U.S. Food and Drug Administration Inspections of Clinical Investigators: Overview of Results from 1977 to 2009

Sonia K. Morgan-Linnell1, David J. Stewart3, and Razelle Kurzrock2

Abstract

The U.S. Food and Drug Administration (FDA) Center for Drug Evaluation and Research is responsible for evaluating drug safety and efficacy, including oversight of clinical trials and principal investigators. The FDA Clinical Investigator Inspection List (CIIL) contains online, detailed, relevant information of all FDA inspections. We reviewed FDA inspections of clinical investigators to ascertain their outcome and included all inspections on the list (July 1977 through December 31, 2009; n = 9,481 inspections). Eighty-eight percent of inspections were "data audit" (primary purpose = verification of data), and the rest (12%) were "for cause." The number of inspections each year significantly increased over time (P < 0.0001) and averaged 350 per year in the past decade. No deficiencies were found in only 11.2% of all "data audit" and 5% of all "for cause" inspections. Only 31% of inspections resulted in "no action indicated." About two thirds of inspections resulted in some finding, requiring either voluntary investigator action (61.3% of inspections) or official FDA action (3.9%). The most frequently cited deficiencies were failure to follow investigational plan (34%), inadequate informed consent form (28%), and inadequate/inaccurate records (27%). In conclusion, over the past decade, the FDA has performed approximately 350 inspections per year, with the number increasing over time. The vast majority of FDA inspections yield deficiency findings and, as a result, only about one third of inspections have an outcome of "no action indicated." Clin Cancer Res; 20(13); 3364–70. ©2014 AACR.

Introduction

The U.S. Food and Drug Administration (FDA) was created in the 1930s (1) and is responsible for protecting human health through its regulation of food, drugs, medical devices, cosmetics, and various other health products (2). The Center for Drug Evaluation and Research is responsible for evaluating drug safety, including new, generic, and over-the-counter drug products (3). Since the 1960s, ensuring efficacy has been added to the FDA mandate.

Part of the FDA jurisdiction involves oversight of clinical trials, as defined by the Code of Federal Regulations (CFR) 312.53 (4). Clinical investigator inspections fall under the Division of Scientific Investigations and are conducted under the Bioresearch Monitoring Program. The FDA may conduct announced or unannounced investigations of clinical trials conducted by clinical investigators. The FDA clinical investigator inspection outcomes are published on a quarterly basis, and the records since 1977 are available online (5).

FDA inspections are typically conducted for the following reasons: (i) to verify the accuracy of clinical investigation data being submitted to the FDA as part the drug approval process, (ii) to respond to a complaint against the clinical investigator, (iii) to respond to concerns by the sponsor of the trial, and/or (iv) to determine that the clinical investigator is protecting research subjects. The FDA also initiates investigations for certain classes of drugs that are considered of particular concern for public welfare and safety.

Clinical investigators who fail to comply with protocol requirements can be subject to revocation of investigator privileges indefinitely or until adequate assurances have been provided to the FDA that they have put in place protections to prevent further deficiencies in trial conduct. The list of investigators that have provided adequate assurances or who are disbarred from conducting clinical trials is provided to the public by the FDA online (6).

Here, we review the results of the FDA Clinical Investigator Inspection List data from July 1977 through December 2009 (5–8). It has been well documented that the complexity of clinical trials has increased substantially in this period of time (9–15). Hence, the purpose of this review was to describe the frequency with which compliance problems are observed, with the rationale for the study being that compliance in the face of increasing protocol complexity might present challenges that should be
ascertained, so that they can be addressed. These challenges might ultimately affect the ability/willingness of investigators and sponsors to perform clinical trials in the United States and are important to the process of ensuring data integrity, ethical and safe conduct of clinical research, and fiscally responsible drug development.

