ORIGINAL RESEARCH

Portfolio of Clinical Research in Adult Cardiovascular Disease as Reflected in ClinicalTrials.gov

Karen P. Alexander, MD; David F. Kong, MD; Aijing Z. Starr, MS; Judith Kramer, MD; Karen Chiswell, PhD; Asba Tasneem, PhD; Robert M. Califf, MD

***Background*—**Cardiovascular medicine is widely regarded as a vanguard for evidence-based drug and technology development. Our goal was to describe the cardiovascular clinical research portfolio from ClinicalTrials.gov.

***Methods and Results*—**We identified 40 970 clinical research studies registered between 2007 and 2010 in which patients received diagnostic, therapeutic, or other interventions per protocol. By annotating 18 491 descriptors from the National Library of Medicine's Medical Subject Heading thesaurus and 1220 free-text terms to select those relevant to cardiovascular disease, we identified studies that related to the diagnosis, treatment, or prevention of diseases of the heart and peripheral arteries in adults (n=2325 [66%] included from review of 3503 potential studies). The study intervention involved a drug in 44.6%, a device or procedure in 39.3%, behavioral intervention in 8.1%, and biological or genetic interventions in 3.0% of the trials. More than half of the trials were postmarket approval (phase 4, 25.6%) or not part of drug development (no phase, 34.5%). Nearly half of all studies (46.3%) anticipated enrolling 100 patients or fewer. The majority of studies assessed biomarkers or surrogate outcomes, with just 31.8% reporting a clinical event as a primary outcome.

***Conclusions*—**Cardiovascular studies registered on ClinicalTrials.gov span a range of study designs. Data have limited verification or standardization and require manual processes to describe and categorize studies. The preponderance of small and late-phase studies raises questions regarding the strength of evidence likely to be generated by the current portfolio and the potential efficiency to be gained by more research consolidation. **(*J Am Heart Assoc.* 2013;2:e000009 doi: 10.1161/JAHA.113.000009)**

Key Words: cardiovascular diseases • cardiovascular medicine • clinical research • clinical trials

In 1997, the Food and Drug Administration Modernization Act (Public Law 105-115) mandated the creation of a public information resource to track clinical trial research. Hosted by the National Library of Medicine (NLM), ClinicalTrials.gov is implemented as a self-reported registry of clinical trial research entered via a Web-based system.[1–3] The registry was originally established to provide specific trial information to help patients identify ongoing studies for their disease or condition. Calls for increased transparency and public access to clinical trials information prompted expansion of the scope and purpose of the registry to accommodate the needs of research regulators, sponsors, reviewers, and policymakers. The Food and Drug Administration Amendments Act of 2007 (FDAAA; Public Law 110 to 85) attempted to reinforce standard reporting of clinical trials in the registry and mandated reporting of basic clinical trial results. Under the FDAAA, interventional studies of drugs, devices, and biologics that are either subject to oversight by the US Food and Drug Administration (FDA) or that include at least 1 US site are required to register in ClinicalTrials.gov. Registration of observational studies, or phase-1 interventional studies of drugs/biologics and small device feasibility studies, is encouraged but not required by law.

Although nobly conceived, the flexible and ambitious ClinicalTrials.gov data model has unstructured "free-text" data, a loosely controlled vocabulary, and no central monitoring or curation. These limitations are amplified by the multitude of investigators who are required to report their own trial data but who are provided with little guidance or oversight for data entry. Standardized data definitions are encouraged but not uniformly enforced. Quality assurance efforts from NLM monitor internal consistency and data

From the Duke Clinical Research Institute (K.P.A., D.F.K., A.Z.S., J.K., K.C., A.T., R.M.C.), and the Duke Translational Medicine Institute (J.K., R.M.C.), Durham, NC.

Correspondence to: Karen P. Alexander, MD, FACC, Duke University, Duke Box 3411, 2400 Pratt St, Suite 7068, Durham, NC 27705. E-mail: karen.alexander@duke.edu

Received November 29, 2012; accepted September 6, 2013.

formats, but are limited by lack of access to study protocols or actual study data.

The Duke Clinical Research Institute and the Clinical Trials Transformation Initiative collaborated with NLM to extract from ClinicalTrials.gov a derivative data set suitable for aggregate analysis.[4,5] This analysis data set describes the portfolio of recent and ongoing clinical research and summarizes study organization and oversight that reflects the quality of the evidence available to inform clinical guidelines and healthcare policy.[6]

As cardiovascular medicine is widely regarded as a vanguard for evidence-based drug and technology development, we systematically surveyed the aggregate portfolio of adult cardiovascular clinical trial research in ClinicalTrials.gov. In so doing, we identified clinical studies relevant to cardiovascular disease from the ClinicalTrials.gov analysis data set and assessed the frequency of design features desirable for generating quality evidence, such as controls, randomization, and blinding.[4]

Methods

ClinicalTrials.gov

ClinicalTrials.gov is the largest existing data repository of clinical trials, containing over 100 000 studies conducted in more than 170 countries. The methods used by ClinicalTrials.gov to register clinical studies have been previously described.[2,3,7] An electronic open access database hosted by NLM collects information for each clinical trial, including the disease or clinical condition being studied, keywords relevant to the focus of the investigation, and study design details. Clinical conditions may be described using the NLM's Medical Subject Heading (MeSH)-controlled vocabulary or as free text. NLM also maintains an algorithm that maps MeSH terms to registered studies by parsing the clinical conditions, keywords, and interventions entered for each trial. The algorithm assigns synonymous terms in the MeSH thesaurus and Supplementary Concept Records to applicable studies. A particular trial may be mapped to several MeSH terms based on the clinical condition(s) being studied. On the other hand, a study may not map to any MeSH terms. Examples include studies with free-text clinical conditions using commonly used terms (eg, cancer) or trials that do not include a reference to a specific disease (eg, healthy volunteer studies).

Under the Clinical Trials Transformation Initiative, a study data set comprising 96 346 clinical studies was downloaded from ClinicalTrials.gov on September 27, 2010. The date of download coincided with the anniversary of the enactment of the FDAAA 3 years earlier. Data were housed in the Oracle relational database management system (RDBMS), version 11.1g (Oracle Corporation). The data set was further limited to interventional studies registered after September 2007 (n=40 970). We chose registration date rather than start date or completion date to determine inclusion, as this field is generated by NLM and has no missing data, whereas some studies have missing or unreliable start date and completion date. An interventional study is defined by ClinicalTrials.gov as one that assigns research subjects to receive specific diagnostic, therapeutic, or other types of interventions based on a protocol, with subsequent outcome assessments.

