safety/efficacy study or safety study).

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

RESULT

From January 1, 2004, to December 13, 2014, 18,642 phase IV trials were registered at ClinicalTrials.gov. Of these, 5,847 were interventional studies which related to drug safety. The number of trials evaluating safety only was 497, which was significantly less than the number of trials evaluating both safety and efficacy (n=5350). 384 trials (6.6%) focused on mental health diseases, 221 trials (3.8%) focused on oncology, and 540 trials (9.2%) focused on cardiovascular diseases (Figure 1).

Basic characteristics of all inclusive 5847 trials registered with ClinicalTrials.gov are shown in Table 2. The median number of participants per trial was 105.0 (IQR, 48.0-260.0). 73.9% of these phase IV trials used randomization and 46.2% used blinding (including double-blind and single-blind). We also noted that 8.6% (n=501) of these phase IV trials were "terminated" or "withdrawn", of which most were small studies (median enrollment: 38.0; IQR, 12.0-116.5). The most common research sites in these phase IV trials were from North American, Asia and Pacific, Europe, which accounted for 37.9%, 30.5% and 30.5%, respectively.

Table 2 Characteristics for Inclusive Trials in different Endpoint Classifications

	No./Total No. (%)		
	All,2004-2014 N=5847	Endpoint classification Safety N=497	Endpoint classification Safety/ Efficacy N=5350
Overall status			
Not yet recruiting	265/5847(4.5)	17/497(3.4)	248/5350(4.6)
Recruiting	1220/5847(20.9)	91/497(18.3)	1129/5350(21.1)
Completed	3424/5847(58.6)	291/497(58.6)	3133/5350(58.6)
Suspended	23/5847(0.4)	5/497(1.0)	18/5350(0.3)
Terminated	389/5847(6.7)	42/497(8.5)	347/5350(6.5)
Withdrawn	112/5847(1.9)	9/497(1.8)	103/5350(1.9)
Active, not recruiting	355/5847(6.1)	41/497(8.2)	314/5350(5.9)
Enrolling by invitation	59/5847(1.0)	1/497(0.2)	58/5350(1.1)
Enrollment, median (IQR),	105.0(48.0,260.0)	104.0(42.0,317.0)	105.0(48.5,256.0)
1-299	4421/5698(77.6)	355/485(73.2)	4066/5153(78.9)

300-599	743/5698(13.0)	66/485(13.6)	577/5153(13.1)
600-2999	454/5698(8.0)	48/485(9.9)	406/5153(7.9)
⩾3000	90/5698(1.6)	16/485(3.3)	74/5213(1.4)
missing	139/5847(2.4)	12/497(2.8)	127/5350(2.6)
Intervention Model			
Crossover Assignment	343/5838(5.9)	55/497(11.1)	288/5341(5.4)
Single Group Assignment	1506/5838(25.8)	187/497(37.6)	1319/5341(24.7)
Parallel Assignment	3914/5838(67)	253/497(50.9)	3661/5341(68.5)
Factorial Assignment	75/5838(1.3)	2/497(0.4)	73/5341(1.4)
NA	9/5847(0.2)	0/497(0.0)	9/5350(0.2)
Allocation			
Randomized	4323/5847(73.9)	302/497(60.8)	4021/5350(75.2)
non-Randomized	853/5847(14.6)	102/497(20.5)	751/5350(14.0)
NA	671/5847(11.5)	93/497(18.7)	578/5350(10.8)
Masking			
Double-Blind	2245/5823(38.6)	151/497(30.4)	2094/5326(39.3)
Single blind	441/5823(7.6)	49/497(9.9)	392/5326(7.4)
Open label	3152/5823(54.1)	297/497(59.8)	2855/5326(53.6)
Missing	24/5847(0.4)	0/497(0.0)	24/5350(0.4)
Sex,%			
Female only	454/5847(7.8)	32/497(6.4)	422/5350(7.9)
Male only	193/5847(3.3)	33/497(6.6)	160/5350(3.0)
Both	5200/5847(88.9)	432/497(86.9)	4768/5350(89.1)
Includes children (<18y)	959/5847(16.4)	95/497(19.1)	864/5350(16.1)
Excludes elderly (>65 y)	1690/5847(28.9)	140/497(28.2)	1550/5350(29.0)
Primary Purpose			
Basic Science	24/5729(0.4)	8/454(1.8)	16/5275(0.3)
Diagnostic	89/5729(1.6)	25/454(5.5)	64/5275(1.2)
Educational/Counseling/Training	3/5729(0.1)	1/454(0.2)	2/5275(0.0)
Health Services Research	23/5729(0.4)	2/454(0.4)	21/5275(0.4)
Prevention	512/5729(8.9)	60/454(13.2)	452/5275(8.6)
Screening	8/5729(0.1)	3/454(0.7)	5/5275(0.1)
Supportive Care	102/5729(1.8)	10/454(2.2)	92/5275(1.7)
Treatment	4968/5729(86.7)	345/454(76.0)	4623/5275(87.6)
NA	118/5847(2.0)	43/497(8.7)	75/5350(1.4)
Leader sponsor			
Industry	3160/5847(54.0)	313/497(63.0)	2847/5350(53.2)
NIH	124/5847(2.1)	9/497(1.8)	115/5350(2.1)
U.S.Federal	20/5847(0.3)	0/497(0.0)	20/5350(0.4)
Other	2543/5847(43.5)	174/497(35.0)	2368/5350(44.3)
Region [a]			
Africa	210/5245(4.0)	17/455(3.7)	193/4790(4.0)
Asia and Pacific	1602/5245(30.5)	118/455(25.9)	1484/4790(31)
Central and South America	387/5245(7.4)	41/455(9.0)	346/4790(7.2)

Europe	1598/5245(30.5)	158/455(34.7)	1440/4790(30.1)
Middle East	295/5245(5.6)	26/455(5.7)	269/4790(5.6)
North America	1988/5245(37.9)	213/455(46.8)	1775/4790(37.1)
Missing	602/5847(10.3)	42/497(8.5)	560/5350(10.5)
Study registration [b]			
Before first participant enrolled	1944/5716(34.0)	205/491(41.8)	1739/5225(33.3)
After first participant enrolled	3772/5716(66.0)	286/491(58.2)	3486/5225(66.7)

Note:

a: Percentages may not sum to 100% as categories are not mutually exclusive.

b: Out of 5847 trials, 131 (2.2%) were missing registration information;6/497 (1.2%) from safety trials., 125/5350 (2.3%) from safety/efficacy trials.

73.2% of the phase IV trials focused on drug safety only had the enrollment less than 300,and only 3.3% (n=16) of the phase IV trials focused on drug safety only enrolled more than 3000 patients. The median number of participants per trial was 104.0 (IQR, 42.0-317.0). The average sample size of the phase IV trials assessed both safety and efficacy was similar, with median enrollment of 105 (IQR, 48.5-256.0). Compared with those evaluating both safety and efficacy, phase IV trials focused on drug safety only showed larger proportion of studies using single group assignment (37.6% *vs* 24.7%) and small proportion using randomization (60.8% *vs* 75.2%). However, the difference of proportion of studies using blinding was relatively small between trials focused on safety only and those assessed safety/efficacy (40.3% *vs* 46.7%). With respect to study registration, the performance of trials focused on safety only was outstanding, with 41.8% registered before enrollment as compared with 33.3% in trials evaluating both safety and efficacy.

Table 3 shows the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular, oncology and mental health). Of these 3 categories, cardiovascular trials were most numerous (n=806, 13.8%). Besides, cardiovascular trials had larger enrollment (median, 175; IQR, 64.0-460.0) than oncology trials (median, 100.0; IQR, 50.0-210.5) and mental health trials (median, 88.0; IQR, 40.0-223.0). Among trials oriented toward prevention, cardiovascular trials contained the largest group: 13.7% *vs* 6.5% for oncology and 3.2% for mental health. Mental health related trials

took more consideration on the treatment than cardiovascular trials and oncology trials (94.5% vs 81.9% and 86.3%, respectively).The usage of randomization were less common in oncology trials than the cardiovascular trials and mental health trials(61.0% *vs* 88.7% for cardiovascular and 80.7% for mental health). The difference in the usage of blinding is similar (20.5% for oncology trials *vs* 48.4% for cardiovascular trials and 57.7% for mental health trials). Among trials including female only, oncology trials comprised the largest group: 12.8% vs 1.6% for cardiovascular trials and 2.5% for mental health trials. It's noteworthy that nearly two thirds of mental health trials (64.3%) excluded elder. Geographical differences were also apparent. Mental health related trials showed the largest proportion of studies with at least one North American research site (52.4%). Whereas, oncology trials showed the largest proportion of studies with at least one Asia and Pacific research site (41.3%). NIH paid more attention on the mental health trials: 9.1% vs 1.2% for cardiovascular trials and 0.6% for oncology trials.

Table 3 Characteristics for Inclusive Trials in different clinical specialties

	No./Total No. (%)			
	All,2004-2014 N=5847	Cardiovascular N=806	Oncology N=312	Mental Health N=672
Overall status				
Not yet recruiting	265/5847(4.5)	51/801(6.4)	13/306(4.2)	23/668(3.4)
Recruiting	1220/5847(20.9)	183/801(22.8)	105/306(34.3)	126/668(18.9)
Completed	3424/5847(58.6)	428/801(53.4)	130/306(42.5)	437/668(65.4)
Suspended	23/5847(0.4)	5/801(0.6)	2/306(0.7)	0/668(0.0)
Terminated	389/5847(6.7)	54/801(6.7)	16/306(5.2)	34/668(5.1)
Withdrawn	112/5847(1.9)	18/801(2.2)	7/306(2.3)	13/668(1.9)
Active, not recruiting	355/5847(6.1)	62/801(7.7)	33/306(10.8)	35/668(5.2)
Enrolling by invitation	59/5847(1.0)	5/806(0.6)	6/312(2.0)	4/672(0.6)
Enrollment, median (IQR)	105.0(48.0,260.0)	175.0(64.0,460.0)	100.0(50.0,210.5)	88.0(40.0,223.0)
1-299	4421/5698()	510/780(65.4)	247/304(81.3)	536/661(81.1)
300-599	743/5698()	108/780(13.8)	33/304(10.9)	89/661(13.6)
600-2999	454/5698()	125/780(16.0)	21/304(6.9)	31/661(4.6)
⩾3000	90/5698()	37/780(4.7)	3/304(1.0)	5/661(0.8)
missing	139/5847(2.4)	26/806(3.2)	8/312(2.6)	11/672(1.6)