Materials and Methods

Data collection

We analyzed data from the FDA Clinical Investigator Inspection List 5, which contains the names, addresses, inspection date, inspection outcome, and other relevant information of all FDA inspections from July 1977. We used a cutoff date of December 31, 2009, to help ensure that all, or at least most, ongoing investigations had been closed. These inspections were conducted by the FDA as part of the Bioresearch Monitoring Program, the purpose of which is to ensure the quality of the clinical studies that provide data to the FDA, as well as ensure the rights of human subjects participating in the clinical trials. Data were downloaded from the FDA text file5 and then were parsed and analyzed.

Statistical analysis

Linear regression analyses were conducted to determine change over time periods. The Student t test was used to determine differences between time periods. Statistical significance was defined as P ≤ 0.05.

Results

Number of inspections

From July 1977 through December 31, 2009, 9,481 unique FDA inspections were reported. For the purpose of this study, if an investigator had multiple inspections that were performed on different dates, each inspection date was considered a separate investigation. Only inspections that were closed are reported by the FDA and thus able to be analyzed.

We found that the number of inspections each year increased significantly over time (P < 0.0001) and in the decade from 1990 to 1999 (n = 3,164, P = 0.02); however, there was not a statistically significant increase in the number of inspections per year for the decade from 2000 to 2009 (n = 3,495, P = 0.12; Fig. 1). There was a statistically insignificant increase in the number of inspections comparing the decade from 2000 to 2009 (n = 3,495) versus 1990 to 1999 (n = 3,164, P = 0.18). The most inspections occurred in 1996 (n = 446), and the average number of inspections per year was 316 over the years from 1990 to 1999 and 350 over the years from 2000 to 2009.

Types of inspections

The FDA classifies the type of inspection as follows: (i) data audit: an inspection in which the focus is on the verification of study data, (ii) for cause: an inspection in which the focus is on the conduct of the study by the Clinical Investigator, or (iii) information gathering (7). None of the inspections in the data over the period we analyzed were categorized as information gathering (data not shown). Of the total 9,481 inspections analyzed, 8,351 were “data audit” inspections (88.1%), 1,129 were “for cause” inspections (11.9%), and one had no category listed. The highest number of “data audit” inspections occurred in 1996 (n = 441), and the highest number of “for cause” inspections occurred in 2003 (n = 97; Fig. 1). Interestingly, the lowest number of “for cause inspections” (n = 5) occurred in 1996, which was the year of the highest number of “data audit” inspections (Fig. 1). We found that the total number of “data audit” and “for cause” inspections increased over time.
Since 1983, there has been a statistically significant increase in the percentage of inspections that were "for cause" ($P = 0.03$), and this increase is even more evident when comparing the percentage of inspections that were "for cause" for the years 1990 to 1999 versus 2000 to 2009 (mean = 4.0% and 18.3%, respectively; $P < 0.0001$).

### Inspection outcomes

The FDA designates inspection outcomes by classification codes, which are shown in Table 1. Table 1 also defines each of the codes. For the purposes of this report, we have grouped all voluntary action indication (AI) and official action indication codes together for analysis. Voluntary action indication was the most common classification code ($n = 5,814; 61.3\%$ of all inspections), followed by no action indication ($n = 2,951; 31.1\%$ of all inspections), official action indication ($n = 372; 3.9\%$ of all inspections), and canceled (CANC) ($n = 283; 3.0\%$; Table 1, Fig. 2A). The percentage of the total inspections, by year, for each classification code is shown in Fig. 2B. Overall, the percentage of inspections classified as no action indication increased