Cardiovascular Study Selection

The ClinicalTrials.gov data encompass a large number of clinical specialties. To subgroup studies into clinical specialties, 18 491 MeSH condition terms from the 2010 MeSH thesaurus and 1220 frequently occurring free-text condition terms from interventional studies registered between 2007 and 2010 were reviewed and annotated by cardiovascular domain experts. Studies were considered for analysis if they had either a free-text condition term or a MeSH term tagged by domain experts as relevant to cardiovascular diseases (n=3503) (Figure 1). Out of these 3503 studies, 2600 studies had at least 1 free-text condition term that was relevant to cardiovascular disease, while 903 studies did not have a cardiovascular free-text condition term. Of these 903 studies, 199 studies had a relevant cardiovascular MeSH condition term.

Individual study records were reviewed to apply additional exclusion criteria. We excluded studies evaluating conditions related to venous and pulmonary embolic disease; general risk factors like diabetes, smoking, or hypertension in patients without coexisting cardiovascular disease; and other noncardiovascular populations or conditions. Also excluded were healthy volunteer studies that did not have a cardiovascular focus. This excluded 256 studies that had cardiovascular clinical conditions and 165 studies without cardiovascular conditions but relevant MeSH mapping. Because we were examining adult cardiovascular studies, 53 studies that enrolled only patients aged <18 years were excluded. This left a final population of 2325 clinical studies that evaluated conditions related to the diagnosis, treatment, or prevention of diseases of the heart and peripheral arteries in adults.

Statistical Analysis

Study characteristics were evaluated for all cardiovascular studies. In addition, studies were evaluated by subgroups within cardiology, by study phase, and by enrollment/completion status. Studies with missing values were excluded from summary statistics involving that data element.

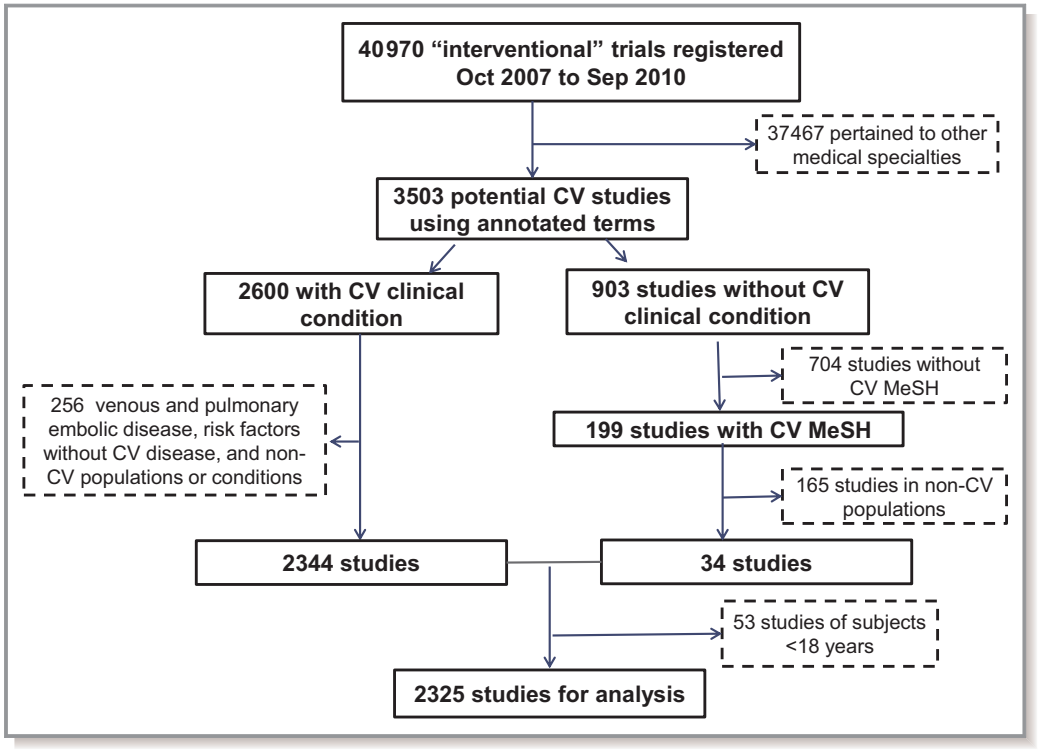

Figure 1. Identification of cardiovascular studies for analysis. CV indicates cardiovascular; MeSH, Medical Subject Heading.

Categorization of Cardiovascular Studies

There were 537 cardiovascular clinical conditions in the raw text data from NLM, which we refined to 179 semantically distinct terms by collapsing permutations in syntax, usage, or spelling (eg, acute myocardial infarction=myocardial infarction, myocardial infarctions, MI, acute MI, NSTEMI, or STEMI). These collapsed terms were used to divide the 2325 cardiovascular studies into 7 categories: (1) coronary artery disease (CAD)/angina, (2) heart failure/cardiomyopathy, (3) electrophysiology (EP)/arrhythmia, (4) valvular disease/surgery/transplant, (5) congenital heart disease (CHD), (6) peripheral arterial disease/aneurysm, and (7) prevention/imaging/other. Studies with only 1 condition term which prevented direct categorization (eg, heart disease or cardiovascular disease; n=96) were manually reclassified on the basis of the complete study record. Because studies could have multiple conditions, all listed conditions and the complete record were reviewed when necessary to break ties. Each study was assigned to only 1 category.

Classification of Sponsor, Intervention Type, and Outcomes

Free-text names of study sponsors were classified into the following groups: academic (US, non-US), government (US/National Institutes of Health [NIH], non-US), industry/business, collaborative group/institute/foundation, health system or provider consortium, and other. The interventions employed by each study were classified using a hierarchy. A trial would be classified as a drug study if at least 1 intervention used a drug. Studies without a drug arm were categorized as a device/procedure/radiation study if at least 1 arm involved a device or procedure. Studies without drug, device, or procedure arms were classified as biological/genetic if at least 1 arm involved a biological or genetic intervention. Dietary/behavioral studies were similarly identified from the remainder, with a classification of other available for studies that examined none of the above. Oversight authorities were classified as US FDA, US non-FDA, and non-US. Outcomes of completed studies were reviewed manually and classified as clinical events, biomarkers, quality of life, or economics, with more than 1 type of outcome possible per study.