Intervention Model				
Crossover Assignment	343/5838(5.9)	36/803(4.5)	5/312(1.6)	43/671(6.4)
Single Group Assignment	1506/5838(25.8)	140/803(17.4)	161/312(51.6)	209/671(31.1)
Parallel Assignment	3914/5838(67)	609/803(75.8)	142/312(45.5)	410/671(61.1)
Factorial Assignment	75/5838(1.3)	18/803(2.2)	4/312(1.3)	9/671(1.3)
NA	9/5847(0.2)	3/806(0.4)	0/312(0.0)	1/672(0.1)
Allocation				
Randomized	4323/5847(73.9)	669/754(88.7)	147/241(61.0)	461/571(80.7)
Non-Randomized	853/5847(14.6)	85/754(11.3)	94/241(39)	110/571(19.3)
NA	671/5847(11.5)	52/806(6.5)	71/312(22.8)	101/672(15)
Masking				
Double-Blind	2245/5823(38.6)	316/805(39.3)	47/312(15.1)	348/670(51.9)
Single blind	441/5823(7.6)	73/805(9.1)	17/312(5.4)	39/670(5.8)
Open label	3152/5823(54.1)	416/805(51.7)	248/312(79.5)	283/670(42.2)
Missing	24/5847(0.4)	1/806(0.1)	0/312(0)	2/672(0.3)
Sex,%				
Female only	454/5847(7.8)	13/806(1.6)	40/312(12.8)	17/672(2.5)
male only	193/5847(3.3)	12/806(1.5)	23/312(7.4)	33/672(4.9)
Both	5200/5847(88.9)	781/806(96.9)	249/312(79.8)	622/672(92.6)
Includes children (<18 y)	959/5847(16.4)	50/806(6.2)	47/312(15.1)	134/672(19.9)
Excludes elder (>65 y)	1690/5847(28.9)	65/806(8.1)	30/312(9.6)	432/672(64.3)
Primary Purpose				
Basic Science	24/5729(0.4)	3/794(0.4)	0/306(0)	3/659(0.5)
Diagnostic	89/5729(1.6)	21/794(2.6)	10/306(3.3)	7/659(1.1)
Educational/Counseling/Training	3/5729(0.1)	0/794(0)	0/306(0)	1/659(0.2)
Health Services Research	23/5729(0.4)	1/794(0.1)	1/306(0.3)	2/659(0.3)
Prevention	512/5729(8.9)	109/794(13.7)	20/306(6.5)	21/659(3.2)
Screening	8/5729(0.1)	2/794(0.3)	1/306(0.3)	0/659(0)
Supportive Care	102/5729(1.8)	8/794(1)	10/306(3.3)	2/659(0.3)
Treatment	4968/5729(86.7)	650/794(81.9)	264/306(86.3)	623/659(94.5)
NA	118/5847(2)	12/806(1.5)	6/312(2)	13/672(2)
Leader sponsor				
Industry	3160/5847(54.0)	385/806(47.8)	172/312(55.1)	392/672(58.3)
NIH	124/5847(2.1)	10/806(1.2)	2/312(0.6)	61/672(9.1)
U.S.Federal	20/5847(0.3)	1/806(0.1)	1/312(0.3)	3/672(0.4)
Other	2543/5847(43.5)	410/806(50.9)	137/312(43.9)	216/672(32.1)
Region				
Africa	210/5245(4)	25/806(3.1)	16/312(5.1)	13/672(1.9)
Asia and Pacific	1602/5245(30.5)	277/806(34.4)	129/312(41.3)	155/672(23.1)
Central and South America	387/5245(7.4)	39/806(4.8)	16/312(5.1)	46/672(6.8)
Europe	1598/5245(30.5)	237/806(29.4)	100/312(32.1)	93/672(13.8)
Middle East	295/5245(5.6)	40/806(5)	20/312(6.4)	26/672(3.9)
North America	1988/5245(37.9)	238/806(29.5)	72/312(23.1)	352/672(52.4)
Missing	602/5847(10.3)	66/806(8.2)	32/312(10.3)	63/672(9.4)

Study registration				
Before first participant enrolled	1944/5716(34.01)	266/806(33)	109/312(34.9)	207/672(30.8)
After first participant enrolled	3772/5716(65.99)	513/806(63.6)	195/312(62.5)	455/672(67.7)

Note:

a: Percentages may not sum to 100% as categories are not mutually exclusive.

b: Out of 5847 trials, 131 (2.2%) were missing registration information;27/806 (3.3%) from cardiovascular trials., 8/312 (2.6%) from Oncology trials,10/672(1.5%) from Mental Health

Table 4 shows the result of regression analysis. This analysis compared the trials 'characteristics which related to the usage of blinding and randomization. When compared with the trials focused on safety only, the trials focused on both safety and efficacy were more likely to use blinding (adjusted OR, 1.35; 95% CI, 1.10-1.66) and randomization (adjusted OR, 1.86; 95% CI, 1.51-2.30). Different clinical specialties can also influence the usage of blinding and randomization. Oncology trials were less likely to use both blinding (adjusted OR, 0.33; 95% CI, 0.24-0.43) and randomization (adjusted OR, 0.31; 95% CI, 0.24-0.40). Mental health trials were more likely to use blinding (adjusted OR, 1.69; 95% CI, 1.42-2.01) but less likely to use randomization (adjusted OR, 0.70; 95% CI, 0.59-0.85). In addition, compared with treatment trials, the prevention trials were more likely to use blinding (adjusted OR, 1.41; 95% CI, 1.16-1.70) and randomization (adjusted OR, 1.44; 95% CI, 1.11-1.85); the diagnostic trials were less likely to use randomization (adjusted OR, 0.60; 95% CI, 0.38-0.95). Compared with the trials which industry was the leader sponsor, the trials funded by other sources were more likely to use blinding (adjusted OR, 1.50; 95% CI, 1.34-1.68).

Table 4 Regression Analyses of Inclusive Trials and the Reported Usage of Blinding and Randomization

Variable	Blinding		Randomization	
	Adjusted OR (95% CI)	*p* Value	Adjusted OR (95% CI)	*p* Value
Leader sponsor (vs industry)				
NIH	1.84(0.77,4.42)	0.172	0.58(0.24,1.41)	0.226
Other	1.50(1.34,1.68)	<0.001	0.87(0.36,2.12)	0.757
US federal	1.87(0.94,3.73)	0.073	1.11(0.36,3.45)	0.858
Study size (per 1000 additional participants)	1.02(0.99,1.04)	0.185	1.02(0.98,1.05)	0.362

Cardiovascular (yes vs no)	1.09(0.93,1.28)	0.292	1.57(1.28,1.94)	<0.001
Oncology (yes vs no)	0.33(0.24,0.43)	<0.001	0.31(0.24,0.40)	<0.001
Mental health (yes vs no)	1.69(1.42,2.01)	<0.001	0.70(0.59,0.85)	<0.001
Start year (increment of 1 y)	1.02(1.01,1.04)	0.008	1.01(0.99,1.03)	0.202
Primary category (vs treatment)				
Diagnostic	1.19(0.76,1.86)	0.441	0.60(0.38,0.95)	0.029
Prevention	1.41(1.16,1.70)	<0.001	1.44(1.11,1.85)	0.005
Other	1.30(0.94,1.80)	0.110	0.97(0.67,1.41)	0.873
Endpoint Classification (safety/efficacy vs safety only)	1.35(1.10,1.66)	0.004	1.86(1.51,2.30)	<0.001

Abbreviations: NIH, National Institutes of Health; OR, odds ratio.

DISSCUSSION

This study provided a descriptive assessment of the current portfolio of phase IV clinical trials which evaluated drug safety. Characteristics of phase IV trials with different endpoint classifications and clinical specialties were compared. And we also analyzed the factors associated with trial quality. Thus this analysis presents a unique opportunity to evaluate of the landscape of phase IV trials referred to drug safety, and identify areas of relative strength or weakness.

Small sample size is the most concerning issue of phase IV trials involving the safety surveillance of an approved drug. Small phase IV trials may be used in evaluation the effectiveness of a given drug in special patient subgroup, or special situations.[5] However, our study only included the phase IV trials with "safety" as an endpoint and most of the trials (77.6%) had an enrollment less than 300. This means that those small trials (n<300) may not have sufficient power to detect adverse events (AEs) with occurrence probability less than 1%.[19] Even in the phase IV trials with safety as primary endpoint, the average sample size was only 104. These findings raise fundamental questions about the ability of the small phase IV trials to provide informative results of the drug safety, especially for less common adverse events. Administration agency may emphasize adequate sample size for phase IV trials with safety assessment as a main task. For example, to observe AE with occurrence probability 1.5‰, the China food and drug administration(CFDA)

required that the enrollment of phase IV trials focus on drug safety should be more than 2000.[20]

Compared with the phase IV clinical trials evaluating both safety and efficacy, those focusing on safety only were less likely to use randomization and blinding. More common single-group design may explain some of the differences in approach with the trials evaluating safety only. The methodological differences of trials were also evident among therapeutic areas. One positive finding was that a relatively large percentage of phase IV cardiovascular trials were conducted with randomization and blinding, which could make their results more subjective and reliable. However, oncology trials were less likely to use randomization and blinding, even after adjustment for other factors.

Compared to prior analyses assessing overall quality of the clinical trials landscape,[15] our results showed some interesting findings. Firstly, Asia and Pacific area plays a more important role in phase IV trials. The 30.5% of phase IV trials including Asia and Pacific area is a significant improvement over prior analyses of all clinical trials (13.5%).[15] Inclusion of various 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%). Robert's research[15] demonstrated that 3.3% of all interventional clinical trials registered from October 2007 through September 2010 were terminated or withdrawn. We further analyzed the condition, endpoint and locations of the terminated or withdrawn phase IV trials, but did not find any special characteristic except for 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 trial to expand the label of an approved drug or look for a completely new indication, which might be a potential explanation of the numerous small phase IV trials. However, the identification and

characterization of the risks associated with the prescribing and use of medications were also essential, which should be based on appropriate designs and sufficiently large sample size

There are some inevitable limitations in this study. First, some clinical trials are not registered in ClinicalTrials.gov. 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 is some missing data for certain data fields, which may induce some bias to the results. Third, as described in the "Methods" section, we use the endpoint classification field from the ClinicalTrials.gov registry to identify phase IV trials referred to drug safety; however, we did not perform additional manual screening to specify the primary endpoint for those evaluating both safety and efficacy.

CONCLUSION

We found that the phase IV trials enterprise referred to drug safety in ClinicalTrials.gov is dominated by small trials with significant heterogeneity in quality. These findings raise questions about the capacity of the phase IV trials enterprise to supply sufficient amounts of high quality evidence for safe medication. Adequate sample size should be emphasized for phase IV trials with safety as primary endpoint. And the integration of evidence from clinical trials and other sources such as researches based on spontaneous self-reports of adverse reactions and electronic medical records would further advance the science of evaluating drug safety in real-world settings.

Figure Legends:

Figure 1 Consortium diagram: Inclusion criteria and methods for study selection

FUNDINGS STATEMENT

This study was sponsored by the National Nature Science Foundation of China (NO. 81502895, 81373105), a grant from the key discipline for construction of evidence-based public health in Shanghai (NO. 12GWZX0602).

COMPETING INTERESTS

The authors declare that they have no competing interests.

CONTRIBUTORSHIP STATEMENT

Xinji Zhang and Yuan Zhang contributed equally in conceiving this project, facilitating protocol, analyzing data and drafting this manuscript. Yexiao Fei leaded the development of performance-based incentives and revised the manuscript critically. Tianyi Zhang and XiaoJing Guo gave their time and effort to modify the programs. Jia He provided expertise for the overall design of the study, and revised and approved the manuscript.

ACKONWLEDGEMENTS

We gratefully acknowledge for his assistance in designing the study. We also acknowledge others who gave their time and effort to this study.

DATA SHARING STATEMENT

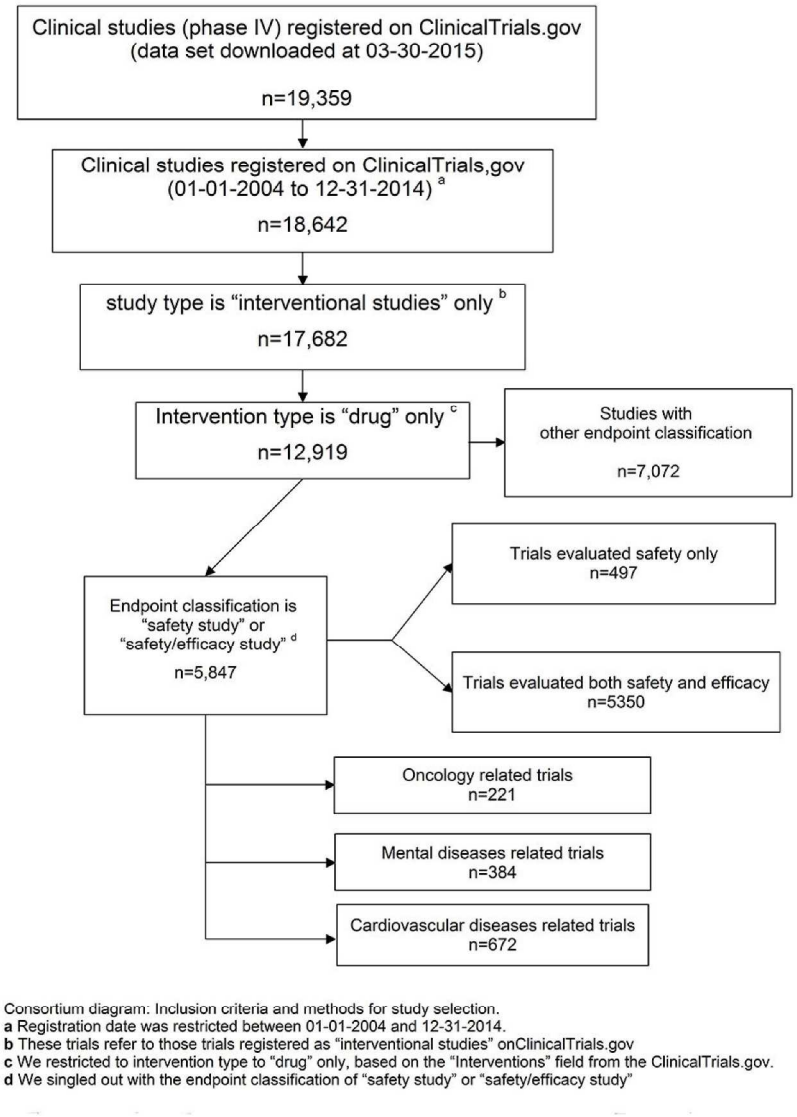
No additional unpublished data are available