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NAI</td>
<td>No action indicated</td>
<td>No objectionable conditions or practices were found during the inspection</td>
<td>2,951</td>
<td>905</td>
<td>1,465</td>
</tr>
<tr>
<td>VAI</td>
<td>Voluntary action indicated</td>
<td>Objectionable conditions were found but the problems do not justify further regulatory action; any corrective plan is left to the investigator to take voluntarily</td>
<td>5,814</td>
<td>1,901</td>
<td>1,819</td>
</tr>
<tr>
<td>VAI1</td>
<td>Voluntary action indicated 1</td>
<td>Correction made on site</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>VAI2</td>
<td>Voluntary action indicated 2</td>
<td>No response requested</td>
<td>2,159</td>
<td>493</td>
<td>61</td>
</tr>
<tr>
<td>VAI2C</td>
<td>Voluntary action indicated 2 consent</td>
<td>Consent problems found</td>
<td>315</td>
<td>127</td>
<td>0</td>
</tr>
<tr>
<td>VAI3</td>
<td>Voluntary action indicated 3</td>
<td>Response requested</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>VAI3C</td>
<td>Voluntary action indicated 3 consent</td>
<td>Case closed</td>
<td>89</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>VAI3F</td>
<td>Voluntary action indicated action indicated 3 follow-up</td>
<td>Follow-up for cause inspection issued</td>
<td>76</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>VAI3R</td>
<td>Voluntary action indicated 3 response</td>
<td>Response received and accepted</td>
<td>199</td>
<td>37</td>
<td>2</td>
</tr>
<tr>
<td>VAIR</td>
<td>Voluntary action indicated response</td>
<td>30-day response requested</td>
<td>39</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>VAIRC</td>
<td>Voluntary action indicated closed</td>
<td>30-day response requested and case closed</td>
<td>25</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>VAIRR</td>
<td>Voluntary action indicated response received</td>
<td>30-day response requested, received, and accepted</td>
<td>157</td>
<td>68</td>
<td>89</td>
</tr>
<tr>
<td>OAI</td>
<td>Official action indicated</td>
<td>Objectionable conditions were found and regulatory and/or administrative sanctions by FDA are indicated</td>
<td>372</td>
<td>93</td>
<td>154</td>
</tr>
<tr>
<td>OAIR</td>
<td>Official action indicated completed</td>
<td>Completed</td>
<td>143</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>OAIRR</td>
<td>Official action indicated response</td>
<td>Response received and accepted</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>OAIW</td>
<td>Official action indicated warning</td>
<td>Warning letter issued</td>
<td>16</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>CANC</td>
<td>Canceled</td>
<td>The inspection was canceled before the inspection was started</td>
<td>283</td>
<td>254</td>
<td>15</td>
</tr>
<tr>
<td>WASH</td>
<td>Washout</td>
<td>An inspection was initiated but no meaningful information could be attained</td>
<td>36</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>REF</td>
<td>Reference</td>
<td>The FDA does not define this code</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MTF</td>
<td>Memo-to-file</td>
<td>Case closed with an MTF</td>
<td>23</td>
<td>0</td>
<td>23</td>
</tr>
</tbody>
</table>
(\(P < 0.0001\)), whereas the percentage of voluntary action indication inspections decreased (\(P < 0.0001\)). There was no significant change in the percentage of inspections classified as official action indication (\(P = 0.20\)).

From 1990 to 1999, the percentage of no action indication increased (\(P = 0.0001\)) and the percentage of voluntary action indication decreased (\(P = 0.001\)), whereas there was no significant change in the percentage of official action indication (\(P = 0.18\)). From 2000 to 2009, the percentage of inspections classified as no action indication and official action indication increased (\(P = 0.01\) and 0.006, respectively), whereas the percentage classified as voluntary action indication increased (\(P = 0.001\)).
indicating decreased ($P = 0.01$). Comparing the years 2000–2009 versus the years 1990–1999, (i) the percentage of no action indication inspections was significantly increased in 2000–2009 ($P = 0.006$); (ii) the percentage of voluntary action indication inspections decreased in 2000–2009 ($P = 0.04$); and (iii) the percentage of official action indication inspections showed a statistically significant upward trend in 2000–2009 ($P = 0.13$).