Results

Of the 2325 cardiovascular studies, 74.0% (n=1722) were ongoing, 22% (n=511) were completed, and 4.0% (n=92) were terminated, withdrawn, or suspended. This sample of ongoing and completed studies had start years between 2001 and 2010 and anticipated end years between 2003 and 2015 (Figure 2). Table 1 describes the study design characteristics in the analysis sample. The majority of studies evaluated at

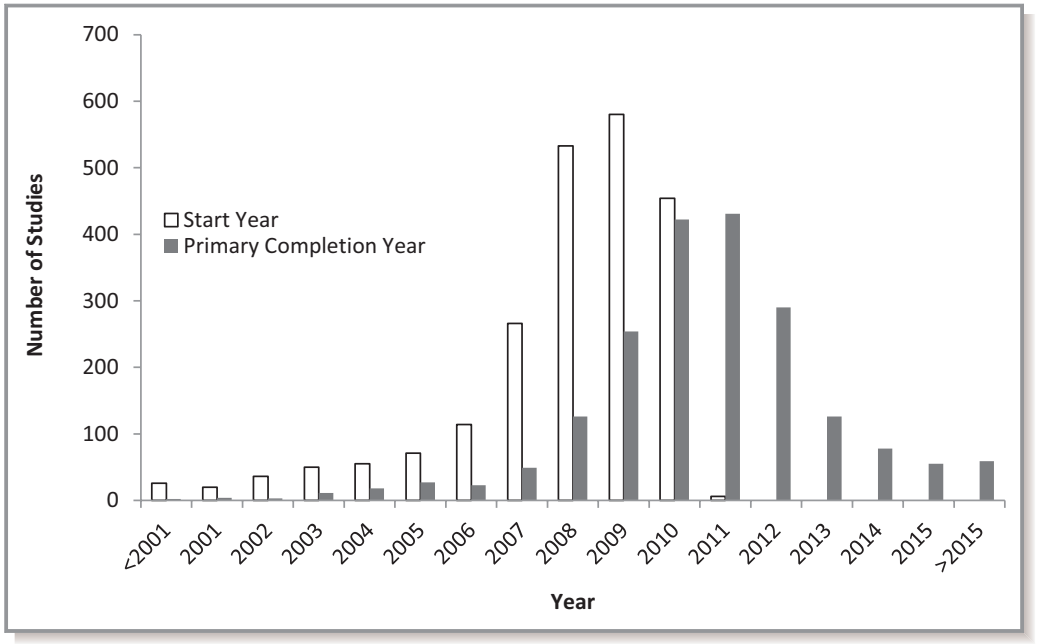

Figure 2. Study start year and primary completion year for 2233 ongoing or completed studies. Out of the 2233 ongoing and completed studies in the analysis data set, 2211 studies have data on start year, and 1978 studies have data on primary complete year.

least 1 drug (44.6%) or a device, procedure, or radiation (39.3%). Far fewer studies evaluated biological, genetic, dietary, or behavioral interventions. Most studies employed a parallel design (67.1%), had 2 experimental arms (59.9%), and used randomization (74.6%). Most studies with 2 or more arms were randomized (94.7%). The majority of studies were open-label (52.9%), with double-blinding present in 32.0% and single-blinding present in 15.1%. A single oversight authority was listed in 85.1%, US oversight was listed in 40.0%, and FDA oversight was listed in 47.9% of studies with US oversight. A little over half (53.3%) had a data monitoring committee.

Development phase was reported as phase 4 in 25.6% (n=596), phase 3 in 19.1% (n=444), phase 2 in 15.9% (n=370), and phase 0 or 1 in 4.9% (n=113); phase was listed as not applicable for the remaining 34.5% (n=802) of studies (Table 2). The vast majority of studies evaluated treatment or supportive care (73.6%), with fewer focusing on prevention, diagnosis, or screening (21. 2%). The median enrollment was 87 (interquartile range [IQR]: 37 to 225) for completed trials, and 120 (50 to 312) for trials that had been registered but not completed. Half the studies overall (46.3%) anticipated enrolling no more than 100 patients; this was more common in phase-0 to 2 studies (68.9%) and least common in phase-3 studies (29.3%) (Table 2). Anticipated enrollment was >1000 patients for 16.8% of phase-3 and 12.5% of phase-4 studies. Studies generally enrolled both sexes. There was a maximum age exclusion for patients >75 years of age in 17.5% of studies, more commonly in phases 0 to 2 (20.7%) than in phase 3 (14.0%).

Industry was the most frequent lead sponsor overall (32.0%) and across study phases (Table 3). Phase-0 to 2 studies were most likely to be funded by industry (50.9%). Academic institutions were sponsors or collaborators in ≈40% of studies, with non-US institutions (24.3%) the leads on more studies than US institutions (15.5%). Government entities were sponsors or collaborators in 5.0% of studies. The remaining studies listed research institutes, foundations, or clinical practice networks as the lead collaborators (Table 3). There was 1 lead collaborator in 64.0%, 2 lead collaborators in 25.2%, and 3 or more in 10.8% of studies; these percentages varied little across study phases (Table 3). The most frequent regions of enrollment across study phases were North America and Europe. There was a higher proportion of phase-3 studies with sites in Central and South America, North Asia, Pacifica, and the Middle East. For phase 4, the proportion of studies with sites in East and Southeast Asia was higher, and the proportion of studies with sites in North America was lower (Table 3).