REFERENCE

1. Harpaz R, ., Dumouchel W, ., Shah NH, et al. Novel data-mining methodologies for adverse drug event discovery and analysis. Clinical Pharmacology & Therapeutics 2012;**91**(6):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. Clinical Pharmacology & Therapeutics 1995;**58**(1):108-17 doi: 10.1016/0009-9236(95)90078-0[published Online First: Epub Date]|.
3. Englev E, Petersen K. [ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives]. Ugeskrift for laeger 2003;**165**(16):1659-62
4. Dixon JR. The international conference on harmonization good clinical practice guideline. Quality Assurance: Good Practice, Regulation, and Law 1999;**6**(2):65-74

5. Gale EAM. Post-marketing studies of new insulins: sales or science? British Medical Journal 2012;**344** doi: 10.1136/bmj.e3974[published Online First: Epub Date]|.
6. Lasser KE, Allen PD, Woolhandler SJ, et al. Timing of new black box warnings and withdrawals for prescription medications. Jama-Journal of the American Medical Association 2002;**287**(17):2215-20 doi: 10.1001/jama.287.17.2215[published Online First: Epub Date]|.
7. Food, Administration D. Food and Drug Administration Amendments Act (FDAAA) of 2007, 2009.
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. The Journal of Clinical Pharmacology 2007;**47**(9):1074-86
9. Schmidt LG, Grohmann R, Helmchen H, et al. ADVERSE DRUG-REACTIONS - AN EPIDEMIOLOGICAL-STUDY AT PSYCHIATRIC-HOSPITALS. Acta Psychiatrica Scandinavica 1984;**70**(1):77-89 doi: 10.1111/j.1600-0447.1984.tb01185.x[published Online First: Epub Date]|.
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**(4):275-77
12. Zarin DA, Tse T. Moving towards transparency of clinical trials. Science (New York, NY) 2008;**319**(5868):1340
13. Cihoric N, Tsikkinis A, van Rhoon G, et al. Hyperthermia-related clinical trials on cancer treatment within the ClinicalTrials.gov registry. International Journal of Hyperthermia 2015;**31**(6):609-14 doi: 10.3109/02656736.2015.1040471[published Online First: Epub Date]|.
14. Shields KE, Lyerly AD. Exclusion of pregnant women from industry-sponsored clinical trials. Obstetrics and gynecology 2013;**122**(5):1077-81 doi: 10.1097/AOG.0b013e3182a9ca67[published Online First: Epub Date]|.
15. Califf RM, Zarin DA, Kramer JM, et al. Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007-2010. Jama-Journal of the American Medical Association 2012;**307**(17):1838-47 doi: 10.1001/jama.2012.3424[published Online First: Epub Date]|.
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**(3) doi: 10.1371/journal.pone.0033677[published Online First: Epub Date]|.
17. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov Results Database - Update and Key Issues. New England Journal of Medicine 2011;**364**(9):852-60 doi: 10.1056/NEJMsa1012065[published Online First: Epub Date]|.
18. Hill KD, Chiswell K, Califf RM, et al. Characteristics of pediatric cardiovascular clinical trials registered on ClinicalTrials.gov. American Heart Journal 2014;**167**(6):921-U195 doi: 10.1016/j.ahj.2014.02.002[published Online First: Epub Date]|.
19. Suvarna V. Phase IV of Drug Development. Perspectives in clinical research 2010;**1**(2):57-60
20. China Food and Drug Administratio: **Drug Registration Management Measures(2007)**.[2015.11.12] http://www.sfda.gov.cn/WS01/CL0053/24529.html.

Figure 1



Consortium diagram: Inclusion criteria and methods for study selection.
a Registration date was restricted between 01-01-2004 and 12-31-2014.
b These trials refer to those trials registered as "interventional studies" onClinicalTrials.gov
c We restricted to intervention type to "drug" only, based on the "Interventions" field from the ClinicalTrials.gov.
d We singled out with the endpoint classification of "safety study" or "safety/efficacy study"

Figure 1 Consortium diagram: Inclusion criteria and methods for study selection
Figure 1
494x700mm (300 x 300 DPI)

BMJ Open

An overview of phase IV clinical trials for post-market drug safety surveillance: analysis of ClinicalTrials.gov registry

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010643.R1
Article Type:	Research
Date Submitted by the Author:	06-May-2016
Complete List of Authors:	Zhang, Xinji; Second Mil Med Univ Zhang, Yuan; Second Mil Med Univ Ye, Xiaofei; Second Mil Med Univ Zhang, Tianyi; Second Mil Med Univ Guo, Xiaojing; Second Mil Med Univ He, Jia; Second Mil Med Univ,
Primary Subject Heading:	Medical management
Secondary Subject Heading:	Health informatics, Pharmacology and therapeutics
Keywords:	Clinical trials < THERAPEUTICS, Adverse events < THERAPEUTICS, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

An overview of phase IV clinical trials for post-market drug safety surveillance: analysis of ClinicalTrials.gov registry

Xinji Zhang [1*], Yuan Zhang [1*], Xiaofei Ye[1], Tianyi Zhang[1], Xiaojing Guo[1], Jia He [1△]

Author Affiliation and Address:

[1] Department of Health Statistics, Second Military Medical University, Shanghai, 200433, China

[*] These authors contributed equally and are co-first authors of this article.

[△]Corresponding author:

Prof Jia He, Department of Health Statistics, Second Military Medical University, No. 800 Xiangyin Road, Shanghai 200433, China

Tel: +86-021-81871441, Fax: +86-021-81871441, E-mail: hejia63@yeah.net

Key words or phrases for search: clinical trials; health and safety; adverse events

Number of words: 2,750

Abstract

Objective: Phase IV trials are often used to survey 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 evaluate drug safety using ClinicalTrials.gov data

Methods: A data set of 19,359 phase IV clinical studies registered with ClinicalTrials.gov was downloaded. The Characteristics of the phase IV trials focusing only on safety were compared with the ones evaluating both safety and efficacy. We also compared the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular diseases, mental health, and oncology). Multivariable logistic regression was used to evaluate factors associated with the use of blinding and randomization.

Results: A total of 4,772 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 enrollment less than 300 with 96.5% less than 3000. A total of 8.2% of trials were terminated or withdrawn. Factors associated with the useage of blinding and randomization included the intervention model , clinical specialty and lead sponsor .

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

Article summary

Strengths and limitations of this study

Strengths

1. We provided a comprehensive descriptive assessment of the current portfolio of phase IV clinical trials evaluating drug safety.

2. We employed logistic regression models to determine the factors that influence the use of blinding and randomization in phase IV clinical trials which evaluated drug safety.

3. We followed a strict analysis process that is widely used in analyzing the data from ClinicalTrials.gov and makes the results convincing.

Limitations

1. Some clinical trials are not registered in ClinicalTrials.gov.

2. There were some unavoidable missing data for certain data fields which may induce some bias into the results.

INTRODUCTION

Drug adverse reaction is a major global health concern accounting for more than 2 million injuries, hospitalizations, and deaths each year in the US alone,[1] and associated with billions of US dollars in costs every year in all developed countries.[2] Although rigorous premarketing studies are required for all new drugs,[3 4] the drug's safety profile of a drug at the time of regulatory approval is often incomplete due to the characteristics of phase I-III trials such as limited sample sizes, short duration and strict inclusion/exclusion criteria.[5] Approximately 20% of drugs acquire new black box warnings postmarketing, and 4% of drugs were ultimately withdrawn for safety reasons[6]. In 2007, the Food and Drug Administration was authorized 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 premarket phase I-III trials, phase IV studies can evaluate drug safety in a related 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 first launched in February 2000.[10] Beginning in 2005, the International Committee of Medical Journal Editors (ICMJE) 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 were obliged by FDAAA to register trials and report key data elements and basic trial results at ClinicalTrials.gov.[12] Hence, ClinicalTrials.gov is considered to be the most widely utilized source for clinical trial information worldwide.[13-15] Harnessing this expansive resource will enable us to gain a deeper understanding of the 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 requiring greater attention. We focus on data elements that are desirable for generating reliable evidence from clinical trials, including factors associated with the use of randomization 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 from the website on March 18, 2015. The data set was locked, and a relational database was designed to facilitate analysis.[15 16]

Study selection

Two authors (XJZ and YZ) each selected the eligible studies and summarized their results. Figure 1 shows the complete process of selection. Our analysis was restricted to phase IV clinical trials registered between January 1, 2004 and December 31, 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 the ClinicalTrials.gov website. Observational studies (n=981), expanded-access studies (n=10) and other studies that investigated “medical devices”, “vaccines” or other products were removed (n=5,878). On the ClinicalTrials.gov website, the “Endpoint Classification” section indicated the primary endpoint of the study, such as bio-equivalence, pharmacokinetics, safety, efficacy and so on. Additionally, based on the information in the “Primary Purpose” section, studies can 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 purpose is "Treatment" and primary endpoint is "Safety study" or "Safety/efficacy study" using ''Primary Purpose'' and "Endpoint Classification" sections. Finally, 4722 eligible phase IV trials assessing drug safety were included in our analysis.

The included trials were regrouped into three derivative databases according to the 3 clinical specialties—mental health, oncology and cardiovascular diseases. For this regrouping, we used the information in the "Conditions" section from ClinicalTrials.gov and the information on the classification of studies based on category by ClinicalTrials.gov. We used the study groups classified by the ClinicalTrials.gov as the matching database and regrouped the original database by matching the NCT number of each study between the original database and the matching database.

Data collection

Trial data are reported by the trial sponsors or investigators, as required by ClinicalTrials.gov.[17] Each record contains a set of data elements describing the study's conditions, enrollment, 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 summarized below. Funding sources were divided using the information in the "Sponsor_Collaborators"" and "Funded_Bys" sections. All trials were divided into 6 groups: industry, NIH, other, US federal (excluding NIH), university and hospital. 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. The funding source was defined as industry if the lead sponsor was from industry or if any collaborators were from industry and none from the NIH.. The funding source was defined as from US federal sources if the sponsor 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. The funding source 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, funding source was defined as other sources. The start dates of trials can be obtained from the "Start_Date" section. Information on the appointment of a data monitoring committee (DMC) became available in April 2007, and is not a required field[18], so in our study, we did not consider this information . The other variables' classification was based on the information in related 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 nonrandomized and the value of blinding as open if the information was missing.[15]

Analysis Methods

The characteristics of the trials were first assessed overall, by two endpoint classifications and by three clinical specialties. The assessments included the overall status, enrollment, intervention model, funding source and so on. The percentage of trials registered before and after the trial start date was determined by comparing the date first received by ClinicalTrials.gov with the start date.

According to the principles of the binomial and Poisson distributions, if investigators plan to observe at least 1 case of adverse events with a probability of occurrence less than 1%, 0.5% or 1‰, the enrollment should be larger than 300, 600 or 3000, respectively (Table 1).[19] Hence, we divided

the included trials into five types: trials with sample size less than 300, between 300 and 599, between 600 and 2999, 3000 or above and missing. Frequencies and percentages are provided for categorical characteristics; medians and interquartile ranges (IQRs) are provided for continuous characteristics.