When we analyzed the outcomes of the “data audit” versus “for cause” inspections, we found significant differences in the outcomes (Supplementary Fig. S1). The largest difference was that the outcome of only 2% of all data audit inspections was official action indication (Supplementary Fig. S1A), whereas 19% of all “for cause” inspections were official action indication ($P < 0.0002$; Supplementary Fig. S1B). In addition, a larger percentage of “for cause” inspections was closed with a memo-to-file (MTF; $P < 0.0002$; almost all inspections eventually classified as MTF at closure were “for cause”). A smaller percentage of “for cause” versus “data audit” inspections were no action indication (21% vs. 33%, $P < 0.0002$), voluntary action indication (52% vs. 60%, $P < 0.0002$), and canceled (2% vs. 4%, $P = 0.0015$). There was no difference in the percentage of inspections classified as washout with no meaningful information ($P = 0.12$); however, these numbers were very small ($n = 129$, “data audit” and $n = 25$, “for cause”).

**Deficiency codes**

The FDA assigns deficiency codes to inspections where issues, according to the CFR, are found [7]. In addition to the coded deficiency (coded as 00-21, with 00 being no deficiency), 2,688 of the 9,481 inspections (28.4%) had a blank in the deficiency column. Of the 2,688 inspections with blanks in the deficiency column, 262 (9.7%) were voluntary action indication ($n = 2,562/9,481, 27.0%$), inadequate drug accountability ($n = 1,437/9,481, 15.2%$ of all inspections), and failure to report adverse drug reactions (15.2% of all inspections). For the period from 2000 to 2009 versus 1990 to 1999, no significant difference was found in the percentage of inspections coded as no deficiencies (00; 8.7% vs. 9.7% of all inspections for the respective decade; $P = 0.35$). Failure to follow investigational plan showed a trend upward between these two decades and was found in 30.4% versus 23% of inspections for the respective decades ($P = 0.07$). Inadequate and inaccurate records (23.1% vs. 18.4% of all inspections for the respective decade; $P = 0.45$) or failure to report adverse drug reactions (5.4% vs. 5.7% of all inspections for the respective decade; $P = 0.63$) were unchanged in the decades from 2000 to 2009 versus 1990 to 1999. For the period from 2000 to 2009 versus 1990 to 1999, the percentage of inspections found to have inadequate informed consent form and inadequate drug accountability was significantly decreased (5.2% vs. 22.2%, $P = 0.004$ and 7.8% vs. 9.1%, 0.0001, respectively).

Almost half of all “for cause” inspections were cited for failure to follow investigational plan (45.1%) and inadequate information consent forms (41.7%; Supplementary Fig. S2B). These were also the two most commonly found deficiencies in “data audit” inspections; however, they occurred in a much lower percentage of all “data audit” inspections (29.6% and 26.9%, respectively). In addition, no deficiencies (coded as 00) were found in only 11.2% of all “data audit” and 5% of all “for cause” inspections.

The FDA also most recently (2005) added three new deficiency codes: (i) failure to supervise or personally conduct the clinical investigation (found in 17 inspections from 2005 through 2009); (ii) failure to protect the rights, safety, and welfare or subjects ($n = 10$); and (iii) failure to permit FDA access to records ($n = 0$).

**Discussion**

We analyzed FDA data from the Clinical Investigator Inspection List for the period of time from July 1977 through December 2009 ($n = 9,481$ inspections). We found that the number of inspections each year increased significantly over time ($P = 0.02$). The majority of FDA inspections were “data audit” inspections (88.1%) as opposed to “for cause” inspections (11.9%), and the percentage of inspections each year that were “for cause” has significantly increased since 1983 ($P = 0.03$), peaking in 2003 when 28.5% of all inspections were “for cause.”

It is likely that the statistically significant increase in the total number of inspections per year is at least in part due to the increase in the number of clinical trials conducted since the 1970s. However, since 2005, the number of clinical trials registered on clinicaltrials.gov has not significantly changed [8]. Between 2000 and 2009, we found a statistically insignificant upward trend in the number of inspections conducted per year ($P = 0.12$).