Cardiovascular conditions under study were CAD/angina in 42.5% (n=987), heart failure/cardiomyopathy in 21.6% (n=502), EP/arrhythmia in 13.1% (n=304), valvular disease/surgery/transplant in 7.6% (n=176), CHD in 1.1% (n=27), peripheral arterial disease or aneurysm in 5.5% (n=129), and prevention/imaging/other in 8.6% (n=200) (Table 4). Distribution of phase of study varied across subtype. Specifically, CAD/angina and EP/arrhythmia had the most phase-4 studies, heart failure/cardiomyopathy and peripheral arterial disease/aneurysm had the most early-phase studies (phase-0 to 2 studies), and CHD and valvular disease/surgery/transplant studies were the subtypes with the most studies not in phases 0–4 (Table 4). Study intervention was most often used

Table 1. Study Design Characteristics by Study Status (n=2325)

Characteristic	All (n=2325)	Ongoing* (n=1722)	Completed (n=511)	Other† (n=92)
Intervention Type				
Drug	44.6 (1036)	40.3 (694)	55.8 (285)	62.0 (57)
Device, procedure, or radiation	39.3 (913)	42.9 (738)	28.6 (146)	31.5 (29)
Dietary supplement/behavioral	8.1 (188)	8.1 (140)	8.8 (45)	3.3 (3)
Biological or genetic	3.0 (69)	3.4 (58)	1.8 (9)	2.2 (2)
Other	5.1 (119)	5.3 (92)	5.1 (26)	1.1 (1)
Primary Purpose				
Treatment/supportive care	73.6 (1653)	74.2 (1233)	70.4 (349)	79.8 (71)
Prevention	12.2 (273)	11.6 (192)	15.3 (76)	5.6 (5)
Diagnostic/screening	9.0 (203)	9.3 (154)	7.9 (39)	11.2 (10)
Health services research	2.5 (57)	2.8 (46)	2.2 (11)	0
Basic science	2.7 (60)	2.2 (36)	4.2 (21)	3.4 (3)
Interventional Model				
Single group	25.8 (594)	26.3 (450)	25.2 (127)	19.5 (17)
Parallel	67.1 (1545)	66.9 (1146)	66.2 (333)	75.9 (66)
Crossover	5.3 (123)	4.7 (81)	7.6 (38)	4.6 (4)
Factorial	1.7 (40)	2.0 (35)	1.0 (5)	0
No. of Arms				
1	26.8 (611)	27.7 (474)	25.5 (122)	16.9 (15)
2	59.9 (1363)	60.3 (1031)	56.7 (271)	68.5 (61)
3	8.3 (188)	8.0 (137)	9.6 (46)	5.6 (5)
≥4	5.0 (114)	3.9 (67)	8.2 (39)	9.0 (8)
Masking				
Open label	52.9 (1217)	54.8 (938)	48.3 (240)	43.3 (39)
Single-blind	15.1 (346)	15.7 (268)	14.5 (72)	6.7 (6)
Double-blind	32.0 (736)	29.6 (506)	37.2 (185)	50.0 (45)
Allocation				
Randomized	74.6 (1707)	73.9 (1256)	76.1 (379)	80.0 (72)
Nonrandomized	25.4 (581)	26.1 (444)	23.9 (119)	20.0 (18)
Primary Endpoint Classification				
Efficacy/safety	49.0 (943)	49.6 (706)	45.4 (194)	57.3 (43)
Efficacy	40.2 (774)	41.6 (592)	36.8 (157)	33.3 (25)
Safety	7.3 (141)	6.0 (86)	12.4 (53)	2.7 (2)
Bioequivalence	0.2 (3)	0.2 (3)	0	0
Pharmacokinetics/dynamics	3.3 (63)	2.5 (35)	5.4 (23)	6.7 (5)
Oversight Authorities				
US (vs non-US)	40.0 (931)	40.9 (704)	33.5 (171)	60.9 (56)
FDA (vs non-FDA), among US	47.9 (446)	44.9 (316)	56.7 (97)	58.9 (33)
No. of Oversight Authorities				
1	85.1 (1979)	84.6 (1457)	86.7 (443)	85.9 (79)
2	7.6 (176)	8.2 (141)	5.9 (30)	5.4 (5)
≥3	7.3 (170)	7.2 (124)	7.5 (38)	8.7 (8)
Study has data monitoring committee	53.3 (1134)	54.7 (869)	47.1 (212)	60.9 (53)

Data are presented as % (n). Missing data elements: primary purpose (n=79, 3.4%); interventional model (n=23, 1%); number of arms (n=49, 2.1%); masking (n=26, 1.1%); allocation (n=37, 1.6%); endpoint classification (n=401, 17.2%); data monitoring committee (n=198, 8.5%). FDA indicates US Food and Drug Administration.

*Ongoing indicates not yet recruiting, recruiting, enrolling by invitation, and active not recruiting.

†Other indicates terminated, withdrawn, or suspended enrollment.

Table 2. Enrollment Details by Study Phase (n=2325)