Table 1 Numbers of patients necessary to enroll

Expected incidence of adverse reaction	Numbers of patients to enroll to detect at least 1 event
1 in 100	300
1 in 200	600
1 in 1000	⩾3000

Logistic regression analysis was used to evaluate factors associated with the use of randomization and blinding. A full model containing 8 characteristics was developed and adjusted odds ratios (ORs) with Wald 95% confidence intervals were calculate for the 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 endpoint 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 version 9.2 (SAS Institute) was used for all statistical analyses.

RESULTS

From January 1, 2004, to December 13, 2014, 18,642 phase IV trials were registered at ClinicalTrials.gov. Of these trials, 5,557 were interventional studies related to drug safety and 4,722 studies' primary purpose was "Treatment". The number of trials evaluating safety alone was 330, which was significantly less 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 4,722 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). A total of 72.7% of these phase IV trials used randomization 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. Of these studies, most were small (median enrollment: 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.

Table 2 Characteristics of Included Trials with Different Endpoint Classifications

	No./Total No. (%)		
	All, 2004-2014	Endpoint classification Safety	Endpoint classification Safety/ Efficacy
	N=4722	N=330	N=4392
Overall status [a]			
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)
Enrollment, 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(.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			

Randomized	3435(72.7)	187(56.7)	3248(74.0)
Non-Randomized	1252(26.5)	135(40.9)	1117(25.4)
Missing	35(0.7)	8(2.4)	27(0.6)
Masking			
Double-Blind	1764(37.4)	90(27.3)	1674(38.1)
Single Blind	332(7.0)	22(6.7)	310(7.1)
Open label	2620(55.5)	218(66.1)	2402(54.7)
Missing	6(0.1)	0(0.0)	6(0.1)
Sex,%			
Female only	337(7.1)	19(5.8)	318(7.2)
Male 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 (<18y)	762(16.1)	72(21.8)	690(15.7)
Excluded elderly (>65 y)	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)
U.S.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 [b]			
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)
Study registration [b]			
Start before submission	131 (2.8)	7(2.1)	124(2.8)
Start after submission	4591(97.2)	323(97.9)	4268(97.2)

Note:

a: "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 reason.

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

A total of 68.5% of the phase IV trials focused on drug safety alone had enrollment less than 300, and only 3.9% (n=13) of the phase IV trials focused on drug safety alone enrolled more than 3000 patients. 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 median enrollment of 103.0 (IQR, 48.00-251.5). Compared with studies evaluating both safety and efficacy, phase IV trials focused only on drug safety showed larger proportion of studies using single group assignment (41.8% *vs* 25.9%) and small proportion using randomization (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 trials assessing safety/efficacy (34.0% *vs* 42.8%).

Table 3 shows the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular, oncology and mental health). Of these 3 categories, the cardiovascular diseases trials were most numerous (n=601, 12.7%). Cardiovascular trials had larger enrollment (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). The use of randomization 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). Among trials including females only, oncology trials comprised the largest group: at 13.5% compared to 1.3% for cardiovascular trials and 2.3% for mental health trials. It is noteworthy that nearly two thirds of mental health trials (65.0%) excluded elderly subjects. Geographical differences were also apparent. Mental health related trials showed 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 paid more attention to mental health trials: 8.9% vs 1.0% for cardiovascular trials and 0.4% for oncology trials.

Table 3 Characteristics of Inclusded Trials in Different Clinical Specialties

	No./Total No. (%)			
	All, 2004-2014	Cardiovascular diseases	Oncology	Mental Health
	N=4722	N=601	N=251	N=594
Overall status [a]				
Not yet recruiting	196(4.2)	39(6.5)	7(2.8)	16(2.7)
Recruiting	941(19.9)	129(21.5)	84(33.5)	106(17.8)
Completed	2858(60.5)	331(55.1)	106(42.2)	404(68)
Suspended	17(0.4)	2(0.3)	2(0.8)	0(0.0)
Terminated	304(6.4)	39(6.5)	13(5.2)	30(5.1)
Withdrawn	87(1.8)	16(2.7)	5(2.0)	10(1.7)
Active, not recruiting	274(5.8)	42(7)	28(11.2)	26(4.4)
Enrolling by invitation	45(1.0)	3(0.5)	6(2.4)	2(0.3)
Enrollment, median (IQR)	104.0(48.0,258.0)	163.0(70.0,400.0)	100.0(48.0,200.0)	88.0(40.0,226.0)
1-299	3585(75.9)	391(65.1)	205(81.7)	475(80)
300-599	629(13.3)	83(13.8)	24(9.6)	80(13.5)
600-2999	344(7.3)	90(15)	14(5.6)	27(4.5)
⩾3000	57(1.2)	17(2.8)	2(0.8)	5(0.8)
Missing	107(2.3)	20(3.3)	6(2.4)	7(1.2)
Intervention Model				
Crossover Assignment	271(5.7)	23(3.8)	5(2.0)	35(5.9)
Single Group Assignment	1276(27.0)	10(1.7)	4(1.6)	8(1.3)
Parallel Assignment	3116(66.0)	451(75.0)	104(41.4)	359(60.4)
Factorial Assignment	52(1.1)	115(19.1)	138(55.0)	191(32.2)
Missing	7(0.1)	2(0.3)	0(0.0)	1(0.2)
Allocation		s		
Randomized	3435(72.7)	489(81.4)	108(43.0)	401(67.5)
Non-Randomized	1252(26.5)	108(18.0)	139(55.4)	190(32.0)
Missing	35(0.7)	4(0.7)	4(1.6)	3(0.5)
Masking				
Double-Blind	1764(37.4)	225(37.4)	35(13.9)	305(51.3)
Single blind	332(7.0)	53(8.8)	9(3.6)	35(5.9)
Open label	2620(55.5)	322(53.6)	207(82.5)	252(42.4)
Missing	6(0.1)	1(0.2)	0(0)	2(0.3)
Sex, %				
Female only	337(7.1)	8(1.3)	34(13.5)	13(2.2)
Male only	159(3.4)	9(1.5)	21(8.4)	31(5.2)
Both	4224(89.5)	584(97.2)	196(78.1)	550(92.6)
Missing	2(0.0)	0(0.0)	0(0.0)	0(0.0)
Included children (<18 y)	762(16.1)	33(5.5)	35(13.9)	112(18.9)
Excluded elder (>65 y)	1362(28.8)	47(7.8)	24(9.6)	386(65.0)

Lead sponsor				
Industry	2711(57.4)	305(50.7)	148(59.0)	360(60.6)
NIH	97(2.1)	6(1.0)	1(0.4)	53(8.9)
U.S.Federal	30(0.6)	2(0.3)	0(0.0)	3(0.5)
Hospitals and similar institutions	682(14.4)	119(19.8)	39(15.5)	55(9.3)
Universities and similar institutions	758(16.1)	108(18.0)	32(12.7)	80(13.5)
Other	444(9.4)	61(10.1)	31(12.4)	43(7.2)
Region [b]				
Africa	168(3.6)	20(3.3)	13(5.2)	13(2.2)
Asia and Pacific	1332(28.2)	210(34.9)	106(42.2)	137(23.1)
Central and South America	324(6.9)	26(4.3)	14(5.6)	41(6.9)
Europe	1250(26.5)	167(27.8)	82(32.7)	76(12.8)
Middle East	239(5.1)	24(4.0)	19(7.6)	25(4.2)
North America	1626(34.4)	172(28.6)	59(23.5)	314(52.9)
Missing	506(10.7)	57(9.5)	26(10.4)	56(9.4)
Study registration				
Start before submission	131 (2.8)	27(4.5)	10(4.0)	12(2.0)
Start after submission	4591(97.2)	574(95.5)	241(96.0)	582(98.0)

Note:

a: "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 reason.

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

Table 4 shows the results of the regression analysis. This analysis compared the characteristics of the trials that related to the use of blinding and randomization. There are 3,361 valid observations in the regression model for Blinding and 3,355 valid observations in the regression model for Randomization. Different clinical specialties can influence the use of blinding and randomization. Oncology trials were less likely to use both blinding (adjusted OR, 0.33; 95% CI, 0.18-0.63) and randomization (adjusted OR, 0.42; 95% CI, 0.28-0.63). Mental health trials were more likely to use blinding (adjusted OR, 3.35; 95% CI, 2.56-4.38). 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, 1.08-1.60).

Table 4 Regression Analyses of Included Trials and the Reported Use of Blinding and Randomization

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

Abbreviations: NIH, National Institutes of Health; OR, odds ratio.

a: 3,361 studies were included in this regression model.

b: 3,355 studies were included in this regression model.

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 endpoint classifications and clinical specialties were compared. We also analyzed the factors associated with trial quality. Thus, this analysis presents 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 is the greatest concern in phase IV trials involving the safety surveillance of an approved drug. Small phase IV trials may be used to evaluate the effectiveness of a given drug in a special patient subgroup, or in special situations.[5] However, our study only included phase IV

trials with “safety” as an endpoint, and most of these trials (77.6%) had an enrollment of less than 300. In the phase IV trials with safety as the primary endpoint, the average sample size was only 104. Thus, these small trials might not have sufficient power to detect adverse events (AEs), especially less common adverse events.[19] Perhaps 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 excepted for each adverse event. For example, to observe an AE with an occurrence probability of 1.5‰, the China food and drug administration (CFDA) requires that the enrollment 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 use the larger value.

For trials with controls, randomization and blinding are both desirable elements for generating reliable evidence. Compared to the phase IV clinical trials sponsored by industry, trials sponsored by a university or college were more likely to use blinding. The methodological differences intrials were also evident among therapeutic areas. Oncology trials were less likely to use randomization and blinding, even after adjustment for other factors. It is unclear that these studies are of sufficient quality to meet the medical needs of the growing oncology population.

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 plays a more important role in phase IV trials. The 30.5% of phase IV trials including the Asia and Pacific area is a significant improvement over prior analyses of all clinical trials (13.5%).[15] The inclusion of various 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%). Robert's research[15] revealed that 3.3% of all interventional clinical trials registered from October 2007 through September 2010 were terminated or withdrawn. We further analyzed the conditions, endpoints 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 characterization 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 are not registered in ClinicalTrials.gov , 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 are some missing data for certain data fields, which may introduce some bias into the results. Third, as described in the "Methods" section, we use the endpoint 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 endpoint for trials evaluating both safety and efficacy.

CONCLUSION

We found that the phase IV trials enterprise related to drug safety in ClinicalTrials.gov are 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 emphasized for phase IV trials with safety as the primary endpoint.

Figure Legends:

Figure 1 Consortium diagram: Inclusion criteria and methods for study selection

FUNDINGS STATEMENT

This study was sponsored by the National Nature Science Foundation of China (NO. 81502895, 81373105), a grant from the key discipline for construction of evidence-based public health in Shanghai (NO. 12GWZX0602) and the Fourth Round of Three-year Action Plan on Public Health Discipline and Talent Program: Evidence-based Public Health and Health Economics (No. 15GWZK0901)

COMPETING INTERESTS

The authors declare that they have no competing interests.

CONTRIBUTORSHIP STATEMENT

Xinji Zhang and Yuan Zhang contributed equally in conceiving this project, facilitating protocol, analyzing data and drafting this manuscript. Yexiao Fei leaded the development of performance-based incentives and revised the manuscript critically. Tianyi Zhang and XiaoJing Guo gave their time and effort to modify the programs. Jia He provided expertise for the overall design of the study, and revised and approved the manuscript.

ACKONWLEDGEMENTS

We gratefully acknowledge for his assistance in designing the study. We also acknowledge

others who gave their time and effort to this study.

