European Heart Journal - Quality of Care and Clinical Outcomes (2017) **3**, 114–122
doi:10.1093/ehjqcco/qcw058

**ORIGINAL ARTICLE**

# Evolving quality standards for large-scale registries: the GARFIELD-AF experience

**Keith A.A. Fox[1]*, Bernard J. Gersh[2], Sory Traore[3], A. John Camm[4], Gloria Kayani[3], Anders Krogh[3], Shweta Shweta[3], and Ajay K. Kakkar[3,5] for the GARFIELD-AF Investigators[†]**

[1]BHF Centre for Cardiovascular Science, University of Edinburgh, Queen's Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, UK; [2]Mayo Clinic College of Medicine, Rochester, 200 1st St SW, Rochester, MN 55905, USA; [3]Thrombosis Research Institute, Emmanuel Kaye Building, Manresa Road, Chelsea, London SW3 6LR, UK; [4]St. George's University of London, Department of Cardiology, St. George s Hospital, Crammer Terrace, London SW17 0RE, UK; and [5]University College London, Gower St, Kings Cross, London WC1E 6BT, UK

*Received 31 August 2016; revised 1 November 2016; editorial decision 3 November 2016; accepted 5 November 2016; online publish-ahead-of-print 15 November 2016*

**See page 99 for the editorial comment on this article (doi:10.1093/ehjqcco/qcx003)**

**Aims** Registries have the potential to capture treatment practices and outcomes in populations beyond the constraints of clinical trial settings. The value of data obtained depend critically upon robust quality standards (including source data verification [SDV] and training); features that are commonly absent from registries. This article outlines the quality standards developed for Global Anticoagulant Registry in the FIELD–Atrial Fibrillation (GARFIELD-AF).

**Methods and results** GARFIELD-AF comprises ∼57 000 patients prospectively recruited over 6.5 years in 35 countries in five successive cohorts. The registry employs a combination of remote and onsite monitoring to ascertain completeness and accuracy of records and by design, SDV is performed on 20% of cases (i.e. ∼11 400 patients). Four performance measures for ranking sites according to data quality and other performance indicators were evaluated (including data quality for 13 quantifiable variables, late data locking, number of missing critical variables, and history of poor data quality from the previous monitoring phase). These criteria facilitated the identification of sites with potentially suboptimal data quality for onsite monitoring. During early phases of the registry, critical variables for data checking were also identified. SDV using these variables (partial SDV in 902 patients) showed similar concordance to SDV of all fields (110 patients): 94.4% vs. 93.1%, respectively. This standard formed the baseline against which ongoing quality improvements were assessed, facilitating corrective action on data quality issues. In consequence, concordance was improved in the next monitoring phase (95.6%; $n = 1172$).

**Conclusion** The quality standards in GARFIELD-AF have the potential to inform a future 'reference' for registries.

**Keywords** Registries • Medical audit • Quality assessment • Atrial fibrillation • Anticoagulation

## Introduction

Atrial fibrillation (AF) is highly prevalent, especially in aging populations, and is currently estimated to affect approximately 5–6.1 million people in USA and 8.8 million people in Europe.[1–4] AF and its complications constitute a major public health burden and account for US$16–26 billion of the annual US health expenditure,[5] and at least 1% of the National Health Service budget in UK (US$2.3 billion).[6] To optimize management of the condition, robust multinational observational programs are needed to characterize patients with AF, their management and their outcomes.

Several large-scale national and international registries are under way, with the aim of defining the management and outcomes in broadly representative populations of patients with AF, including: GARFIELD-AF

* Corresponding author. Tel: +44 1479831771, Email: k.a.a.fox@ed.ac.uk

[†] A complete list of investigators is given in Supplementary material online, *Appendix*.

(clinical trial identifier: NCT01090362), GLORIA-AF (NCT01468701), ORBIT-AF (NCT01701817) and PREFER in AF.[7–11] These observational registries have the potential to capture the burden of disease in large-scale populations[12] by employing wide inclusion criteria and representation of historically under-researched groups, such as the elderly and patients with comorbidities.[13]

However, registries differ in their design, recruitment strategies, care settings, geographic representation and duration of follow-up. Only some registries collect data prospectively and employ systematic quality assurance methods to check the validity of data against sources, and use independent adjudication to ensure the robustness of outcome and safety measures. Thus, to allow robust interpretations, there is a compelling need for appropriate quality standards to be applied to registries.

Evidence generated from multisite randomized controlled trials (RCTs) is recognized as the 'gold standard' for comparing treatment options. Such trials employ robust methods for comparing treatment strategies, but they have restrictive inclusion/exclusion criteria. This constitutes a form of entry bias since eligibility for both forms of therapy is mandatory in order to ethically justify randomization.[14] Hence, trial patients do not necessarily reflect the full range of baseline characteristics, nor the frequency of outcome and safety events observed in unrestricted 'real-world' populations.

The study design also needs to be appropriate for the research questions. While the processes employed in RCTs may represent the appropriate quality standards and performance measures for comparing treatments for product registration, these measures are not designed for large-scale observational registries. Nonetheless quality standards and performance measures are required for registry programs, but these are not well defined, and can differ substantially from one program to the next.

The International Conference on Harmonization (ICH) Good Clinical Practice (GCP) guideline defines trial monitoring as 'the act of overseeing the progress of a clinical trial and of ensuring that it is conducted, recorded and reported in accordance with the protocol, Standard Operating Procedures, GCP, and the applicable regulatory requirement(s)'.[15] ICH GCP does not specify the methods for monitoring but suggests that 'in general there is a need for on-site monitoring, before, during and after the study'.