We found that there was a significant difference in the outcomes of the “data audit” and “for cause” inspections. It is perhaps not surprising that a larger percentage of “for cause” audits were official action indication (19% vs. 2%, $P < 0.0002$), whereas a smaller percentage were no action indication (21% vs. 33%, $P < 0.0002$). The reason for this observation might be that “for cause” audits were conducted because a known problem with a clinical investigator was reported to the FDA. However, it also may be that if a problem with an investigator is suspected, the inspector has a propensity to investigate more closely.

The most common deficiencies cited for inspections were failure to follow investigational plan ($n = 3,202, 33.8%$ of all inspections), followed by inadequate informed consent
form (n = 2,661, 28.1%), inadequate and inaccurate records (n = 2,562, 27.0%), inadequate drug accountability (n = 1,437, 15.2%), and failure to report adverse drug reactions (n = 562, 5.9%). For the period from 2000 to 2009 versus 1990 to 1999, there was no significant change in the percentage of inspections with no deficiencies noted (P = 0.35), inadequate and inaccurate records (P = 0.45), or failure to report adverse drug reactions (P = 0.63). However, the percentage of inspections found to have inadequate informed consent form and inadequate drug accountability were significantly decreased (P = 0.0001 and 0.004, respectively) for the years 2000–2009 as compared with 1990–1999. The percentage of inspections noted as failure to follow investigational plan showed an upward trend in the more recent decade versus 1990–1999 (P = 0.07).

In recent years, the complexity of clinical trials has increased (9–13). For instance, eligibility criteria for industry-sponsored clinical trials have increased greatly; the Tufts Center for the Study of Drug Development found that 48% of patients screened for clinical trials over the period from 1990 to 1999 went on to complete the trial, whereas only 23% did so for the period from 2000 to 2009 (10, 13). Assuming that the decrease is at least partly due to screen failures, this may be because the average number of inclusion criteria increased from 10 in 1999 to 26 in 2005. Another study found that for the period from 2003 to 2006 versus the period from 1999 to 2003, the number of eligibility criteria tripled, which would significantly restrict inclusion (11). Excessive eligibility restrictions can be problematic, especially when imposed without a clear scientific rationale. Furthermore, if a drug is approved, the patients receiving the drug in the community may not resemble the healthier cohort of individuals who participated in the trials that led to approval. While the increase in “failure to follow the investigational plan” is either a physician or patient compliance issue, it is conceivable that a number of these studies are poorly written and place unfair burdens on the subjects, or are very difficult for the investigators to manage. Investigators should also not undertake studies that they lack the resources to conduct. However, in the context of a large number of studies with increasing eligibility requirements (as mentioned above; ref. 11), and the fact that some of the drugs in these trials are perceived to have significant salutary effects, it may at times be hard for investigators to turn down these studies despite their difficult-to-manage restrictions (14).

In addition to the increased stringency in eligibility criteria, the number of study-related procedures has increased substantially as well (9–13). Getz and colleagues (11) reported that between 1999 and 2005, the number of study procedures per protocol for all phases and therapeutic areas increased by an average of 6.5% annually, and the number of procedures per protocol for gastrointestinal, anesthesia, and ophthalmology protocols increased annually by 30.5%, 22.5%, and 21.2%, respectively. A separate study found that the average number of protocol-required procedures for industry-sponsored studies increased from an average of 90 between 1999 and 2002 to an average of 150 from 2003 to 2005 (10). A review of phase I oncology trials in our department found similar increases in study-related procedures [45 mandated events in the first 4 weeks on trial for studies between 2004 and 2007 (ref. 9) vs. 105 events for pharmaceutically sponsored studies activated after January 2009; (Kurzrock, unpublished data)]. This increased complexity, especially for industry-sponsored studies, also increases the work burden on clinical investigators, as well as the burden on the participants enrolling in these trials. The average length of a case report form increased from 55 pages in 1999–2002 to an average of 180 pages in 2003–2005, an increase of more than 300% (10, 11). One of the reasons for this increase is that current regulatory practices often demand excessively detailed recording of imprecise symptoms, such as start and stop dates of subjective low-grade adverse events, for example, fatigue. Even so, there are few (if any) instances in which such documentation for grade 1 to 2 toxicity has proven important in deciding whether or not a drug is of value in a given disease (13).