DATA SHARING STATEMENT

The analyzed dataset was upgraded on the Datadryad.org website .The title of the dataset used in this revision is “phase IV clinical studies received by ClinicalTrials.gov between 2004-2014 ”

URL: http://datadryad.org/review?doi=doi:10.5061/dryad.3t6sc

REFERENCE

1. Harpaz R, ., Dumouchel W, ., Shah NH, et al. Novel data-mining methodologies for adverse drug event discovery and analysis. Clinical Pharmacology & Therapeutics 2012;**91**(6):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. Clinical Pharmacology & Therapeutics 1995;**58**(1):108-17 doi: 10.1016/0009-9236(95)90078-0
3. Englev E, Petersen K. [ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives]. Ugeskrift for laeger 2003;**165**(16):1659-62
4. Dixon JR. The international conference on harmonization good clinical practice guideline. Quality Assurance: Good Practice, Regulation, and Law 1999;**6**(2):65-74
5. Gale EAM. Post-marketing studies of new insulins: sales or science? British Medical Journal 2012;**344** doi: 10.1136/bmj.e3974.
6. Lasser KE, Allen PD, Woolhandler SJ, et al. Timing of new black box warnings and withdrawals for prescription medications. Jama-Journal of the American Medical Association 2002;**287**(17):2215-20 doi: 10.1001/jama.287.17.2215.
7. Food and Drug Administration. Food and Drug Administration Amendments Act (FDAAA) of 2007 .2007. http://www.fda.gov/downloads/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FoodandDrugAdministrationAmendmentsActof2007/FDAAAImplementationChart/UCM213016.pdf (accessed August 4, 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. The Journal of Clinical Pharmacology 2007;**47**(9):1074-86
9. Schmidt LG, Grohmann R, Helmchen H, et al. ADVERSE DRUG-REACTIONS - AN EPIDEMIOLOGICAL-STUDY AT PSYCHIATRIC-HOSPITALS. Acta Psychiatrica Scandinavica 1984;**70**(1):77-89 doi: 10.1111/j.1600-0447.1984.tb01185.x
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**(4):275-77
12. Zarin DA, Tse T. Moving towards transparency of clinical trials. Science (New York, NY) 2008;**319**(5868):1340
13. Cihoric N, Tsikkinis A, van Rhoon G, et al. Hyperthermia-related clinical trials on cancer treatment within the ClinicalTrials.gov registry. International Journal of Hyperthermia 2015;**31**(6):609-14 doi: 10.3109/02656736.2015.1040471.
14. Shields KE, Lyerly AD. Exclusion of pregnant women from industry-sponsored clinical trials. Obstetrics and

gynecology 2013;**122**(5):1077-81 doi: 10.1097/AOG.0b013e3182a9ca67[published Online First: Epub Date]|.

15. Califf RM, Zarin DA, Kramer JM, et al. Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007-2010. Jama-Journal of the American Medical Association 2012;**307**(17):1838-47 doi: 10.1001/jama.2012.3424.
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**(3) doi: 10.1371/journal.pone.003367717. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov Results Database - Update and Key Issues. New England Journal of Medicine 2011;**364**(9):852-60 doi: 10.1056/NEJMsa1012065.
18. Hill KD, Chiswell K, Califf RM, et al. Characteristics of pediatric cardiovascular clinical trials registered on ClinicalTrials.gov. American Heart Journal 2014;**167**(6):921-U195 doi: 10.1016/j.ahj.2014.02.002
19. Suvarna V. Phase IV of Drug Development. Perspectives in clinical research 2010;**1**(2):57-60
20. China Food and Drug Administratio: **Drug Registration Management Measures(2007)**.[2007] http://www.sfda.gov.cn/WS01/CL0053/24529.html. (accessed August 4, 2015) [in Chinese]

Figure 1 Consortium diagram

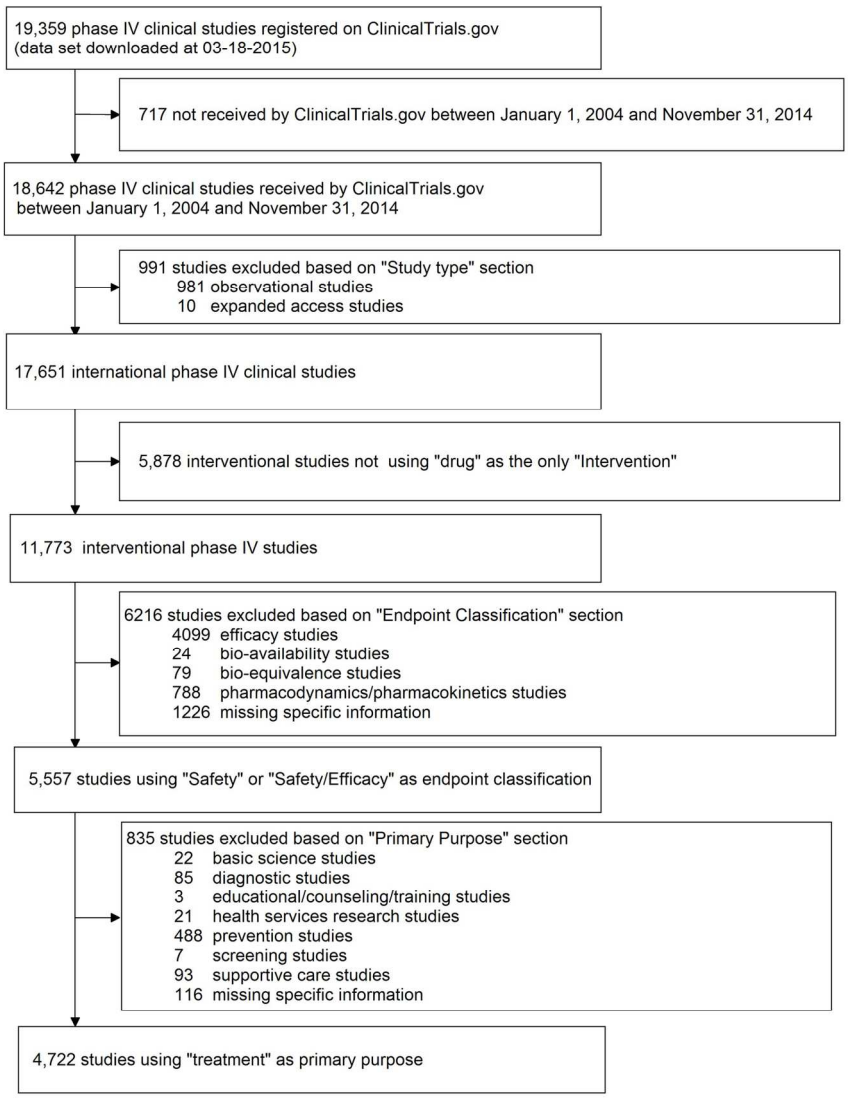


Figure 1 Consortium diagram: Inclusion criteria and methods for study selection
Figure 1
167x236mm (300 x 300 DPI)

BMJ Open

An overview of phase IV clinical trials for post-market drug safety surveillance: a status report from ClinicalTrials.gov registry

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010643.R2
Article Type:	Research
Date Submitted by the Author:	21-Jun-2016
Complete List of Authors:	Zhang, Xinji; Second Mil Med Univ Zhang, Yuan; Second Mil Med Univ Ye, Xiaofei; Second Mil Med Univ Guo, Xiaojing; Second Mil Med Univ Zhang, Tianyi; Second Mil Med Univ He, Jia; Second Mil Med Univ,
Primary Subject Heading:	Medical management
Secondary Subject Heading:	Health informatics, Pharmacology and therapeutics
Keywords:	Clinical trials < THERAPEUTICS, Adverse events < THERAPEUTICS, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

An overview of phase IV clinical trials for post-market drug safety surveillance: a status report from ClinicalTrials.gov registry

Xinji Zhang [1*], Yuan Zhang [1*], Xiaofei Ye[1], Xiaojing Guo[1],Tianyi Zhang[1], Jia He [1△]

Author Affiliation and Address:

[1] Department of Health Statistics, Second Military Medical University, Shanghai, 200433, China

[*] These authors contributed equally and are co-first authors of this article.

[△]Corresponding author:

Prof Jia He, Department of Health Statistics, Second Military Medical University, No. 800 Xiangyin Road, Shanghai 200433, China

Tel: +86-021-81871441, Fax: +86-021-81871441, E-mail: hejia63@yeah.net

Key words or phrases for search: clinical trials; health and safety; adverse events

Number of words: 2,849

Abstract

Objective: Phase IV trials are often used to survey 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 evaluate drug safety using ClinicalTrials.gov data

Methods: A data set of 19,359 phase IV clinical studies registered with ClinicalTrials.gov was downloaded. The Characteristics of the phase IV trials focusing only on safety were compared with the ones evaluating both safety and efficacy. We also compared the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular diseases, mental health, and oncology). Multivariable logistic regression was used to evaluate factors associated with the use of blinding and randomization.

Results: A total of 4,772 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 enrollment less than 300 with 96.5% less than 3000. A total of 8.2% of trials were terminated or withdrawn. Factors associated with the usage of blinding and randomization included the intervention model, clinical specialty and lead sponsor .

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

Article summary

Strengths and limitations of this study

Strengths

1. We provided a comprehensive descriptive assessment of the current portfolio of phase IV clinical trials evaluating drug safety.

2. We employed logistic regression models to determine the factors that influence the use of blinding and randomization in phase IV clinical trials which evaluated drug safety.

3. We followed a strict analysis process that is widely used in analyzing the data from ClinicalTrials.gov and makes the results convincing.

Limitations

1. Some clinical trials are not registered in ClinicalTrials.gov.

2. There were some unavoidable missing data for certain data fields which may induce some bias into the results.

INTRODUCTION

Drug adverse reaction is a major global health concern accounting for more than 2 million injuries, hospitalizations, and deaths each year in the US alone,[1] and associated with billions of US dollars in costs every year in all developed countries.[2] Although rigorous premarketing studies are required for all new drugs,[3 4] the drug's safety profile of a drug at the time of regulatory approval is often incomplete due to the characteristics of phase I-III trials such as limited sample sizes, short duration and strict inclusion/exclusion criteria.[5] Approximately 20% of drugs acquire new black box warnings postmarketing, and 4% of drugs were ultimately withdrawn for safety reasons[6]. In 2007, the Food and Drug Administration was authorized 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 premarket phase I-III trials, phase IV studies can evaluate drug safety in a related 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 first launched in February 2000.[10] Beginning in 2005, the International Committee of Medical Journal Editors (ICMJE) 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 were obliged by FDAAA to register trials and report key data elements and basic trial results at ClinicalTrials.gov.[12] Hence, ClinicalTrials.gov is considered to be the most widely utilized source for clinical trial information worldwide.[13-15] Harnessing this expansive resource will enable us to gain a deeper understanding of the 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 requiring greater attention. We focus on data elements that are desirable for generating reliable evidence from clinical trials, including factors associated with the use of randomization 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 from the website on March 18, 2015. The data set was locked, and a relational database was designed to facilitate analysis.[15 16]

Study selection

Two authors (XJZ and YZ) each selected the eligible studies and summarized their results. Figure 1 shows the complete process of selection. Our analysis was restricted to phase IV clinical trials registered between January 1, 2004 and December 31, 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 the ClinicalTrials.gov website. Observational studies (n=981), expanded-access studies (n=10) and other studies that investigated "medical devices", "vaccines" or other products were removed (n=5,878). On the ClinicalTrials.gov website, the "Endpoint Classification" section indicated the primary endpoint of the study, such as bio-equivalence, pharmacokinetics, safety, efficacy and so on. Additionally, based on the information in the "Primary Purpose" section, studies can 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 purpose is "Treatment" and primary endpoint is "Safety study" or "Safety/efficacy study" using ''Primary Purpose'' and "Endpoint Classification" sections. Finally, 4722 eligible phase IV trials assessing drug safety were included in our analysis.

The included trials were regrouped into three derivative databases according to the 3 clinical specialties—mental health, oncology and cardiovascular diseases. For this regrouping, we used the information in the "Conditions" section from ClinicalTrials.gov and the information on the classification of studies based on category by ClinicalTrials.gov. We used the study groups classified by the ClinicalTrials.gov as the matching database and regrouped the original database by matching the NCT number of each study between the original database and the matching database.

Data collection

Trial data are reported by the trial sponsors or investigators, as required by ClinicalTrials.gov.[17] Each record contains a set of data elements describing the study's conditions, enrollment, 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 summarized below. Funding sources were divided using the information in the "Sponsor_Collaborators"" and "Funded_Bys" sections. All trials were divided into 6 groups: industry, NIH, other, US federal (excluding NIH), university and hospital. 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. The funding source was defined as industry if the lead sponsor was from industry or if any collaborators were from industry and none from the NIH.. The funding source was defined as from US federal sources if the sponsor 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. The funding source 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, funding source was defined as other sources. The start dates of trials can be obtained from the "Start_Date" section. Information on the appointment of a data monitoring committee (DMC) became available in April 2007, and is not a required field[18], so in our study, we did not consider this information . The other variables' classification was based on the information in related 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 nonrandomized and the value of blinding as open if the information was missing.[15]

Analysis Methods

The characteristics of the trials were first assessed overall, by two endpoint classifications and by three clinical specialties. The assessments included the overall status, enrollment, intervention model, funding source and so on. The percentage of trials registered before and after the trial start date was determined by comparing the date first received by ClinicalTrials.gov with the start date.