One of the most common procedures undertaken during onsite monitoring is source data verification (SDV), in order to check that data recorded within electronic case report form (eCRF) matches the primary source data. Such quality standards in registries are key to the veracity of the findings and their generalizability. However, the extent of SDV varies and some registries and observational programs avoid SDV completely, as it is considered too intrusive during the collection of real-world data. Post-marketing surveys, e.g. tend not to use SDV, and instead opt for risk-management approaches, relying entirely on the results of remote monitoring to check data reliability and to trigger onsite management and training.

Comprehensive (i.e. 100%) SDV of the whole record for all patients, as performed in RCTs designed for product registration, is impractical and beyond the financial scope of large-scale registries. More cost-effective alternatives to 100% SDV are needed for registries.[16–18] A few registries have adopted a combination of remote monitoring and risk-based onsite audits and SDV to ensure that data aligned to routine practice are correctly reported according to the study protocol.[19,20]

# Theoretical requirements for high-quality registries

## Recording of data and quality assurance

A criticism of registries is that they may lack the stringent quality assurance, as seen for key safety and efficacy endpoints in clinical trials. What constitutes 'acceptable' data quality for a given clinical trial or registry depends on multiple factors, including the variables themselves, the size of the dataset, the type and extent of errors, and the accuracy of the statistical analysis.[21] The quality of data is frequently assessed centrally using Kappa summary statistics or by dividing the number of errors observed by the number of data fields inspected. In RCTs, an error rate of 5% or less within electronic datasets[22] and outstanding queries on 1% or less of the data are generally considered acceptable standards.[16,23] Are the same standards achievable and appropriate for registries?

## Data management and remote monitoring

In order to mitigate the risks of reduced data quality, large studies are increasingly dependent on remote monitoring and quality assurance. Data discrepancies that are identified by remote adjudication can be queried in 'data management' processes that involve the application of screening rules and internal consistency checks. In large registries, achieving a balance between data integrity and ease of enrolment and follow-up is an important consideration during the planning of audits.

## Quality assurance protocols

Internal data quality assurance protocols are needed to assess completeness, consistency and accuracy.[24] However, registries vary in their quality assurance procedures. There are key differences between registries that employ routinely collected data (with variable clinical interpretations of endpoints and bleeding events) vs. those with predefined endpoints, which are adjudicated and audited for accuracy.[25] Large-scale epidemiological studies are valuable, but they do not collect all the variables needed in assessing treatments and outcomes in patients with AF and instead rely on routinely collected clinical data with neither standardized definitions of disease nor consistently defined outcomes.[26,27]

With the absence of a consensus, and only limited discussion in the literature of a reasonable and cost-effective approach for the audit of registry data,[28,29] the authors reflected on how they might achieve the quality standards and performance measures within large-scale registries such as the Global Anticoagulant Registry in the FIELD–Atrial Fibrillation (GARFIELD-AF). This article outlines the quality standards that were developed for the ongoing GARFIELD-AF registry, and derivation and validation of an electronic data quality score for study sites so that risk-based SDV could be implemented.

# Garfield-AF registry design

A detailed description of the design of the GARFIELD-AF registry has been reported previously.[7] In brief, GARFIELD-AF is an ongoing non-interventional registry of adults (≥18 years) with newly diagnosed non-valvular AF (diagnosis was established within 6 weeks of enrolment) and with one or more additional risk factor for stroke, as judged by the investigator, regardless of therapy. These risk factors were not prespecified in the protocol, nor were they limited to the components of existing risk stratification schemes. Prospective enrolment of consecutive patients meeting the inclusion criteria began in March 2010 in 19 countries worldwide. The roll-out of the GARFIELD-AF registry across five phases (cohorts) has now extended to 35 countries and more than 50 000 patients have been recruited, prospectively, over 6.5 years. The follow-up period will be a minimum of 2 years and a maximum of 8 years.

GARFIELD-AF is an independent academic research initiative sponsored by the Thrombosis Research Institute (London, UK) and supported by an unrestricted research grant from Bayer Pharma AG (Berlin, Germany). The quality assurance processes employed in the GARFIELD-AF registry are subject to independent review by an Audit Committee which, in turn, reports to the scientific Steering Committee. The statistical analyses are conducted by the Thrombosis Research Institute and independently reviewed by a leading statistician from a North American academic research centre (KP).

The primary aim of the registry is to define initial and ongoing management strategies and clinical and economic outcomes in patients with non-valvular AF in the clinical practice (non-trial) setting. The key outcomes for the registry are all-cause mortality, stroke or systemic embolism, major bleeding and healthcare utilization (including any hospitalization and emergency department visit). Data captured the management of AF in consecutive patients from the time of diagnosis to the end of follow-up. Patients are enrolled and data are captured on the eCRF by the physicians who are responsible for the patients' ongoing care. For example, if the cardiologists makes the initial diagnosis but the patient is then referred back to primary care, the primary care physician will enrol and follow-up the patient. A sufficient number of sites, both globally and on a national level were identified from hospital, community, and anticoagulation clinic settings to ensure proportional representation of AF treating care settings in all countries (office-based practice; hospital departments—neurology, cardiology, geriatrics, internal medicine, and emergency; anticoagulation clinics; and general or family practice).

# The audit process

## Central and onsite data monitoring

The registry employs a combination of remote electronic monitoring and more conventional onsite monitoring (including SDV) at ~10% of sites. The milestones for study recruitment, reporting and audit are outlined in *Figure 1*.