There are limitations to the analysis presented in this article, as it is based on the data available on the FDA website. For instance, a detailed breakdown by phase of trial was not possible. Furthermore, we do not know the absolute trial number per year and, therefore, some percentages cannot be calculated. It is likely that the FDA would exert more oversight of those trials submitted in support of approval of a new indication or a first-in-human trial that has potentially life-threatening adverse effects. This aspect could not be covered in our analysis because the data were not in the public database. Other aspects of the analysis, such as the reason for the increase in “no action indicated” inspections after about 1996, might better be ascertained with direct FDA feedback. Finally, numbers such as those pertaining to the frequency of FDA inspections could be influenced by extraneous factors such as government budgetary constraints.

Although not mandated by the FDA, clinical trial monitors and inspectors have recently adopted a policy in which 100% source document verification is the standard (10). Furthermore, distinctions are blurred between minor data imperfections and systematic trial negligence—in both instances, the investigator is cited for failure to follow the investigational plan. With the rapid increase in study-mandated procedures (9–13), deviations, especially minor ones, from the protocol are virtually guaranteed because of patient, medical, or logistical issues. Yet, the differences between protocol deviations (minor departures) and violations (that could affect patient safety or data integrity) are poorly delineated, and the current regulatory approach frequently classifies any deviation from the protocol as a violation. The complexity of some studies, both sponsor- and investigator-initiated, encourages larger numbers of deviations. It may be easier to ameliorate this problem in investigator-initiated studies. It also seems reasonable and indeed may enhance both the generalizability of the results and ease of trial conduct, if eligibility criteria were streamlined to address only critical inclusion and exclusion
criteria. The number of procedures associated with studies, especially phase I, has climbed steadily, and procedures per patient might benefit from a formal need for justification. In summary, over the last decade, the FDA has performed approximately 350 inspections per year, of which 82% are "data audit" and 18% are "for cause." Not surprisingly, a higher percentage of "for cause" audits find objectionable conditions. Whether this is because there are more underlying problems or because inspectors are predisposed toward finding problems in such an inspection is not clear. There has been a significant increase in inspections in the last decade (2000–2009) as compared with the prior decade (1990–1999), and the percentage of inspections each year that were "for cause" has also significantly increased since 1983, peaking in 2003 at 28.5%. The percentage of inspections that were "official action indicated" did not change with time. The most common deficiency has been failure to follow investigational plan, noted in about 34% of all inspections; this may be due to the increasing complexity of protocols, with a tripling of eligibility criteria and a rapid increase in protocol-mandated procedures. The vast majority of FDA inspections result in findings, with only about 31% of inspections having an outcome of no action. Furthermore, no deficiencies (code 00) were found in only 11.2% of all "data audit" and 5% of all "for cause" inspections. The impact and implications of increasingly complicated and stringent protocol requirements in the context of frequent FDA findings of deficiencies warrant examination.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: S.K. Morgan-Linnell, R. Kurzrock Development of methodology: S.K. Morgan-Linnell, R. Kurzrock Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S.K. Morgan-Linnell, R. Kurzrock Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S.K. Morgan-Linnell, R. Kurzrock Writing, review, and/or revision of the manuscript: S.K. Morgan-Linnell, D.J. Stewart, R. Kurzrock Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S.K. Morgan-Linnell, R. Kurzrock Study supervision: R. Kurzrock.

Received December 4, 2013; revised March 4, 2014; accepted March 7, 2014; published OnlineFirst April 15, 2014.

References
Clinical Cancer Research

U.S. Food and Drug Administration Inspections of Clinical Investigators: Overview of Results from 1977 to 2009

Sonia K. Morgan-Linnell, David J. Stewart and Razelle Kurzrock