According to the principles of the binomial and Poisson distributions, if investigators plan to observe at least 1 case of adverse events with a probability of occurrence less than 1%, 0.5% or 1‰, the enrollment should be larger than 300, 600 or 3000, respectively (Table 1).[19] Hence, we divided

the included trials into five types: trials with sample size less than 300, between 300 and 599, between 600 and 2999, 3000 or above and missing. Frequencies and percentages are provided for categorical characteristics; medians and interquartile ranges (IQRs) are provided for continuous characteristics.

Table 1 Numbers of patients necessary to enroll

Expected incidence of adverse reaction	Numbers of patients to enroll to detect at least 1 event
1 in 100	300
1 in 200	600
1 in 1000	⩾3000

Logistic regression analysis was used to evaluate factors associated with the use of randomization and blinding. A full model containing 8 characteristics was developed and adjusted odds ratios (ORs) with Wald 95% confidence intervals were calculate for the 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 endpoint 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 version 9.2 (SAS Institute) was used for all statistical analyses.

RESULTS

From January 1, 2004, to December 13, 2014, 18,642 phase IV trials were registered at ClinicalTrials.gov. Of these trials, 5,557 were interventional studies related to drug safety and 4,722 studies' primary purpose was "Treatment". The number of trials evaluating safety alone was 330, which was significantly less 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 4,722 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). A total of 72.7% of these phase IV trials used randomization 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. Of these studies, most were small (median enrollment: 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.

Table 2 Characteristics of Included Trials with Different Endpoint Classifications

	No./Total No. (%)		
	All, 2004-2014	Endpoint classification Safety	Endpoint classification Safety/ Efficacy
	N=4722	N=330	N=4392
Overall status [a]			
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)
Enrollment, 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(.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			

Randomized	3435(72.7)	187(56.7)	3248(74.0)
Non-Randomized	1252(26.5)	135(40.9)	1117(25.4)
Missing	35(0.7)	8(2.4)	27(0.6)
Masking			
Double-Blind	1764(37.4)	90(27.3)	1674(38.1)
Single Blind	332(7.0)	22(6.7)	310(7.1)
Open label	2620(55.5)	218(66.1)	2402(54.7)
Missing	6(0.1)	0(0.0)	6(0.1)
Sex,%			
Female only	337(7.1)	19(5.8)	318(7.2)
Male 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 (<18y)	762(16.1)	72(21.8)	690(15.7)
Excluded elderly (>65 y)	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)
U.S.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 [b]			
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)
Study registration [b]			
Start before submission	131 (2.8)	7(2.1)	124(2.8)
Start after submission	4591(97.2)	323(97.9)	4268(97.2)

Note:

a: "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 reason.

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

A total of 68.5% of the phase IV trials focused on drug safety alone had enrollment less than 300, and only 3.9% (n=13) of the phase IV trials focused on drug safety alone enrolled more than 3000 patients. 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 median enrollment of 103.0 (IQR, 48.00-251.5). Compared with studies evaluating both safety and efficacy, phase IV trials focused only on drug safety showed larger proportion of studies using single group assignment (41.8% *vs* 25.9%) and small proportion using randomization (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 trials assessing safety/efficacy (34.0% *vs* 42.8%).

Table 3 shows the characteristics of the phase IV trials in 3 major therapeutic areas (cardiovascular, oncology and mental health). Of these 3 categories, the cardiovascular diseases trials were most numerous (n=601, 12.7%). Cardiovascular trials had larger enrollment (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). The use of randomization 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). Among trials including females only, oncology trials comprised the largest group: at 13.5% compared to 1.3% for cardiovascular trials and 2.3% for mental health trials. It is noteworthy that nearly two thirds of mental health trials (65.0%) excluded elderly subjects. Geographical differences were also apparent. Mental health related trials showed 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 paid more attention to mental health trials: 8.9% vs 1.0% for cardiovascular trials and 0.4% for oncology trials.

Table 3 Characteristics of Inclusded Trials in Different Clinical Specialties

	No./Total No. (%)			
	All, 2004-2014	Cardiovascular diseases	Oncology	Mental Health
	N=4722	N=601	N=251	N=594
Overall status [a]				
Not yet recruiting	196(4.2)	39(6.5)	7(2.8)	16(2.7)
Recruiting	941(19.9)	129(21.5)	84(33.5)	106(17.8)
Completed	2858(60.5)	331(55.1)	106(42.2)	404(68)
Suspended	17(0.4)	2(0.3)	2(0.8)	0(0.0)
Terminated	304(6.4)	39(6.5)	13(5.2)	30(5.1)
Withdrawn	87(1.8)	16(2.7)	5(2.0)	10(1.7)
Active, not recruiting	274(5.8)	42(7)	28(11.2)	26(4.4)
Enrolling by invitation	45(1.0)	3(0.5)	6(2.4)	2(0.3)
Enrollment, median (IQR)	104.0(48.0,258.0)	163.0(70.0,400.0)	100.0(48.0,200.0)	88.0(40.0,226.0)
1-299	3585(75.9)	391(65.1)	205(81.7)	475(80)
300-599	629(13.3)	83(13.8)	24(9.6)	80(13.5)
600-2999	344(7.3)	90(15)	14(5.6)	27(4.5)
⩾3000	57(1.2)	17(2.8)	2(0.8)	5(0.8)
Missing	107(2.3)	20(3.3)	6(2.4)	7(1.2)
Intervention Model				
Crossover Assignment	271(5.7)	23(3.8)	5(2.0)	35(5.9)
Single Group Assignment	1276(27.0)	10(1.7)	4(1.6)	8(1.3)
Parallel Assignment	3116(66.0)	451(75.0)	104(41.4)	359(60.4)
Factorial Assignment	52(1.1)	115(19.1)	138(55.0)	191(32.2)
Missing	7(0.1)	2(0.3)	0(0.0)	1(0.2)
Allocation		s		
Randomized	3435(72.7)	489(81.4)	108(43.0)	401(67.5)
Non-Randomized	1252(26.5)	108(18.0)	139(55.4)	190(32.0)
Missing	35(0.7)	4(0.7)	4(1.6)	3(0.5)
Masking				
Double-Blind	1764(37.4)	225(37.4)	35(13.9)	305(51.3)
Single blind	332(7.0)	53(8.8)	9(3.6)	35(5.9)
Open label	2620(55.5)	322(53.6)	207(82.5)	252(42.4)
Missing	6(0.1)	1(0.2)	0(0)	2(0.3)
Sex, %				
Female only	337(7.1)	8(1.3)	34(13.5)	13(2.2)
Male only	159(3.4)	9(1.5)	21(8.4)	31(5.2)
Both	4224(89.5)	584(97.2)	196(78.1)	550(92.6)
Missing	2(0.0)	0(0.0)	0(0.0)	0(0.0)
Included children (<18 y)	762(16.1)	33(5.5)	35(13.9)	112(18.9)
Excluded elder (>65 y)	1362(28.8)	47(7.8)	24(9.6)	386(65.0)

Lead sponsor				
Industry	2711(57.4)	305(50.7)	148(59.0)	360(60.6)
NIH	97(2.1)	6(1.0)	1(0.4)	53(8.9)
U.S.Federal	30(0.6)	2(0.3)	0(0.0)	3(0.5)
Hospitals and similar institutions	682(14.4)	119(19.8)	39(15.5)	55(9.3)
Universities and similar institutions	758(16.1)	108(18.0)	32(12.7)	80(13.5)
Other	444(9.4)	61(10.1)	31(12.4)	43(7.2)
Region [b]				
Africa	168(3.6)	20(3.3)	13(5.2)	13(2.2)
Asia and Pacific	1332(28.2)	210(34.9)	106(42.2)	137(23.1)
Central and South America	324(6.9)	26(4.3)	14(5.6)	41(6.9)
Europe	1250(26.5)	167(27.8)	82(32.7)	76(12.8)
Middle East	239(5.1)	24(4.0)	19(7.6)	25(4.2)
North America	1626(34.4)	172(28.6)	59(23.5)	314(52.9)
Missing	506(10.7)	57(9.5)	26(10.4)	56(9.4)
Study registration				
Start before submission	131 (2.8)	27(4.5)	10(4.0)	12(2.0)
Start after submission	4591(97.2)	574(95.5)	241(96.0)	582(98.0)

Note:

a: "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 reason.

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

Table 4 shows the results of the regression analysis. This analysis compared the characteristics of the trials that related to the use of blinding and randomization. There are 3,361 valid observations in the regression model for Blinding (blinded n=1950 vs open label n=1411) and 3,361 valid observations in the regression model for Randomization (randomized n=3234 ; not randomized n=127). Different clinical specialties can influence the use of blinding and randomization. Oncology trials were less likely to use both blinding (adjusted OR, 0.33; 95% CI, 0.18-0.63) and randomization (adjusted OR, 0.42; 95% CI, 0.28-0.63). Mental health trials were more likely to use blinding (adjusted OR, 3.35; 95% CI, 2.56-4.38). 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, 1.08-1.60).

Table 4 Regression Analyses of Included Trials and the Reported Use of Blinding and Randomization

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

Abbreviations: NIH, National Institutes of Health; OR, odds ratio.

a: 3,361 studies were included in this regression model.

b: 3,355 studies were included in this regression model.

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 endpoint classifications and clinical specialties were compared. We also analyzed the factors associated with trial quality. Thus, this analysis presents 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 is the greatest concern in phase IV trials involving the safety surveillance of an approved drug. Small phase IV trials may be used to evaluate the effectiveness of a given drug in

a special patient subgroup, or in special situations.[5] However, our study only included phase IV trials with "safety" as an endpoint, and most of these trials (77.6%) had an enrollment of less than 300. In the phase IV trials with safety as the primary endpoint, the average sample size was only 104. Thus, these small trials might not have sufficient power to detect adverse events (AEs), especially less common adverse events.[19] Perhaps 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 excepted for each adverse event. For example, to observe an AE with an occurrence probability of 1.5‰, the China food and drug administration (CFDA) requires that the enrollment 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 use the larger one.

Phase IV clinical trials can have various designs and single-arm, non-randomized or open-label studies are accepted. If randomization and blinding are feasible in the studies with control 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 randomization and blinding, which was consisted with the results of previous studies. One possible reason is that some of the oncology trials investigate individualized or personalized treatment and cannot use randomization or blinding. Due to the limitation of information on ClinicalTrials.gov website, it is difficult to check whether all the phase IV trials with control have appropriate design. However, researcher should pay attention to

randomization 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 plays a more important role in phase IV trials. The 30.5% of phase IV trials including the Asia and Pacific area is a significant improvement over prior analyses of all clinical trials (13.5%).[15] The inclusion of various 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%). Robert's research[15] revealed that 3.3% of all interventional clinical trials registered from October 2007 through September 2010 were terminated or withdrawn. We further analyzed the conditions, endpoints 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 characterization 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 are not registered in ClinicalTrials.gov, 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 are some missing data for certain data fields, which may introduce some bias into the results. Third, as described in the "Methods" section, we use the

endpoint 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 endpoint for trials evaluating both safety and efficacy.

CONCLUSION

We found that the phase IV trials enterprise related to drug safety in ClinicalTrials.gov are 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 emphasized for phase IV trials with safety as the primary endpoint.

Figure Legends:

Figure 1 Flow Chart of Inclusion and Exclusion

FUNDINGS STATEMENT

This study was sponsored by the National Nature Science Foundation of China (NO. 81502895, 81373105), a grant from the key discipline for construction of evidence-based public health in Shanghai (NO. 12GWZX0602) and the Fourth Round of Three-year Action Plan on Public Health Discipline and Talent Program: Evidence-based Public Health and Health Economics (No. 15GWZK0901)

COMPETING INTERESTS

The authors declare that they have no competing interests.