## Quality standards for the GARFIELD-AF registry

Quality standards for the GARFIELD-AF registry, as defined by the protocol, are summarized in *Table 1*. Sites are given access to online data entry only after formal training; all sites receive regular retraining depending on site performance and have ongoing access to a training web portal. In addition, 20% of all eCRFs (i.e. ~11 400 of 57 000 patients) are monitored against source documentation during six phases of audit between 2010 and 2018 (*Figure 1*). All modifications to the data are recorded electronically in an audit trail. At study completion, 5% of data for each of the critical variables for baseline data and follow-up (as defined in Supplementary material online, *Appendix Table S1*) are audited during the statistical analysis.

The eCRFs are monitored remotely to check for consistency, to identify implausible and outlying data, evaluate data quality and completeness, and to analyze patterns and trends. Monitoring and tracking of site-specific issues occurs either on a monthly, quarterly or 6-monthly basis, depending on the performance of each site. Site-level performance data, including patient recruitment numbers and rate of recruitment, are recorded by the clinical research associate (CRA) who also ensures query resolution and data locking (of baseline data and 4-monthly intervals thereafter); data are reviewed and audited at 12-monthly milestones. As outlined in *Figure 1*, the number of patients targeted for onsite monitoring during each phase of the audit is proportional to the number of patients recruited into the trial at the time of audit. The number of audited patients' records is cumulative over time so that by the end of the study, 20% or ~11 400 of patients will have been SDV'd (*Figure 1*).

## SDV of critical variables

Electronic data capture allows large volumes of data to be analyzed concurrently and to produce the summary statistics (e.g. missing data, data error rates, protocol violations) in real time in order to assess the magnitude of discrepancies between the electronic records and the site-verified source data.

The most efficient processes for onsite monitoring (in order to achieve 20% SDV of cases) evolved over several phases of monitoring. The process of onsite monitoring started early in the study, with an evaluation of 10 sites recruiting patients into cohort 1 (C1) in two countries (UK and France). Onsite monitors conducted 100% SDV of all fields in 15 patients (1–2 at each site). On an average, a complete SDV (i.e. 100% SDV) of all fields took 8–10 h to perform for each patient record. This first phase was both labour- and time-intensive and it led to revision of the monitoring strategy.

## Complete vs. partial SDV

During phase II monitoring, an abridged SDV process involving the assessment of only variables that were critical to the clinical dataset and statistical analyses was developed (see Supplementary material online, *Appendix Table S2*). Overall, 110 sites were monitored during phase II from 24 countries between November 2013 and April 2014. About 80% of the 110 sites were selected using performance-related criteria (data quality, GCP compliance issues, patient enrolment, outliers and other statistical anomalies) and ~20% of sites (which served as a control) were selected using random selection techniques.

In order to compare complete SDV with partial SDV, complete SDV was conducted in the first randomly selected patient at each site (i.e. 10% of patients) and compared with partial SDV of critical variables for nine patients at each site (i.e. ~90% of patients); SDV was only conducted on fields that were relevant to the patient or analysis at the time of study. Assessment of the results across all 110 sites showed that the level of concordance between the source data and the eCRF was similar following both complete SDV (93.1% of 7259 fields in 110 of 1012 patients) and partial SDV of critical variables (94.4% of 15 272 fields in 902 of 1012 patients), thus supporting the validity of partial SDV.