Characteristic	Overall	Phase 0 to 2 (n=483)	Phase 3* (n=444)	Phase 4 (n=596)	Other† (n=802)
Actual enrollment, median (IQR)	87 (37 to 225)	45 (26 to 116)	140 (50 to 424)	95 (44 to 280)	100 (40 to 240)
Patients, % (n)					
<50	35.2 (180)	52.9 (72)	28.4 (25)	29.4 (37)	28.6 (46)
51 to 100	19.6 (100)	17.6 (24)	13.6 (12)	22.2 (28)	22.4 (36)
101 to 500	32.7 (167)	24.3 (33)	37.5 (33)	33.3 (42)	36.6 (59)
501 to 1000	6.3 (32)	2.9 (4)	14.8 (13)	6.3 (8)	4.3 (7)
>1000	6.3 (32)	2.2 (3)	5.7 (3)	8.8 (6)	8.1 (13)
Anticipated enrollment, median (IQR)	120 (50 to 312)	60 (30 to 145)	200 (100 to 562)	182 (70 to 460)	104 (50 to 300)
Patients, % (n)					
<50	25.8 (458)	47.1 (162)	12.5 (44)	16.8 (77)	28.0 (175)
51 to 100	20.5 (364)	21.8 (75)	16.8 (59)	20.6 (94)	21.8 (136)
101 to 500	37.6 (668)	25.3 (87)	44.9 (158)	42.0 (192)	37.0 (231)
501 to 1000	6.8 (120)	3.2 (11)	9.1 (32)	8.1 (37)	6.4 (40)
>1000	9.4 (167)	2.6 (9)	16.8 (59)	12.5 (57)	6.8 (42)
Sex, % (n)					
Both male and female	97.5 (2268)	97.3 (470)	98.2 (436)	96.6 (576)	98.0 (786)
Male only	1.3 (31)	1.4 (7)	1.1 (5)	2.2 (13)	0.7 (6)
Female only	1.1 (26)	1.2 (6)	0.7 (3)	1.2 (7)	1.2 (10)
Minimum Age (y), % (n)					
≤17	1.7 (41)	1.4 (7)	2.0 (9)	1.3 (8)	2.1 (17)
18 to 21	77.1 (1792)	80.5 (389)	80.6 (358)	78.7 (469)	71.8 (576)
22+	15.6 (363)	14.7 (71)	14.2 (63)	15.1 (90)	17.3 (139)
None listed	5.5 (129)	3.3 (16)	3.2 (14)	4.9 (29)	8.7 (70)
Maximum Age (y), % (n)					
≤69	4.1 (96)	4.3 (21)	2.3 (10)	4.0 (24)	5.0 (41)
70 to 79	13.9 (323)	17.2 (83)	12.8 (57)	14.4 (86)	12.1 (97)
80+	20.8 (484)	26.7 (129)	16.7 (74)	20.3 (121)	20.0 (160)
None listed	61.2 (1422)	51.8 (250)	68.2 (303)	61.2 (365)	62.8 (504)
Age Exclusions (y), % (n)					
>65	3.8 (89)	4.3 (21)	2.0 (9)	3.9 (23)	4.5 (36)
>75	17.5 (406)	20.7 (100)	14.0 (62)	18.3 (109)	16.8 (135)
Intervention Type, % (n)					
Drug	44.6 (1036)	52.6 (254)	55.4 (246)	57.9 (345)	23.8 (191)
Device, procedure, or radiation	39.3 (913)	31.1 (150)	33.1 (147)	38.4 (229)	48.3 (387)
Biological or genetic	3.0 (69)	9.1 (44)	4.1 (18)	0.3 (2)	0.6 (5)
Dietary supplement/behavioral	8.1 (188)	5.0 (24)	5.6 (25)	0.8 (5)	16.7 (134)
Other	5.1 (119)	2.3 (11)	1.8 (8)	2.5 (15)	10.6 (85)
Primary Purpose, % (n)					
Treatment/supportive care	73.6 (1653)	79.6 (375)	77.6 (340)	79.7 (458)	63.0 (480)
Prevention	12.2 (273)	11.0 (52)	14.8 (65)	11.1 (64)	12.1 (92)
Diagnostic/screening	9.0 (203)	5.5 (26)	5.9 (26)	7.4 (43)	14.2 (108)
Health services research	2.5 (57)	0.6 (3)	0.5 (2)	0.7 (4)	6.3 (48)
Basic science	2.7 (60)	3.2 (15)	1.1 (5)	1.0 (6)	4.5 (34)

Studies submitted their actual enrollment by September 27, 2010 (n=511 studies, minimum enrollment=1 patient, maximum enrollment=18 277 patients); studies submitted their anticipated enrollment by September 27, 2010 (n=1777, minimum enrollment 2 patients, maximum enrollment=50 000 patients). IQR indicates interquartile range.
*Includes combined phase-2/3 studies.
†Other indicates terminated, withdrawn, or suspended enrollment.

Table 3. Study Oversight and Location Details (n=2325)

Characteristic	Overall	Phase 0 to 2 (n=483)	Phase 3* (n=444)	Phase 4 (n=596)	Other† (n=802)
Lead Sponsor					
Government (US/NIH)	3.6 (83)	5.2 (25)	4.5 (20)	1.8 (11)	3.4 (27)
Government (non-US)	1.4 (32)	1.4 (6)	0.9 (4)	0.8 (5)	2.1 (17)
Institute, foundation, network	22.8 (530)	12.2 (59)	24.6 (109)	31.7 (189)	21.6 (173)
Industry	32.0 (745)	50.9 (246)	39.6 (176)	23.5 (140)	22.8 (183)
Academic (US)	15.5 (360)	14.5 (70)	9.2 (41)	10.7 (64)	23.1 (185)
Academic (non-US)	24.3 (564)	15.5 (75)	20.3 (90)	31.0 (185)	26.7 (214)
Other	0.5 (11)	0.4 (2)	0.9 (4)	0.3 (2)	0.4 (3)
No. of Collaborator/Sponsors					
1	64.0 (1488)	65.0 (314)	62.4 (277)	64.4 (384)	64.0 (513)
2	25.2 (586)	24.8 (120)	25.0 (111)	24.0 (143)	26.4 (212)
≥3	10.8 (251)	10.1 (49)	12.6 (56)	11.6 (69)	9.6 (77)
Region‡					
North America	44.5 (959)	53.4 (241)	47.2 (188)	29.7 (165)	48.6 (365)
Central/South America	5.2 (113)	4.7 (21)	13.1 (52)	4.5 (25)	2.0 (15)
Europe	42.6 (919)	38.6 (174)	49.2 (196)	46.3 (257)	38.9 (292)
East/Southeast Asia	12.8 (276)	9.5 (43)	11.1 (44)	20.5 (114)	10.0 (75)
North Asia	2.9 (62)	4.9 (22)	7.8 (31)	1.1 (6)	0.4 (3)
South Asia	2.7 (59)	3.3 (15)	5.8 (23)	1.4 (8)	1.7 (13)
Pacifica	3.6 (77)	3.8 (17)	10.1 (40)	1.8 (10)	1.3 (10)
Middle East	4.5 (97)	5.1 (23)	9.3 (37)	3.4 (19)	2.4 (18)
Africa	1.5 (32)	1.6 (7)	5.5 (22)	0.4 (2)	0.1 (1)
No. of Regions					
1	85.2 (1980)	81.6 (394)	76.1 (338)	87.8 (523)	90.4 (725)
2	3.8 (89)	7.5 (36)	3.4 (15)	3.7 (22)	2.0 (16)
≥3	3.7 (86)	4.3 (21)	10.1 (45)	1.7 (10)	1.2 (10)
NA	7.3 (170)	6.6 (32)	10.4 (46)	6.9 (41)	6.4 (51)

Data are presented as % (n). Sponsor is primary organization that oversees implementation of study and is responsible for data analysis. Collaborator is an organization providing supervision, including funding, design, implementation, and reporting. NA indicates not applicable (region missing); NIH, National Institutes of Health.

*Includes combined phase-2/3 studies.

†Other=terminated, withdrawn, or suspended enrollment.