CONTRIBUTORSHIP STATEMENT

Xinji Zhang and Yuan Zhang contributed equally in conceiving this project, facilitating protocol,

analyzing data and drafting this manuscript. Yexiao Fei leaded the development of performance-based incentives and revised the manuscript critically. Tianyi Zhang and XiaoJing Guo gave their time and effort to modify the programs. Jia He provided expertise for the overall design of the study, and revised and approved the manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge the valuable advices on revision from Prof. Adelaide Doussau; Prof. Terence Campbell; Prof. Stephen Glasserthe and Prof. Mary Wiktorowicz .We also thank Jian Lu, PhD, for his assistance in designing the study. Besides, we acknowledged American Journal Experts, LLC for its professional copyediting service.

DATA SHARING STATEMENT

The analyzed dataset was upgraded on the Datadryad.org website .The title of the dataset used in this revision is “phase IV clinical studies received by ClinicalTrials.gov between 2004-2014 ”

URL: http://datadryad.org/review?doi=doi:10.5061/dryad.3t6sc

REFERENCE

1. Harpaz R, ., Dumouchel W, ., Shah NH, et al. Novel data-mining methodologies for adverse drug event discovery and analysis. Clinical Pharmacology & Therapeutics 2012;**91**(6):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. Clinical Pharmacology & Therapeutics 1995;**58**(1):108-17 doi: 10.1016/0009-9236(95)90078-0
3. Englev E, Petersen K. [ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives]. Ugeskrift for laeger 2003;**165**(16):1659-62
4. Dixon JR. The international conference on harmonization good clinical practice guideline. Quality Assurance: Good Practice, Regulation, and Law 1999;**6**(2):65-74
5. Gale EAM. Post-marketing studies of new insulins: sales or science? British Medical Journal 2012;**344** doi: 10.1136/bmj.e3974.
6. Lasser KE, Allen PD, Woolhandler SJ, et al. Timing of new black box warnings and withdrawals for prescription medications. Jama-Journal of the American Medical Association 2002;**287**(17):2215-20 doi: 10.1001/jama.287.17.2215.
7. Food and Drug Administration. Food and Drug Administration Amendments Act (FDAAA) of 2007 .2007. http://www.fda.gov/downloads/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FoodandDrugAdministrationAmendmentsActof2007/FDAAAImplementatio

nChart/UCM213016.pdf (accessed August 4, 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. The Journal of Clinical Pharmacology 2007;**47**(9):1074-86
9. Schmidt LG, Grohmann R, Helmchen H, et al. ADVERSE DRUG-REACTIONS - AN EPIDEMIOLOGICAL-STUDY AT PSYCHIATRIC-HOSPITALS. Acta Psychiatrica Scandinavica 1984;**70**(1):77-89 doi: 10.1111/j.1600-0447.1984.tb01185.x
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**(4):275-77
12. Zarin DA, Tse T. Moving towards transparency of clinical trials. Science (New York, NY) 2008;**319**(5868):1340
13. Cihoric N, Tsikkinis A, van Rhoon G, et al. Hyperthermia-related clinical trials on cancer treatment within the ClinicalTrials.gov registry. International Journal of Hyperthermia 2015;**31**(6):609-14 doi: 10.3109/02656736.2015.1040471.
14. Shields KE, Lyerly AD. Exclusion of pregnant women from industry-sponsored clinical trials. Obstetrics and gynecology 2013;**122**(5):1077-81 doi: 10.1097/AOG.0b013e3182a9ca67[published Online First: Epub Date]|.
15. Califf RM, Zarin DA, Kramer JM, et al. Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007-2010. Jama-Journal of the American Medical Association 2012;**307**(17):1838-47 doi: 10.1001/jama.2012.3424.
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**(3) doi: 10.1371/journal.pone.003367717. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov Results Database - Update and Key Issues. New England Journal of Medicine 2011;**364**(9):852-60 doi: 10.1056/NEJMsa1012065.
18. Hill KD, Chiswell K, Califf RM, et al. Characteristics of pediatric cardiovascular clinical trials registered on ClinicalTrials.gov. American Heart Journal 2014;**167**(6):921-U195 doi: 10.1016/j.ahj.2014.02.002
19. Suvarna V. Phase IV of Drug Development. Perspectives in clinical research 2010;**1**(2):57-60
20. China Food and Drug Administratio: **Drug Registration Management Measures(2007)**.[2007] http://www.sfda.gov.cn/WS01/CL0053/24529.html. (accessed August 4, 2015) [in Chinese]

Figure 1 Flow Chart of Inclusion and Exclusion

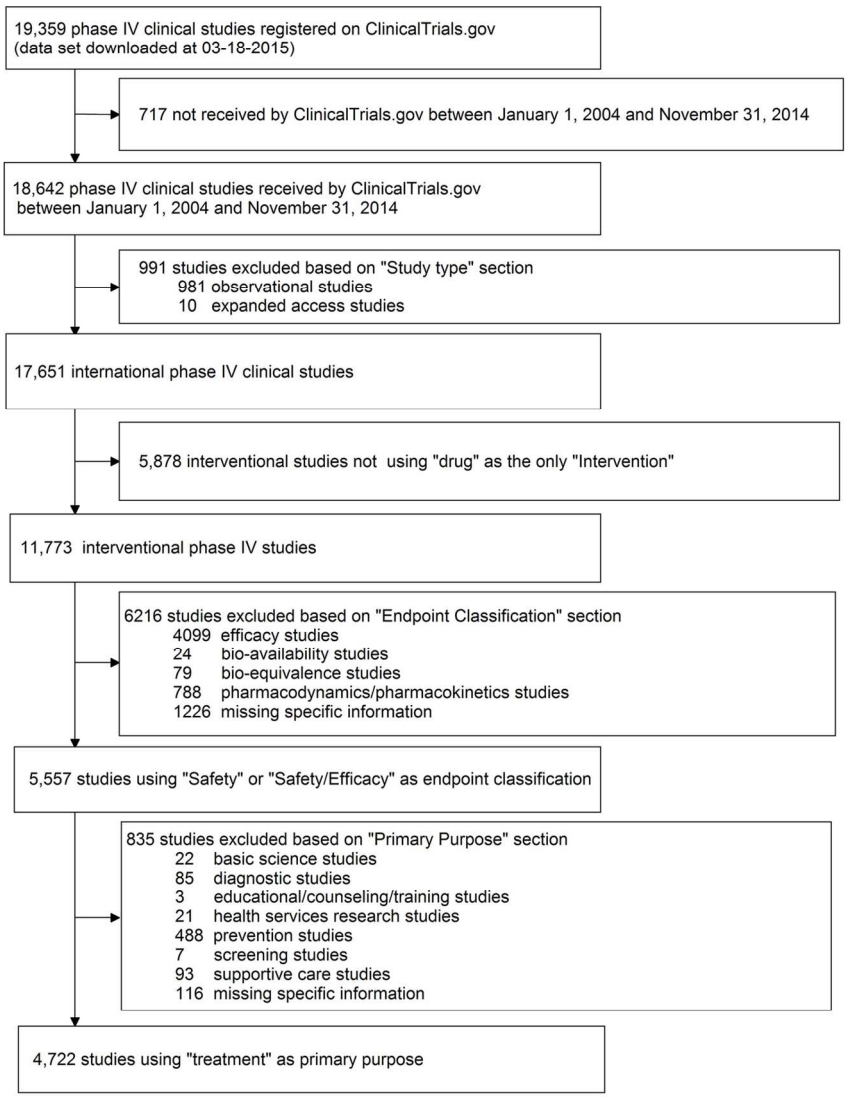


Figure 1 Flow Chart of Inclusion and Exclusion
Figure 1
146x207mm (300 x 300 DPI)

BMJ Open

An overview of phase IV clinical trials for post-market drug safety surveillance: a status report from the ClinicalTrials.gov registry

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010643.R3
Article Type:	Research
Date Submitted by the Author:	03-Aug-2016
Complete List of Authors:	Zhang, Xinji; Second Mil Med Univ Zhang, Yuan; Second Mil Med Univ Ye, Xiaofei; Second Mil Med Univ Guo, Xiaojing; Second Mil Med Univ Zhang, Tianyi; Second Mil Med Univ He, Jia; Second Mil Med Univ,
Primary Subject Heading:	Medical management
Secondary Subject Heading:	Health informatics, Pharmacology and therapeutics
Keywords:	Clinical trials < THERAPEUTICS, Adverse events < THERAPEUTICS, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

An overview of phase IV clinical trials for post-market drug safety surveillance: a status report from the ClinicalTrials.gov registry

Xinji Zhang [1*], Yuan Zhang [1*], Xiaofei Ye[1], Xiaojing Guo[1],Tianyi Zhang[1], Jia He [1△]

Author Affiliation and Address:

[1] Department of Health Statistics, Second Military Medical University, Shanghai, 200433, China

[*] These authors contribute equally and are co-first authors of this article.

[△]Corresponding author:

Prof Jia He, Department of Health Statistics, Second Military Medical University, No. 800 Xiangyin Road, Shanghai 200433, China

Tel: +86-021-81871441, Fax: +86-021-81871441, E-mail: hejia63@yeah.net

Key words or phrases for search: clinical trials; health and safety; adverse events

Number of words: 2,834

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 3 major therapeutic areas (cardiovascular diseases, mental health, and oncology). Multivariable logistic regression was used to evaluate factors associated with the use of blinding and randomization.

Results: A total of 4,772 phase IV trials were identified, including 330 focusing on drug safety alone and 4,392 evaluating both safety and efficacy. Most of the phase IV trials evaluating drug safety (75.9%) had enrollment less than 300 with 96.5% less than 3,000. Among these trials, 8.2% were terminated or withdrawn. Factors associated with the use of blinding and randomization 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 emphasized for phase IV trials with safety surveillance as main task.

Article summary

Strengths and limitations of this study

Strengths

1. We provided a comprehensive descriptive assessment of the current portfolio of phase IV clinical trials evaluating drug safety in the ClinicalTrials.gov registry.

2. We employed logistic regression models to determine the factors associated with the use of blinding and randomization in phase IV clinical trials which evaluated drug safety.

3. We followed a strict analysis process that was widely used in analyzing the data from ClinicalTrials.gov to arrive at convincing results.

Limitations

1. Some clinical trials were not registered in ClinicalTrials.gov.

2. There were some unavoidable missing data for certain data fields which might induce some bias into the results.

INTRODUCTION

Drug adverse reaction is a major global health concern accounting for more than 2 million injuries, hospitalizations, and deaths each year in the US alone,[1] and associated with billions of US dollars in cost every year in the developed countries.[2] Although rigorous 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 drugs were ultimately withdrawn for safety reasons.[6] In 2007, the Food and Drug Administration was authorized 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 (ICMJE) 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 the 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 randomization 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 March 18, 2015. A database was designed to facilitate analysis.[15 16]

Study selection

Two authors (XJZ and YZ) selected the eligible studies and summarized their results independently. Figure 1 showed the complete process of selection. Our analysis was restricted to phase IV clinical trials registered between January 1, 2004 and December 31, 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=5,878). On ClinicalTrials.gov, the "Endpoint Classification" section indicated the primary endpoint 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 endpoints were "Safety study" or "Safety/efficacy study" using ''Primary Purpose'' and "Endpoint Classification" sections. Finally, 4,722 eligible phase IV trials assessing drug safety alone or both safety and efficacy were included in our analysis.

The included trials were then categorized 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, enrollment, 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 summarized below. All trials were divided into 6 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 nonrandomized and the value of blinding as open if the information was missing.[15] In addition, the allocation or masking was reported as "Uncertain" if single-arm trials were registered as randomized or blind.

Statistical Analysis

The characteristics of the trials were assessed overall, by two endpoint classifications (safety only and safety/efficacy) and by three clinical specialties (mental health, oncology and cardiovascular diseases). The assessments included the study status, enrollment, 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 have a probability of occurrence 1%, 0.5% or 1‰, the enrollment 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 adverse events.[19] Hence, we divided the included trials into five types: trials with sample size less than 300, between 300 and 599, between 600 and 2,999 and 3,000 or above and missing. Frequencies and percentages were provided for categorical characteristics; medians and interquartile ranges (IQRs) were provided for continuous characteristics.