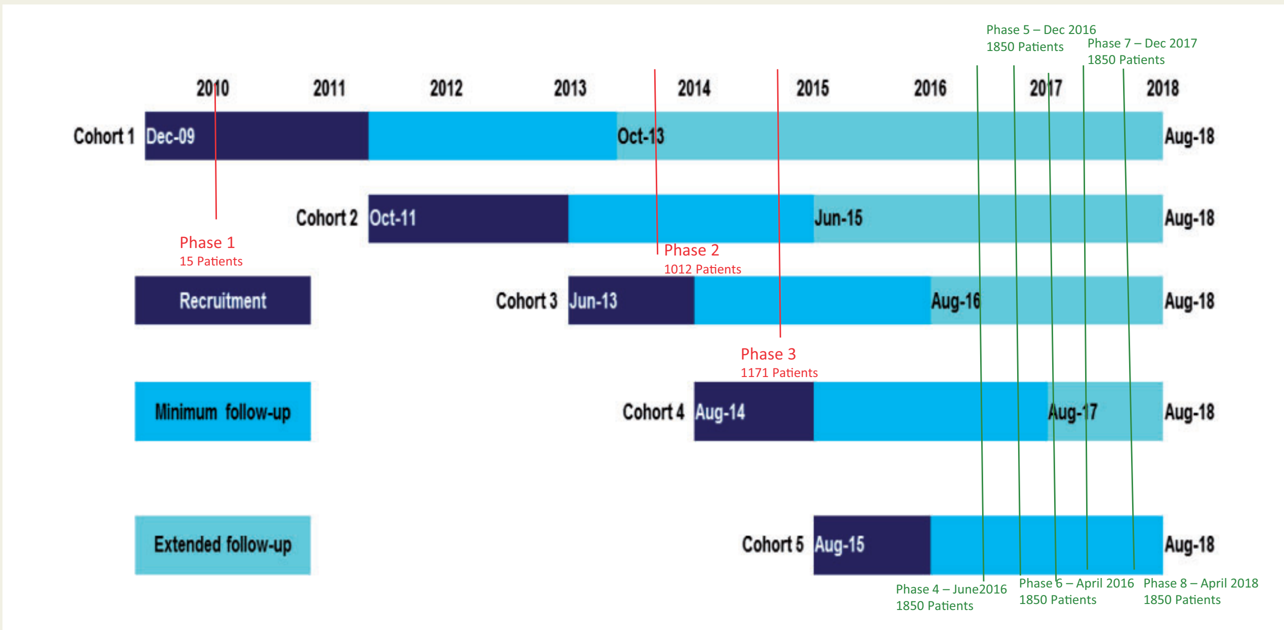


**Figure 1** GARFIELD-AF recruitment, monitoring, and reporting milestones.

**Table 1 Quality standards for the GARFIELD-AF registry**

- Audits are conducted at regular intervals, starting early during the recruitment phase and continuing until the end of follow-up, thereby allowing cost-effective and real-time trend analysis across the whole study.
- Audits include remote monitoring as well as onsite monitoring targeted at sites with potential suboptimal quality data.
- Sites with potential suboptimal quality data are identified using the following four performance measures for: data quality (using quantifiable variables), late data locking, number of missing critical variables, and a history of poor data quality from the previous phase of monitoring.
- The target for source data verification (SDV) is 20% of all records. Only critical variables that are considered essential to overall data quality are assessed.
- Audit is followed by feedback and training, then reassessment and additional onsite monitoring of poorly performing sites.
- At each audit, 80% of sites with data quality issues are selected for onsite monitoring and the results compared with the quality of data at 20% of sites which are randomly selected.
- All modifications to the data are recorded electronically in an audit trail.
- An independent professional statistician and steering committee monitors the data collection and analysis.
- Results are reported to the steering and audit committees at pre-agreed milestones to ensure proper oversight and management of the study.
- Annually, national data are fed back to sites to incentivize the ongoing recruitment and/or follow-up of patients.
- At study completion, 5% of the data for each of the critical variables for baseline data and follow-up are audited during the statistical analysis.
- All sites receive training before the start of the study and regular re-training depending on site performance, and have ongoing access to a training web portal.

## Risk-based site selection for onsite monitoring

During phase III monitoring, ∼104 of ∼1040 (i.e. 10%) sites were scheduled for onsite monitoring. Consistent with phase II monitoring, the goal was to identify ∼80 of 104 (∼80%) sites with poor quality so that resources could be targeted for onsite monitoring and for partial SDV at sites where there were potentially the greatest problems. The remaining 20% of sites were randomly selected (and served as a control for comparison with the data from poor quality sites).

An equal number of poorly performing sites (i.e. $4 \times 20 = 80$) were identified based on each of the following four complementary measures of data quality and other performance indicators:

*Data quality for 13 quantifiable variables within the eCRF* (see Supplementary material online, *Appendix Table S2*):

**Table 2** Summary of results for phase II and phase III of onsite monitoring (including source data verification)

| Variable | Phase II | Phase III |
|---|---|---|
| Patients | 1012 | 1172 |
| Sites | 110 | 104 |
| Countries | 24 | 28 |
| Source data not available* | 4475/37 243 (12.0) | 2550/92 507 (2.8) |
| Data verified (excluding blank fields) | 21 178/27 006 (78.4) | 29 121/33 005 (88.2) |
| Data verified (excluding blank fields and source data not available) | 21 178/22 531 (94.0) | 29 121/30 455 (95.6) |
| Queries (excluding blank fields and source data not available) | 1065/22 531 (4.7) | 1361/30 455 (4.5) |

*Source data were not available if information was added to the case report form that could not be confirmed within the source data provided to the monitor.

For each critical variable, the mean value across all patients at a site for that variable was assessed and, if it was not within the defined interval of $\pm 2 \times$ SEM, then the site was flagged as 'out of control' for that variable. A score from 1 to 13 was assigned to the site depending on the number of 'out of control' flags. All sites were then ranked according to the data quality score and ~20 (i.e. one-quarter of 80 sites) of the worst performing sites were selected for onsite monitoring.

*Late locking of data:* A site was designated a late lock score if late data locking occurred >30% of the time for the key milestones (i.e. baseline, 12 months and 24 months). Approximately 20 sites with the highest proportion of late-lock defaults were selected for onsite monitoring.

*Total number of missing critical events*: Sites were given a score proportional to the number of critical missing events. Missing events were identified by the data discrepancy between the event summary page of the eCRF and the completed events in the eCRF for each of the patients. Approximately 20 sites were selected for onsite monitoring on this basis.

*The findings (GCP critical and SDV discrepancies) during the previous monitoring phase.* GCP (non-compliance) findings were weighted for each site as either 'critical', 'major' or 'minor' and the 20 sites with the highest default scores were selected for onsite monitoring.

Based on the findings from these analyses, the poorest performing 78 sites from each component of the score were selected and 26 sites were also randomly selected for onsite monitoring. There was some overlap in the sites selected by the above measures: three sites selected for the missing event score were also selected by the 13-item score, the late data locking score and the GCP compliance score. For a minority of sites (~5%) that were unable or unwilling to participate in the audit, replacement sites were identified.

In total, 1172 patients at 104 sites (9.9% of the total of 1046 sites) with potential poor data quality in 28 countries were identified for onsite monitoring between December 2014 and May 2015 during phase III of the monitoring process. The distribution of sites by country is summarized on the horizontal axis of *Figure 2*.

### Testing data quality at the sites selected for onsite monitoring

The effectiveness of this site selection process was assessed by comparing the error rate (based on partial SDV) in patients at sites selected using each of the individual components of the site selection process. The results showed that the error rate based on partial SDV was greater in patients at sites selected using the 13-item score (7.1% [95% confidence interval {CI}: 6.0–8.2%] of 2031 fields in 144 patients) compared with patients from sites selected with late locking (error rate: 4.4% [95% CI: 3.5–5.3%] of 2069 fields in 91 patients) or those selected for total number of missing events (error rate: 3.2% [95% CI: 2.8–3.7%] of 5771 fields in 185 patients). The 13-item score (which provided an indicator of potential outliers in the dataset) was the most effective in identifying sites with poor data quality (based on partial SDV), with an error rate almost twice the average rate observed at the remaining sites (7.1% vs. 3.7%).