‡Top enrolling countries per region: East/Southeast Asia (China, Japan, Hong Kong, Korea, Taiwan, Vietnam, Singapore, Philippines), North Asia (Ukraine, Russian Federation, Belarus), South Asia (India, Pakistan, Bangladesh), North America (Canada, United States, Mexico), South/Central America (Puerto Rico, Guatemala, Costa Rica, Panama, Brazil, Argentina, Columbia), Middle East (Israel, Iran, Turkey), Pacifica (Australia, New Zealand, New Guinea), Africa (South Africa, Kenya, Uganda, Egypt); for a complete list of countries in each region, see www.clinicaltrials.gov.

for drug or device/procedure/radiation across the subtypes of cardiovascular disease (Table 4). However, EP/arrhythmia and peripheral arterial disease/aneurysm had the most device/procedure/radiation studies, and prevention/imaging/other had the most dietary and behavioral studies. Lead sponsors followed the overall distribution by cardiovascular subtype. Peripheral arterial disease/aneurysm and EP/arrhythmia had the most industry lead sponsors. Non-US academic centers were lead sponsors more often than US academic centers, particularly for valvular disease/surgery/transplant, CHD, and prevention/imaging/other studies (Table 4).

Primary outcomes for the completed studies (n=511) were most often studying a biomarker for a clinical condition or outcome (72.0%), and only 31.8% listed a clinical event for their primary outcome (Table 5). Quality-of-life and economic analyses were performed in 2.6% of studies.

Discussion

Our analysis of adult cardiovascular research registered in ClinicalTrials.gov demonstrates that prevalent diseases and diverse patient enrollment do not necessarily translate into a

Table 4. Study Details by Subtype (n=2325)

Variable	CAD/Angina (n=987)	HF/CMP (n=502)	EP/Arrhythmia (n=304)	Valvular Disease/ Surgery/Transplant (n=176)	Congenital Heart Disease (n=27)	PAD/Aneurysm (n=129)	Prevention/ Imaging/Other (n=200)
Study Status							
Ongoing	72.3 (714)	73.9 (371)	80.9 (246)	74.4 (131)	77.8 (21)	74.4 (96)	71.5 (143)
Completed	23.9 (236)	19.9 (100)	17.8 (54)	21.6 (38)	18.5 (5)	23.3 (30)	24.0 (48)
Other	3.7 (37)	6.2 (31)	1.3 (4)	4.0 (7)	3.7 (1)	2.3 (3)	4.5 (9)
Intervention Type							
Drug	49.8 (492)	43.8 (220)	37.5 (114)	45.5 (80)	55.6 (15)	20.9 (27)	44.0 (88)
Device, procedure, or radiation	35.6 (351)	35.5 (178)	52.6 (160)	39.8 (70)	40.7 (11)	64.3 (83)	30.0 (60)
Biological, genetic	2.0 (20)	4.2 (21)	0.3 (1)	6.3 (11)	0	10.1 (13)	1.5 (3)
Dietary, behavioral	8.0 (79)	9.0 (45)	3.6 (11)	4.5 (8)	3.7 (1)	3.9 (5)	19.5 (39)
Other	4.6 (45)	7.6 (38)	5.9 (18)	4.0 (7)	0	0.8 (1)	5.0 (10)
Primary Purpose							
Treatment/supportive care	70.0 (671)	81.2 (397)	73.8 (214)	70.5 (117)	63.0 (17)	92.8 (116)	63.7 (121)
Prevention	11.4 (109)	6.3 (31)	18.6 (54)	20.5 (34)	18.5 (5)	2.4 (3)	19.5 (37)
Diagnostic/screening	12.1 (116)	7.2 (35)	6.2 (18)	5.4 (9)	14.8 (4)	3.2 (4)	8.9 (17)
Health services research	3.2 (31)	3.1 (15)	0.7 (2)	0.6 (1)	0	0.8 (1)	3.7 (7)
Basic science	3.3 (32)	2.2 (11)	0.7 (2)	3.0 (5)	3.7 (1)	0.8 (1)	4.2 (8)
Phase							
Phase 0 to 2	17.8 (176)	28.3 (142)	15.8 (48)	15.3 (27)	3.7 (1)	38.0 (49)	20.0 (40)
Phase 3*	20.3 (200)	17.3 (87)	20.4 (62)	19.3 (34)	14.8 (4)	9.3 (12)	22.5 (45)
Phase 4	31.2 (308)	19.9 (100)	27.6 (84)	19.9 (35)	22.2 (6)	18.6 (24)	19.5 (39)
Other	30.7 (303)	34.5 (173)	36.2 (110)	45.5 (80)	59.3 (16)	34.1 (44)	38.0 (76)
Lead Sponsor							
Government (US/NIH)	3.3 (33)	5.2 (26)	2.6 (8)	2.3 (4)	0	3.9 (5)	3.5 (7)
Government (non-US)	1.8 (18)	2.0 (10)	1.0 (3)	0.6 (1)	0	0	0
Institute, foundation, network	26.8 (265)	18.3 (92)	23.0 (70)	21.1 (37)	22.4 (6)	15.5 (20)	20.0 (40)
Industry	26.7 (264)	33.9 (170)	40.8 (124)	32.4 (57)	25.9 (7)	46.5 (60)	31.5 (63)
Academic (US)	14.0 (138)	18.9 (95)	12.8 (39)	13.6 (24)	25.9 (7)	14.0 (18)	19.5 (39)
Academic (non-US)	26.8 (265)	21.3 (107)	19.4 (59)	28.4 (50)	25.9 (7)	19.4 (25)	25.5 (51)
Other	0.4 (4)	0.4 (2)	0.3 (1)	1.7 (3)	0	0.8 (1)	0

Data are presented as % (n). CAD indicates coronary artery disease; EP, electrophysiology; HF/CMP, heart failure/cardiomyopathy; NIH, National Institutes of Health; PAD, peripheral arterial disease.
*Includes combined phase-2/3 studies.

portfolio of large controlled trials powered to test meaningful clinical endpoints. Despite over 6 million admissions for cardiovascular disease in the United States annually,[8] the cardiovascular research landscape is dominated by smaller trials and postmarketing studies. This may be the result of fragmented allocation of resources to research, interest in incubator and early technology development, or challenges in generating the large-scale research needed to demonstrate efficacy and safety in the general population. Although early development is crucial to innovation, investigators, sponsors, and policymakers must also avoid the pitfalls of bias and the overestimation of effects from combining multiple small trials as a strategy.[9] Lastly, the preponderance of smaller trials represents an opportunity to encourage collaboration and foster research networks capable of enrolling larger numbers of patients, thereby improving the likelihood that the clinical research portfolio will make a meaningful impact on the care of patients.