Table 1 Numbers of patients necessary to enroll

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

Logistic regression analysis was used to evaluate factors associated with the use of randomization and blinding. A full model containing 8 characteristics was developed and adjusted odds ratios (ORs) with Wald 95% confidence intervals 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 endpoint 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 version 9.2 (SAS Institute) was used for all statistical analyses.

RESULTS

From January 1, 2004, to December 13, 2014, 18,642 phase IV trials were registered at ClinicalTrials.gov. Of these trials, 4,722 phase IV trials related to drug safety were included in our study. Figure 1 showed the search process. The number of trials evaluating safety alone was 330,

which was much less 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 4,722 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). 72.7% of these phase IV trials conducted randomization 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 4,722 studies were small (median enrollment: 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.

Table 2 Characteristics of Included Trials by different Endpoint Classifications

	No (%)		
	All, 2004-2014 N=4722	Safety Alone N=330	Safety/ Efficacy N=4392
Overall status [a]			
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)
Enrollment, 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(.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			
Randomized	3310(70.1)	179(54.2)	3131(71.3)
Non-Randomized	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)
Masking			
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,%			
Female only	337(7.1)	19(5.8)	318(7.2)
Male 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 (<18y)	762(16.1)	72(21.8)	690(15.7)
Excluded elderly (>65 y)	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)
U.S.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 [b]			
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)
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)

Note:

a: "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.
b: Percentages may not sum to 100% , as categories are not mutually exclusive.

68.5% of the phase IV trials evaluating drug safety alone had enrollment less than 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 enrollment 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 small proportion using randomization (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 3 major therapeutic areas (cardiovascular, oncology and mental health). The cardiovascular diseases trials accounted for the most among these 3 categories (n=601, 12.7%). Also cardiovascular trials had more enrollment (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). Randomization 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 female-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 subjects. 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 3 Characteristics of Included Trials in Different Clinical Specialties

	No (%)		
	Cardiovascular diseases	Oncology	Mental Health
	N=601	N=251	N=594
Overall status [a]			
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)
Enrollment, 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	s		
Randomized	469(78.0)	101(40.2)	390(65.7)
Non-Randomized	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)
Masking			
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, %			

Female only	8(1.3)	34(13.5)	13(2.2)
Male 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 y)	33(5.5)	35(13.9)	112(18.9)
Excluded elder (>65 y)	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)
U.S.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 [b]			
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)
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)

Note:

a: "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.

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

Table 4 shows the results of the regression analyses. These analyses compared the trial characteristics that are related to the use of blinding and randomization. A total of 1,276 single-arm trials and 78 studies with any of the data elements missing were excluded from the regression analysis. Hence, there were 3,361 trials considered in the regression model. Of these trials, 1950(58.02%) studied were blind and 3234(96.22%) were randomized. Different clinical specialties could affect the use of blinding and randomization. Oncology trials were less likely to use both blinding (adjusted OR: 0.33; 95% CI: 0.18-0.63) and randomization (adjusted OR: 0.42; 95% CI:

0.28-0.63). Mental health trials were more likely to implement blinding (adjusted OR: 3.35; 95% CI: 2.56-4.38). 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: 1.08-1.60).

Table 4 Regression Analyses of Included Trials and the Reported Use of Blinding and Randomization

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

Abbreviations: NIH, National Institutes of Health; OR, odds ratio.

a: 1950(58.02%) of the included 3,361 studied were blind in this regression model.

b: 3234(96.22%) of the included 3,361 studied were randomized in this regression model.

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 endpoint classifications and clinical specialties were compared. We also analyzed 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 only included phase IV trials with "safety" as an endpoint, and most of these trials (77.6%) had an enrollment of less than 300. In the phase IV trials with safety as the primary endpoint, the average sample size was only 104. Thus, these small trials might not have sufficient power to detect adverse events (AEs), especially less common adverse events.[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 adverse event. For example, to observe an AE with an occurrence probability of 1.5‰, the China Food and Drug Administration (CFDA) requires that the enrollment 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-randomized or open-label studies are accepted. If randomization and blinding are feasible in the studies with control 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 randomization and blinding, which was consisted with the results of previous studies.[15] One possible reason is that some of the

oncology trials are conducted to investigate individualized or personalized treatment and randomization or blinding is not feasible. Due 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, researcher should adopt randomization 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. 30.5% of the phase IV trials including the Asia and Pacific area was 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 research [15] revealed that 3.3% of all interventional clinical trials registered from October 2007 through September 2010 were terminated or withdrawn. We further analyzed the conditions, endpoints 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 characterization 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 endpoint 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 endpoint 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 emphasized for phase IV trials with safety as the primary endpoint.

Figure Legends: Figure 1 Flow Chart of Inclusion and Exclusion

FUNDINGS STATEMENT

This study was sponsored by the National Nature Science Foundation of China (NO. 81502895, 81373105), a grant from the key discipline for construction of evidence-based public health in Shanghai (NO. 12GWZX0602) and the Fourth Round of Three-year Action Plan on Public Health Discipline and Talent Program: Evidence-based Public Health and Health Economics (No. 15GWZK0901)

COMPETING INTERESTS

The authors declare that they have no competing interests.

CONTRIBUTORSHIP STATEMENT

Xinji Zhang and Yuan Zhang contributed equally in conceiving this project, facilitating protocol, analyzing data and drafting this manuscript. Yexiao Fei leaded the development of performance-based incentives and revised the manuscript critically. Tianyi Zhang and XiaoJing Guo gave their time and effort to modify the programs. Jia He provided expertise for the overall design of the study, and revised and approved the manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge the valuable advices on revision from the reviewers. .We also thank Jian Lu, PhD, for his assistance in designing the study. Besides, we acknowledged Meijing Wu and American Journal Experts, LLC for its professional copyediting service.

DATA SHARING STATEMENT

The analyzed dataset was upgraded on the Datadryad.org website .The title of the dataset used in this revision is “phase IV clinical studies received by ClinicalTrials.gov between 2004-2014”

URL: http://datadryad.org/review?doi=doi:10.5061/dryad.3t6sc

REFERENCE

1. Harpaz R, Dumouchel W, Shah NH, *et al*. Novel data-mining methodologies for adverse drug event discovery and analysis. Clinical Pharmacology & Therapeutics 2012;91(6):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. Clinical Pharmacology & Therapeutics 1995;58(1):108-17.
3. Englev E, Petersen K. ICH-GCP Guideline: quality assurance of clinical trials. Status and perspectives. Ugeskrift for laeger 2003;165(16):1659-62.
4. Dixon JR. The international conference on harmonization good clinical practice guideline. Quality Assurance: Good Practice, Regulation, and Law 1999;6(2):65-74.
5. Gale EAM. Post-marketing studies of new insulins: sales or science?. British Medical Journal 2012;344.
6. Lasser KE, Allen PD, Woolhandler SJ, *et al*. Timing of new black box warnings and withdrawals for prescription medications. Jama-Journal of the American Medical Association

2002;287(17):2215-20.
7. Food and Drug Administration. Food and Drug Administration Amendments Act (FDAAA) of 2007.http://www.fda.gov/downloads/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FoodandDrugAdministrationAmendmentsActof2007/FDAAAImplementationChart/UCM213016.pdf (accessed August 4, 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. The Journal of Clinical Pharmacology 2007;47(9):1074-86.
9. Schmidt LG, Grohmann R, Helmchen H, *et al*. Adverse drug-reactions - an epidemiological study at psychiatric hospitals. Acta Psychiatrica Scandinavica 1984;70(1):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(4):275-77.
12. Zarin DA, Tse T. Moving towards transparency of clinical trials. Science (New York, NY) 2008;319(5868):1340.
13. Cihoric N, Tsikkinis A, van Rhoon G, *et al*. Hyperthermia-related clinical trials on cancer treatment within the ClinicalTrials.gov registry. International Journal of Hyperthermia 2015;31(6):609-14. .
14. Shields KE, Lyerly AD. Exclusion of pregnant women from industry-sponsored clinical trials. Obstetrics and gynecology 2013;122(5):1077-81.
15. Califf RM, Zarin DA, Kramer JM, *et al*. Characteristics of Clinical Trials Registered in ClinicalTrials.gov, 2007-2010. Jama-Journal of the American Medical Association 2012;307(17):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(3).
17. Zarin DA, Tse T, Williams RJ, *et al*. The ClinicalTrials.gov Results Database - Update and Key Issues. New England Journal of Medicine 2011;364(9):852-60.
18. Hill KD, Chiswell K, Califf RM, *et al*. Characteristics of pediatric cardiovascular clinical trials registered on ClinicalTrials.gov. American Heart Journal 2014;167(6):921-U195.
19. Suvarna V. Phase IV of Drug Development. Perspectives in clinical research 2010;1(2):57-60.
20. China Food and Drug Administration(CFDA).Drug Registration Management Measures(2007). http://www.sfda.gov.cn/WS01/CL0053/24529.html. (accessed August 4, 2015) [in Chinese].

Figure 1 Flow Chart of Inclusion and Exclusion

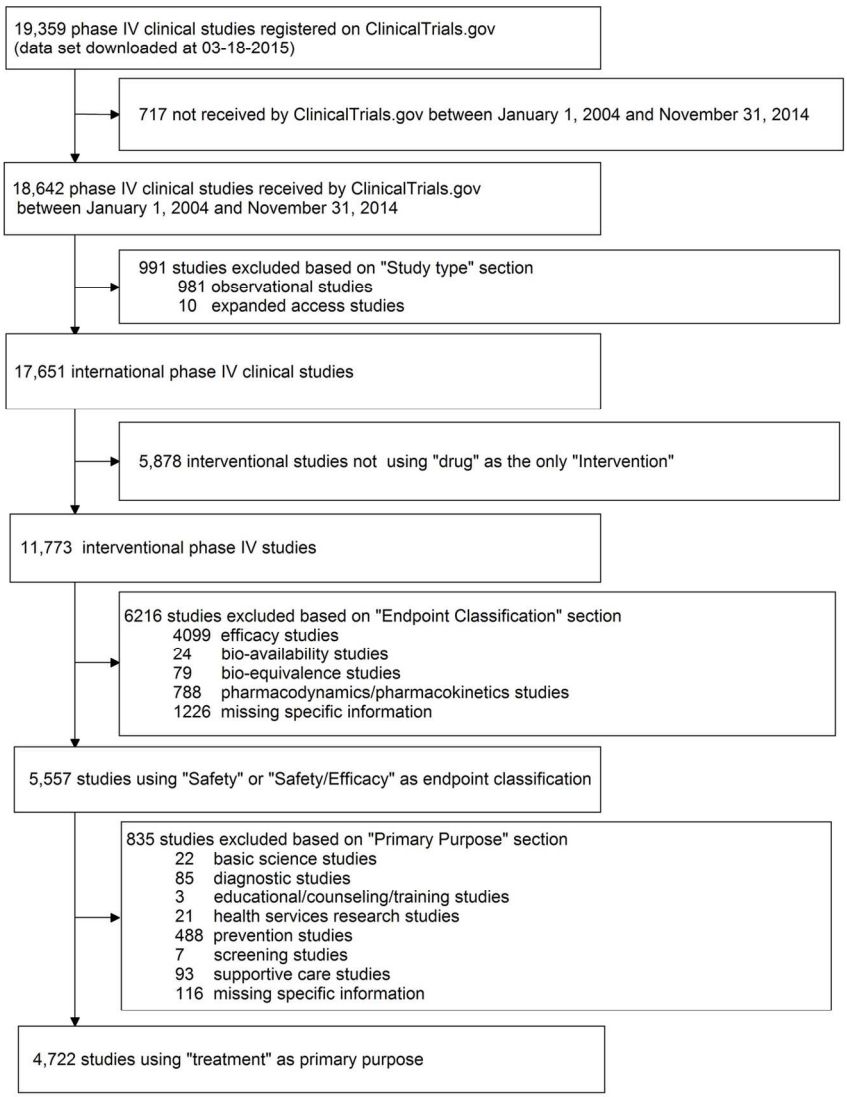


Figure 1 Flow Chart of Inclusion and Exclusion
Figure 1
146x207mm (300 x 300 DPI)