## Audit results for key quality assurance performance measures

### Concordance with source data

The process for phase III monitoring (which included partial SDV of critical variables) enabled many more patients' records and fields to be verified during phase III (92 507 fields overall in 1172 patients) than during phase II (37 243 fields in 1012 patients) (*Table 2*). Even though a greater number of patients were identified from sites with poor data quality in phase III, the level of concordance was similar in both phases—eCRFs matched patient records 94.0% of cases in phase II and 95.6% in phase III (*Table 2* and *Figure 3*). Overall, the level of concordance was >96% for all sections of the eCRF in phase II and phase III, except for two sections of the eCRF where concordance was <90%. These were 'hospitalization/procedure/consultation' (89.3%) and 'treatment change/interruption' (80.8%), in phase II. The concordance between source data and the eCRF for the 'hospitalization/procedure/consultation' section was substantially improved in phase III (97.2%); this was attributed, in part, to the improved training of sites following phase II. The proportion of cases where source data were not available decreased from 12.1% in phase II to 2.8% in phase III. Missing source data were attributed, in some cases, to legal and

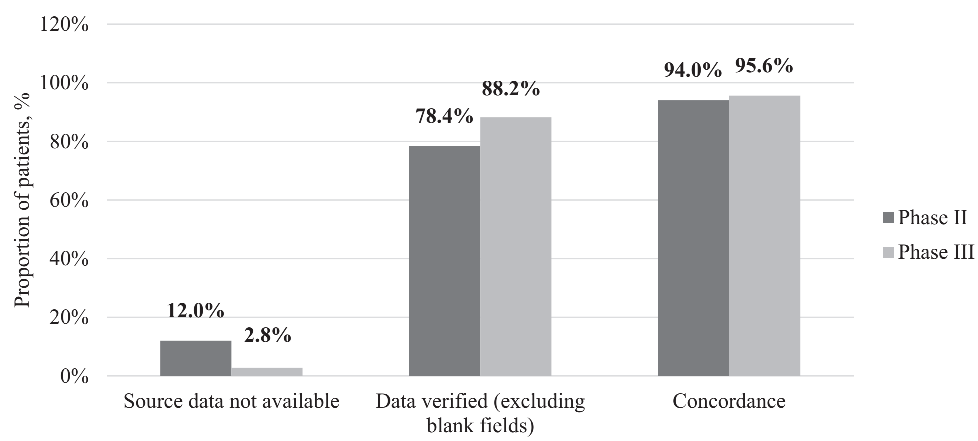


**Figure 2** Country differences in concordance (based on partial source data verification of critical variables) during phase III monitoring.

administrative restrictions and, in other cases, to the unpreparedness of patient records at the site due to the short notice prior to onsite visits. There was no indication that the eCRF had been completed in the absence of source data.

Concordance with source data was similar in all geographic regions during the monitoring in both phase II (Americas 96.7%; Asia-Pacific 94.9%; Europe, the Middle East, and Africa [EMEA] 95.3%) and phase III (Americas 94.9%; Asia-Pacific 95.1%; EMEA 95.9% Phase III). Comparison of concordance by country (*Figure 2*) identified some countries where there was a notable deviation from the protocol in the recording of data. In some countries, this discrepancy was attributed to the low number of sites in the registry at the time of phase III monitoring.

## Monitoring of missing events

The monitoring of missing events (i.e. events recorded in the source data but not in the eCRF) found that the number of missing events was low and diminished from phase II to phase III. During phase II onsite monitoring, 23 bleed events, 14 stroke events and 12 deaths were identified as missing in 1012 patients. During phase III monitoring, 10 bleed events, 3 stroke events and 12 deaths in 1172 patients were identified as missing and the eCRFs updated.

Missed hospitalization events were also frequent during phase II monitoring (350/1156 [30.3%]), but less frequent in phase III (402/2288 [14.9%]) in the worst performing sites and randomly selected sites (138/842 [16.4%]). These missing events were predominantly due to sites recording only AF-related hospitalizations. Following this finding, the coordinating centre (Thrombosis Research Institute, London, UK) provided additional training to all sites. Additional onsite monitoring of poor performing sites was also performed in order to ensure that the recording of data on key endpoints (stroke, all-cause mortality, bleeding) was adequately addressed. To assess the impact of training, further audit of missing events at poorly performing sites will also be captured in the next audit phase.

## Other GCP findings

In phase II, most of the findings identified in a subset of 30 out of 110 sites related to use of the incorrect version of the informed consent form (ICF). At two sites (in five patients), the missing ICF was not recovered from the patients' files. These five patient records were not included the database and the related data were not analyzed. In phase III, most of the discrepant findings at 31 of 104 sites were also related to use of the incorrect version of the ICF. One finding was related to a site breach of eligibility criteria and five to breaches of the GCP informed consent process. As a part of the Corrective Action Preventative Action (CAPA), the patient data were excluded from the analyses for these records. The site staff have received retraining in the protocol and regulatory requirements.

# Discussion

The overall objectives of the audit process in GARFIELD-AF were to: evaluate the compliance of the protocol with GCP and local regulations, the quality and completeness of data and source documentation, and the concordance between eCRF data and source documents, and to identify potentially unreported outcome events of interest.

Whereas standards for the design and conduct of large-scale randomized trials have evolved and gained acceptance,[30] registries and observational datasets vary substantially in their design, their conduct, the extent to which they utilize routinely reported data, and the extent to which they audit and validate outcome and safety data.[28] Some registry and observational programs employ retrospective data derived from routinely collected information. In such retrospective programs there are key challenges, including the variability in defining clinically recorded outcomes, inconsistency of recording of baseline characteristics and outcome measures, and uncertain or absent verification of key data. Prospective registries with predefined baseline characteristics and outcome measures and with defined quality standards have the potential to provide more robust datasets.