Representation of cardiovascular conditions within the adult cardiovascular portfolio was as expected, reflecting the highly

Table 5. Primary Outcomes Listed for Completed Studies (n=511)

Primary Outcome*	Overall (n=511)	Phase 0 to 2 (n=133)	Phase 3† (n=81)	Phase 4 (n=147)	Other‡ (n=150)
Clinical event	31.8 (157)	34.4 (45)	30.4 (24)	27.1 (38)	35.0 (50)
Biomarker	72.0 (355)	74.0 (97)	75.9 (60)	72.1 (101)	67.8 (97)
Quality of life	2.0 (10)	1.5 (2)	0	1.4 (2)	4.2 (6)
Economic analysis	0.6 (3)	0	0	1.4 (2)	0.7 (1)

Data are presented as % (n). Outcome missing in 3.5% (n=18) studies.
*Outcome categories describe clinical events (eg, MI, death, stroke), biomarkers (eg, blood pressure, lab or imaging findings), quality of life (eg, patient-reported outcomes), or economic evaluations (eg, cost or cost-effectiveness).
†Includes combined phase-2/3 studies.
‡Other indicates terminated, withdrawn, or suspended enrollment.

prevalent pathology of CAD, EP, and heart failure. Peripheral arterial disease, CHD, and preventive cardiology studies were less frequently studied. Early-phase (0 to 1) pipeline studies constituted just 5% of the adult cardiovascular portfolio, likely reflecting the legal exemption of phase-1 and device feasibility studies from registration. One-third of studies reported that development phase was not applicable, reflecting a substantial portion of registered research activity occurring outside typical development programs. The limited degree of government sponsorship (<5%) reported in ClinicalTrials.gov likely reflects the policy that NIH-funded studies register the institution as sponsor. Industry was the most frequent sponsor for studies of EP and peripheral vascular disease, while academia was the most frequent sponsor for studies of CAD, heart failure, and surgery. Additionally, the registry reflects substantial globalization of cardiovascular clinical research, with over half the trials in the registry recruiting subjects outside North America. The majority of trials sponsored by academia listed a non-US institution as the lead sponsor. Sixty percent of studies were subject to non-US oversight, with 14.9% of studies reporting oversight by multiple authorities. Among studies subject to US oversight, 48% reported oversight by the FDA. Regulatory oversight for non-US–sponsored studies was not collected. The substantial number of lead study sponsors from outside the United States signifies the use of ClinicalTrials.gov beyond its original mandate of helping patients identify trials in which they might want to participate. Study personnel entering data into ClinicalTrials.gov from around the globe may have limited familiarity with ClinicalTrials.gov definitions, or difficulty with clinical research terms. Definitions in the registry also limit clear attribution of funding and sponsors for cardiovascular research in ClinicalTrials.gov, particularly from government or multiple-sponsor studies.

The expected enrollment for registered cardiovascular trials was surprisingly small, given the need for studies with ample statistical power using conventional hypothesis testing to inform evidence-based practice. Among studies with an enrollment target entered, 46% reported anticipated enrollment of fewer than 100 patients. Early-phase (0 to 2) studies had the greatest proportion of very small studies, with 67.8% of trials enrolling fewer than 100 patients. Yet, the median anticipated enrollment for phase-3 studies was 200 patients. Similarly, more than one-third of phase-4 studies (37.4%) anticipated fewer than 100 subjects. The mega-trial designs considered the bulwark of cardiovascular evidence were decidedly uncommon, with <10% of all registered studies and only 17% of phase-3 studies anticipating a sample of more than 1000 patients. This suggests that a large number of patients are being enrolled in small-sample-size studies that may be underpowered to answer important clinical questions. This could reflect a tendency to register more studies, including small studies, due to the desire to ensure the future publication of results. This may also reflect the challenge of adequate research funding and organization, as well as competing priorities to advance individual academic advancement, which may incentivize individual projects over larger collaborative studies.

Only 30% of registered studies identified a hard clinical endpoint such as death or myocardial infarction as their primary outcome measure, with the vast majority of studies using a surrogate measure. The use of surrogate endpoints in small studies may play a key role in evidence generation, but surrogate endpoints may also be selected in studies with insufficient power to otherwise test clinical outcomes. Although surrogate measures may reduce the sample size and cost required to detect a measurable effect, their limitations are well recognized.[10–12] Their frequent use as a primary endpoint counters the common assumption that clinical endpoints are "extensively studied as part of the primary analysis of a trial large enough to collect useful clinical endpoint data."[10] The use of surrogates in the cardiovascular domain ignores the appeals in editorials over the past decade for large clinical trials assessing key clinical outcomes.[11,13,14] The low rate of quality-of-life assessment as an outcome in cardiovascular trials is also counter to strategic initiatives from the American Heart Association to improve the cardiovascular health of the population by 2020.[15] A systematic approach in the ClinicalTrials.gov data set to

collect study endpoints would help clarify the selection of clinical endpoints, surrogate markers, and patient-reported outcomes in gathering evidence.

The trials reported within ClinicalTrials.gov are heterogeneous, reflecting the efforts and encouragement of the NIH to voluntarily register a broad spectrum of research studies. Although free-text data entry allows submitters considerable flexibility for describing study attributes, the unstructured data and the paucity of indexing and metadata limit search algorithms, information retrieval, synthesis of the research evidence, and knowledge development.