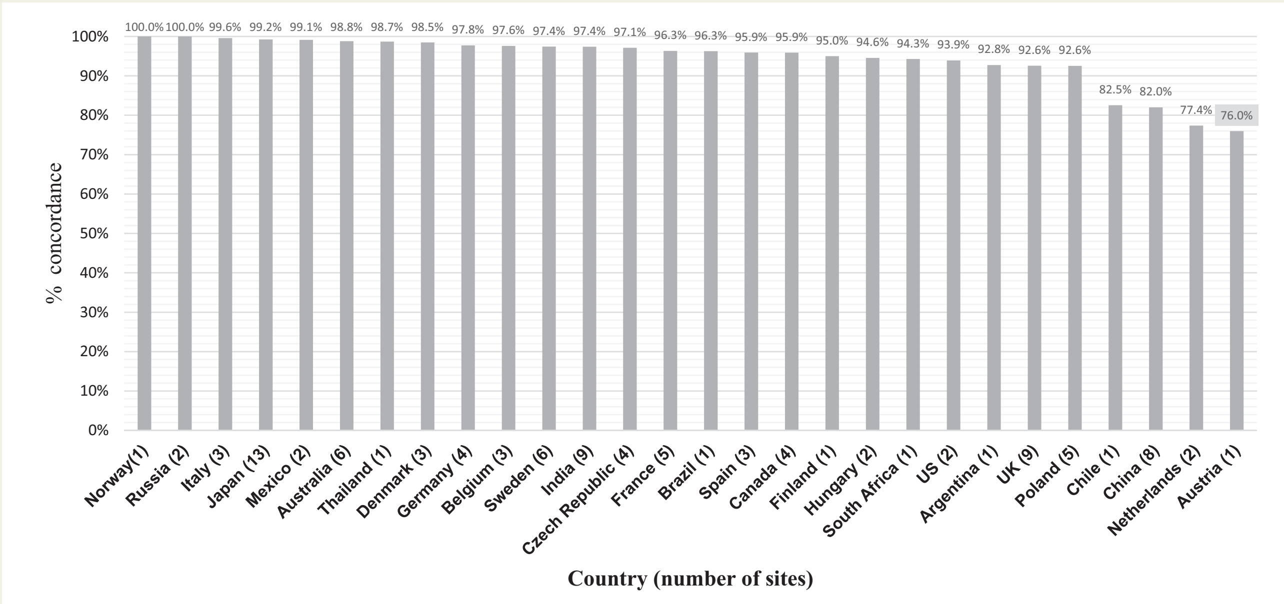


**Figure 3** Audit results from phase II (1012 patients at 110 sites in 28 countries) and phase III (1172 patients at 104 sites in 35 countries) monitoring.

There have been attempts to establish key criteria for the validity of registries,[31] but as yet, no consensus exists. However, independent reviewers have proposed a number of key criteria to help establish the validity of registries and their interpretation (see Supplementary material online, *Appendix Table S3*).[29] The extent to which large-scale registries and observational studies fulfil these criteria varies substantially and this impacts on the interpretation of reports from the respective studies. The independent Audit Committee has reviewed these criteria in the GARFIELD-AF program [12,13] and has determined that the registry meets these criteria. In addition, the committee has implemented further quality standards (see Supplementary material online, *Appendix Table S3*).

Complete SDV in registries is neither practical nor cost-effective. Published studies have shown that only a very small percentage of data is changed due to SDV of all fields within a record, and the effect of this change on the primary analysis is minimal.[16,18,32] Robust sampling strategies with SDV of up to 20% of records may be sufficient, without clinically significant differences in the primary analysis.[22] In the audit of GARFIELD-AF data, we found that there was a similar level of concordance for SDV of whole records compared with SDV of the critical variables. During onsite monitoring, for the 9.9% of the poorest sites in terms of data quality (derived according the GARFIELD-AF data quality score), <5% discrepancies between the electronic records and the site-verified source data were found during phase III monitoring. This is considered to be within the acceptable bounds in the field of clinical trials for regulatory approval. Houston *et al.*[22] recently determined that an error rate of 5% or less within electronic datasets for RCTs should be the 'gold standard' for determining data quality within a clinical setting. The GARFIELD-AF registry met this standard. Where data quality issues were identified in certain countries and centres, early corrective action including onsite training (improved knowledge of the data management system) and further clarification of the eCRF ensured ongoing quality improvements (e.g. to ensure that data on all-cause hospitalizations were appropriately captured at all sites). Regular audit, annual deadlines for data locks and additional onsite monitoring of poorly performing sites in between audits has also been an essential element of the registry design to ensure that all events are captured (*Figure 1*). The data submitted for publication from GARFIELD-AF are based only on locked data where efforts are made to ensure that all events are captured; while the data presented in this article include information on interim data (i.e. 'unlocked' data where either the whole, or part, of the eCRF data were not finalized).

In summary, no single monitoring approach is applicable to all studies. The frequency and extent of monitoring need to be appropriate, and achieve a balance between reliable data integrity and ease of enrolment and follow-up. Audit approaches should be tailored to the objectives of the study and may combine a number of different monitoring methods that allow cost-effective and real-time trend analysis. GARFIELD-AF adopted a dual auditing scheme using remote monitoring as well as onsite monitoring targeted at sites with potential suboptimal quality data. This approach may be useful for other large-scale registries. The GARFIELD-AF sets high standards for a large-scale registry (summarized in *Table 1*). Starting early in the recruitment of patients into the registry, eight audits were planned across all phases of the recruitment and monitoring so that by the end of the study, 20% of all eCRFs will have been monitored (*Figure 1*). Only critical variables that are considered essential to overall data quality are assessed during SDV. For example, in GARFIELD-AF, baseline characteristics important to the research question (such as components of the $CHA_2DS_2$-VASc score for assessing the risk of stroke) and outcomes (stroke, bleeding events and death) are audited. Audits should be followed by feedback and training, then reassessment. In GARFIELD-AF, the results of the previous audit is used to facilitate corrective action on data quality issues and forms a baseline against which quality improvements are assessed in the next monitoring phase. All sites receive regular re-training depending on site performance, and have ongoing access to a training web portal. At regular intervals, results are reported to the steering and audit committees to ensure proper oversight and management of the study. Finally, national data are also fed back to sites to incentivize the ongoing recruitment and/or follow-up of patients. Through the implementation of the standards outlined in *Table 1*, we believe that GARFIELD-AF has the potential to inform a future 'reference standard' for the successful delivery of high-quality data from registries.

# Supplementary material

Supplementary material is available at *European Heart Journal – Quality of Care and Clinical Outcomes* online.

## Acknowledgements

We would like to thank Karen Pieper from the Statistical Research Science department at Duke Clinical Research Institute (Durham, NC) for her support in the independent adjudication of the statistical analyses. The eCRF was designed by Dendrite Clinical Systems Ltd (Henley-on-Thames, UK). Oversight of operations and data management was conducted by the sponsor and coordinating centre (Thrombosis Research Institute, London, UK), with support from Quintiles (Durham, NC), The University of Birmingham Department of Primary Care Clinical Sciences (UK), Thrombosis Research Group—Brigham and Women's Hospital, Boston (USA), and AIXIAL (France). Submitted data were examined for completeness and accuracy by the Thrombosis Research Institute (London, UK). Editorial assistance was provided by Rae Hobbs from the Thrombosis Research Institute (London, UK).

## Funding

The GARFIELD-AF registry is an independent academic research initiative sponsored by the Thrombosis Research Institute (London, UK) and supported by an unrestricted research grant from Bayer Pharma AG (Berlin, Germany).

**Conflict of interest:** K.A.A.F. has received personal fees from: Bayer, Johnson and Johnson, Lilly, Astra Zeneca, and Sanofi/Regeneron and research grants from Bayer, Johnson and Johnson, and Astra Zeneca; B.J.G. has received personal fee from TRI A.J.C. is an advisor, and has received personal fees, from Bayer, Boehringer Ingelheim, Pfizer/BMS and research grants from Daiichi Sankyo. A.K.K. is a principal investigator/received research grants from Bayer Healthcare; he received personal fee from Bayer Healthcare, Boehringer-Ingelheim Pharma, Daiichi Sankyo Europe, Sanofi SA and Janssen Pharma. G.K., S.T., A.K., and S.S. have no conflict of interest.