Data inaccuracies were a limitation of our analysis. Without clarification, fields with seemingly inconsistent data are reported as entered. Also, lexical variants or misspelled free text could cause an uncommon cardiovascular condition to remain unrecognized by an automated search algorithm. Further, keyword strategies are confounded by submission of overly general terms. For example, a keyword or clinical condition of aneurysm might refer to a study of cerebral aneurysm, abdominal aortic aneurysm, both, or neither. General keywords such as heart disease and atherosclerosis identify studies potentially relevant to a cardiovascular domain but do not provide insight into the relevance to categories within cardiovascular disease. The registry also lacks structural features to facilitate aggregate descriptions of similar clinical studies. These shortcomings were addressed by simultaneous review of MeSH keywords and manual evaluation of the full text of the registry entries. Based on an internal validation sample of 1000 randomly selected studies, we expect that the chance of misclassifying a large number of cardiovascular studies in our analysis data set is remote.[4] However, wider implementation of controlled vocabularies, such as MeSH and the Clinical Data Interchange Standards Consortium, would improve the accuracy of registry queries. In the cardiovascular domain, increasing implementation of uniform endpoint definitions and data element definitions in ClinicalTrials.gov records would facilitate data pooling, meta-analysis, and systematic overviews. Clarification of terminology and data structure would improve the understanding of the portfolio.

A robust evidence base for developing guidelines and healthcare policy will benefit from increased consolidation and clearer classification of the cardiovascular research portfolio. Registration in ClinicalTrials.gov, while an important step forward, provides little assurance of study quality, scientific validity, or practical relevance. Lack of uniform data entry further limits its use for research classification or review. Increasingly constrained research resources should drive a greater emphasis on networks, collaboration, and improved data standards for interoperability. National registry infrastructures that coordinate data collection across projects (such as the Swedish Coronary Angiography and Angioplasty Register[16]) have been valuable for critical evaluation of cardiovascular therapeutics and for postmarket safety surveillance. A move toward common data elements and standard terminology will facilitate data sharing among researchers, regulators, and policymakers. This could also prevent redundancy and improve assessment of safety and effectiveness. By prioritizing important research ideas, coupled with a robust national information infrastructure that extends ClinicalTrials.gov, the cardiovascular clinical research community can more rapidly advance therapeutics from concept to evidence-based implementation.

Acknowledgments

We acknowledge the Clinical Trials Transformation Initiative and support from the National Library of Medicine ClinicalTrials.gov team, Deborah Zarin and Nick Ide.

Sources of Funding

Financial support provided by grant U19 FD003800 from the US Food and Drug Administration awarded to Duke University for the Clinical Trials Transformation Initiative (https://www.trialstransformation.org/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures

Dr Kong received a research grant from Abbott Laboratory; consulting fees from Allmed Healthcare, American College of Cardiology Foundation, Alpha Medica, Abbott Laboratory, Primahealth, Astra Zeneca, and King Pharmaceuticals; and has equity in Irvine Scientific. Dr Kramer received consulting fees from GlaxoSmithKline. Dr Califf received research grants from Amylin, Scios, Merck/Schering Plough, Novartis, and Bristol-Myers Squibb Foundation; consulting fees from Scios, Novartis, Bayer, Roche, Pfizer, and Bristol-Myers Squibb Foundation; and has equity in Nitrox LLC. The other authors have no disclosures.

References

1. Tse T, Williams RJ, Zarin DA. Reporting "basic results" in ClinicalTrials.gov. *Chest*. 2009;136:295–303.
2. Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DA. Issues in the registration of clinical trials. *JAMA*. 2007;297:2112–2120.
3. Gillen JE, Tse T, Ide NC, McCray AT. Design, implementation and management of a web-based data entry system for ClinicalTrials.gov. *Stud Health Technol Inform*. 2004;107:1466–1470.
4. Tasneem A, Aberle L, Ananth H, Chakraborty S, Chiswell K, McCourt BJ, Pietrobon R. The database for aggregate analysis of ClinicalTrials.gov (AACT) and subsequent regrouping by clinical specialty. *PLoS One*. 2012;7:e33677.
5. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007–2010. *JAMA*. 2012;307:1838–1847.

6. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC Jr. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA*. 2009;301:831–841.
7. Zarin DA, Tse T, Ide NC. Trial registration at ClinicalTrials.gov between May and October 2005. *N Engl J Med*. 2005;353:2779–2787.
8. Roger VL, Go AS, Lloyd-Jones DM, Adams RJ, Berry JD, Brown TM, Carnethon MR, Dai S, de Simone G, Ford ES, Fox CS, Fullerton HJ, Gillespie C, Greenlund KJ, Hailpern SM, Heit JA, Ho PM, Howard VJ, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Makuc DM, Marcus GM, Marelli A, Matchar DB, McDermott MM, Meigs JB, Moy CS, Mozaffarian D, Mussolino ME, Nichol G, Paynter NP, Rosamond WD, Sorlie PD, Stafford RS, Turan TN, Turner MB, Wong ND, Wylie-Rosett J; American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2011 update: a report from the American Heart Association. *Circulation*. 2011;123:e18–e209.
9. Fleming TR. Clinical trials: discerning hype from substance. *Ann Intern Med*. 2010;153:400–406.
10. Buhr KA. Surrogate end points in secondary analyses of cardiovascular trials. *Prog Cardiovasc Dis*. 2012;54:343–350.
11. Psaty BM, Weiss NS, Furberg CD, Koepsell TD, Siscovick DS, Rosendaal FR, Smith NL, Heckbert SR, Kaplan RC, Lin D, Fleming TR, Wagner EH. Surrogate end points, health outcomes, and the drug-approval process for the treatment of risk factors for cardiovascular disease. *JAMA*. 1999;282:786–790.
12. Wittes J, Lakatos E, Probstfield J. Surrogate endpoints in clinical trials: cardiovascular diseases. *Stat Med*. 1989;8:415–425.
13. Califf RM, Kramer JM. What have we learned from the calcium channel blocker controversy? *Circulation*. 1998;97:1529–1531.
14. Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA*. 1999;282:790–795.
15. Lloyd-Jones DM, Hong Y, Labarthe D, Mozaffarian D, Appel LJ, Van Horn L, Greenlund K, Daniels S, Nichol G, Tomaselli GF, Arnett DK, Fonarow GC, Ho PM, Lauer MS, Masoudi FA, Robertson RM, Roger V, Schwamm LH, Sorlie P, Yancy CW, Rosamond WD. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. *Circulation*. 2010;121:586–613.
16. Jernberg T, Attebring MF, Hambraeus K, Ivert T, James S, Jeppsson A, Lagerqvist B, Lindahl B, Stenestrand U, Wallentin L. The Swedish Web-system for enhancement and development of evidence-based care in heart disease evaluated according to recommended therapies (SWEDEHEART). *Heart*. 2010;96:1617–1621.