## References


1. Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. *Am J Cardiol* 2013;**112**:1142–1147.
2. Krijthe BP, Kunst A, Benjamin EJ, Lip GY, Franco OH, Hofman A, Witteman JC, Stricker BH, Heeringa J. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J* 2013;**34**:2746–2751.
3. Rahman F, Kwan GF, Benjamin EJ. Global epidemiology of atrial fibrillation. *Nat Rev Cardiol* 2014;**11**:639–654.
4. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Despres JP, Fullerton HJ, Howard VJ, Huffman MD, Isasi CR, Jimenez MC, Judd SE, Kissela BM, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Magid DJ, McGuire DK, Mohler ER, 3rd, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Rosamond W, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Woo D, Yeh RW, Turner MB. Executive summary: heart disease and stroke statistics–2016 update: a report from the American Heart Association. *Circulation* 2016;**133**:447–454.
5. Lee WC, Lamas GA, Balu S, Spalding J, Wang Q, Pashos CL. Direct treatment cost of atrial fibrillation in the elderly American population: a medicare perspective. *J Med Econ* 2008;**11**:281–298.
6. Stewart S, Murphy NF, Walker A, McGuire A, McMurray JJ. Cost of an emerging epidemic: an economic analysis of atrial fibrillation in the UK. *Heart* 2004;**90**:286–292.
7. Kakkar AK, Mueller I, Bassand JP, Fitzmaurice DA, Goldhaber SZ, Goto S, Haas S, Hacke W, Lip GY, Mantovani LG, Verheugt FW, Jamal W, Misselwitz F, Rushton-Smith S, Turpie AG. International longitudinal registry of patients with atrial fibrillation at risk of stroke: Global Anticoagulant Registry in the FIELD (GARFIELD). *Am Heart J* 2012;**163**:13–9 e1.
8. Huisman MV, Lip GY, Diener HC, Dubner SJ, Halperin JL, Ma CS, Rothman KJ, Teutsch C, Zint K, Ackermann D, Clemens A, Bartels DB. Design and rationale of global registry on long-term oral antithrombotic treatment in patients with atrial fibrillation: a global registry program on long-term oral antithrombotic treatment in patients with atrial fibrillation. *Am Heart J* 2014;**167**:329–334.
9. Piccini JP, Fraulo ES, Ansell JE, Fonarow GC, Gersh BJ, Go AS, Hylek EM, Kowey PR, Mahaffey KW, Thomas LE, Kong MH, Lopes RD, Mills RM, Peterson ED. Outcomes registry for better informed treatment of atrial fibrillation: rationale and design of ORBIT-AF. *Am Heart J* 2011;**162**:606–12 e1.
10. Steinberg BA, Blanco RG, Ollis D, Kim S, Holmes DN, Kowey PR, Fonarow GC, Ansell J, Gersh B, Go AS, Hylek E, Mahaffey KW, Thomas L, Chang P, Peterson ED, Piccini JP, Investigators O-ASC. Outcomes registry for better informed treatment of atrial fibrillation II: rationale and design of the ORBIT-AF II registry. *Am Heart J* 2014;**168**:160–167.
11. Kirchhof P, Ammentorp B, Darius H, De Caterina R, Le Heuzey JY, Schilling RJ, Schmitt J, Zamorano JL. Management of atrial fibrillation in seven European countries after the publication of the 2010 ESC Guidelines on atrial fibrillation: primary results of the PREvention oF thromboemolic events–European Registry in Atrial Fibrillation (PREFER in AF). *Europace* 2014;**16**:6–14.
12. Cohen AT, Goto S, Schreiber K, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting? *Eur Heart J Supplements* 2015;**17**(Suppl D):D2–D8.
13. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br J Cancer* 2014;**110**:551–555.
14. Brown ML, Gersh BJ, Holmes DR, Bailey KR, Sundt TM III. From randomized trials to registry studies: translating data into clinical information. *Nat Clin Pract Cardiovasc Med* 2008;**5**:613–620.
15. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline: *Guideline for Good Clinical Practice* 1996;**E6**.
16. Andersen JR, Byrjalsen I, Bihlet A, Kalakou F, Hoeck HC, Hansen G, Hansen HB, Karsdal MA, Riis BJ. Impact of source data verification on data quality in clinical trials: an empirical post hoc analysis of three phase 3 randomized clinical trials. *Br J Clin Pharmacol* 2015;**79**:660–668.
17. Olsen R, Bihlet AR, Kalakou F, Andersen JR. The impact of clinical trial monitoring approaches on data integrity and cost – a review of current literature. *Eur J Clin Pharmacol* 2016;**76**:660–668.
18. De S. Hybrid approaches to clinical trial monitoring: practical alternatives to 100% source data verification. *Perspect Clin Res* 2011;**2**:100–104.
19. Rationale and design of the GRACE (Global Registry of Acute Coronary Events) Project: a multinational registry of patients hospitalized with acute coronary syndromes. *Am Heart J* 2001;**141**:190–199.
20. Verhulst K, Artiles-Carloni L, Beck M, Clarke JT, Neto JC, Cox GF, Fernhoff PM, Guffon N, Kong Y, Martins AM, Tylki-Szymanska A, Whitley CB, Wijburg FA, Wraith EJ, Koepper CM. Source document verification in the Mucopolysaccharidosis Type I Registry. *Pharmacoepidemiol Drug Saf* 2012;**21**:749–752.
21. Clinical Trials Transformation Initiative. Effective and efficient monitoring as a component of quality assurance in the conduct of clinical trials. 2012. http://www.ctti-clinicaltrials.org/ (March 2016).
22. Houston L, Probst Y, Humphries A. Measuring data quality through a source data verification audit in a clinical research setting. *Stud Health Technol Informatics* 2015;**214**:107–113.

23. Sheetz N, Wilson B, Benedict J, Huffman E, Lawton A, Travers M, Nadolny P, Young S, Given K, LF. Evaluating source data verification as a quality control measure in clinical trials. *Ther Innov Regul Sci* 2014;**48.6**:671–680.
24. Messenger JC, Ho KK, Young CH, Slattery LE, Draoui JC, Curtis JP, Dehmer GJ, Grover FL, Mirro MJ, Reynolds MR, Rokos IC, Spertus JA, Wang TY, Winston SA, Rumsfeld JS, Masoudi FA, Science N, Quality Oversight Committee Data Quality W. The National Cardiovascular Data Registry (NCDR) data quality brief: the NCDR Data Quality Program in 2012. *J Am Coll Cardiol* 2012;**60**:1484–1488.
25. Fox KA. Registries and surveys in acute coronary syndrome. *Eur Heart J* 2006;**27**:2260–2262.
26. Alves-Cabratosa L, Garcia-Gil M, Comas-Cufi M, Ponjoan A, Marti R, Parramon D, Blanch J, Ramos R. Incident atrial fibrillation hazard in hypertensive population: a risk function from and for clinical practice. *Hypertension* 2015;**65**:1180–1186.
27. LaPar DJ, Speir AM, Crosby IK, Fonner E Jr, Brown M, Rich JB, Quader M, Kern JA, Kron IL, Ailawadi G. Postoperative atrial fibrillation significantly increases mortality, hospital readmission, and hospital costs. *Ann Thorac Surg* 2014;**98**:527–533; discussion 33.
28. Roovers JP. Registries: what level of evidence do they provide? *Int Urogynecol J Pelvic Floor Dysfunct* 2007;**18**:1119–1120.
29. Alpert JS. Are data from clinical registries of any value? *Eur Heart J* 2000;**21**:1399–1401.
30. Clinical Trials Transformation Initiative. Quality by design project – critical to qulity (CTQ) factor principles document. http://www.ctti-clinicaltrials.org/sites/www.ctti-clinicaltrials.org/files/QbD_toolkit/Principles%20Document_finaldraft_19MAY5.pdf (May 2015).
31. Dreyer NA, Velentgas P, Westrich K, Dubois R. The GRACE checklist for rating the quality of observational studies of comparative effectiveness: a tale of hope and caution. *J Manag Care Spec Pharm* 2014;**20**:301–308.
32. Tudur Smith C, Stocken DD, Dunn J, Cox T, Ghaneh P, Cunningham D, Neoptolemos JP. The value of source data verification in a cancer clinical trial. *PLoS One* 2012;**7**:e51623.