Result disclosure on ClinicalTrials.gov — first experiences and challenges

Ralf Minkenberg

Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany

The disclosure of clinical trial 'basic' results on the website ClinicalTrials.gov has been legally required since September 2008 for all FDA-approved trials or cleared drugs and devices. Since that time, many additions (e.g. adverse event reporting in September 2009) and corrections (both content and format related) have been implemented. As most of the information needed for ClinicalTrials.gov is created anyway during the analysis and reporting of a clinical trial, a well-defined process can be established to collect the needed results during the programming of tables and listings for the trial report. In addition, the ongoing changes in the disclosure requirements and the different responsibilities and functions involved in trial disclosure lead to many challenges in defining and establishing such a process. This paper will focus on the process definition and execution from the programmer's perspective. We will also discuss common pain points of interpreting the ClinicalTrials.gov requirements.

Keywords: Results disclosure, ClincialTrials.gov, FDA, EMA, Clinical Study Report

Introduction

With the FDA Amendment Act of 2007, the disclosure of basic results on the ClincialTrials.gov site managed by the National Institutes of Health became a legal requirement. Title VIII, Section 801 stated a 'responsible party' (i.e. the study sponsor or designated principal investigator) register and report results of certain 'applicable clinical trials' that were initiated or ongoing as of 27 September 2007.[1]

'Applicable clinical trials' include interventional studies of drugs, biological products, or devices that are subject to FDA regulation. In general, the submission of the 'basic results' information has to be done not later than 1 year after the date of an intervention for purposes of final data collection for the primary outcome has been performed, regardless of whether the clinical trial was concluded according to the protocol or was terminated.[2]

The ClincialTrials.gov 'basic results' database was launched in September 2008 and two types of results information are to be reported. There is scientific information, organized in four sections (participant flow, baseline characteristics, outcome measures and statistical analyses, and adverse events), as well as administrative information (e.g. a point of contact to obtain additional results information). The results can be reported either via interactive data entry using online forms, or via an automatic upload of results data files. The format of these results data files is defined on the ClinicalTrials.gov sites and it is given in XML format.[3]

The format eligible for upload to Clinicaltrials.gov differs from the formats which are currently used when clinical trial results are presented, either in the Clinical Study Report or in a scientific publication or presentation.

The ClinicalTrials.gov Website

ClinicalTrials.gov organizes information for each registered study as an integrated unit, displaying the study protocol record and, if available, the corresponding results record under different tabs on the same page to facilitate cross-referencing (see also Fig. 1).

At the end of July 2010, there were 93 445 studies registered on ClinicalTrials.gov: 77 134 (83%) were interventional studies, and 15 899 (17%) were observational studies. Today about 300–350 trials per week are being newly registered.

Of these 93 445 studies only 2035 have reported results (end of July 2010). Currently there are about 40 new results from trials per week, and ClinicalTrials.gov anticipates about 150 per week.[4]

The basic results part of ClinicalTrials.gov (Fig. 2) is divided into four main sections, which are described briefly in the following paragraphs:

1. participant flow — a table is displayed including the number of patients who dropped out of the clinical trial and the number of patients excluded from the analysis, if any. Therefore, for every defined milestone during a trial the numbers of

Correspondence to: Ralf Minkenberg, Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany. Email: ralf.minkenberg@boehringer-ingelheim.com

Published by Maney
DOI 10.1179/175709310X12845438122791

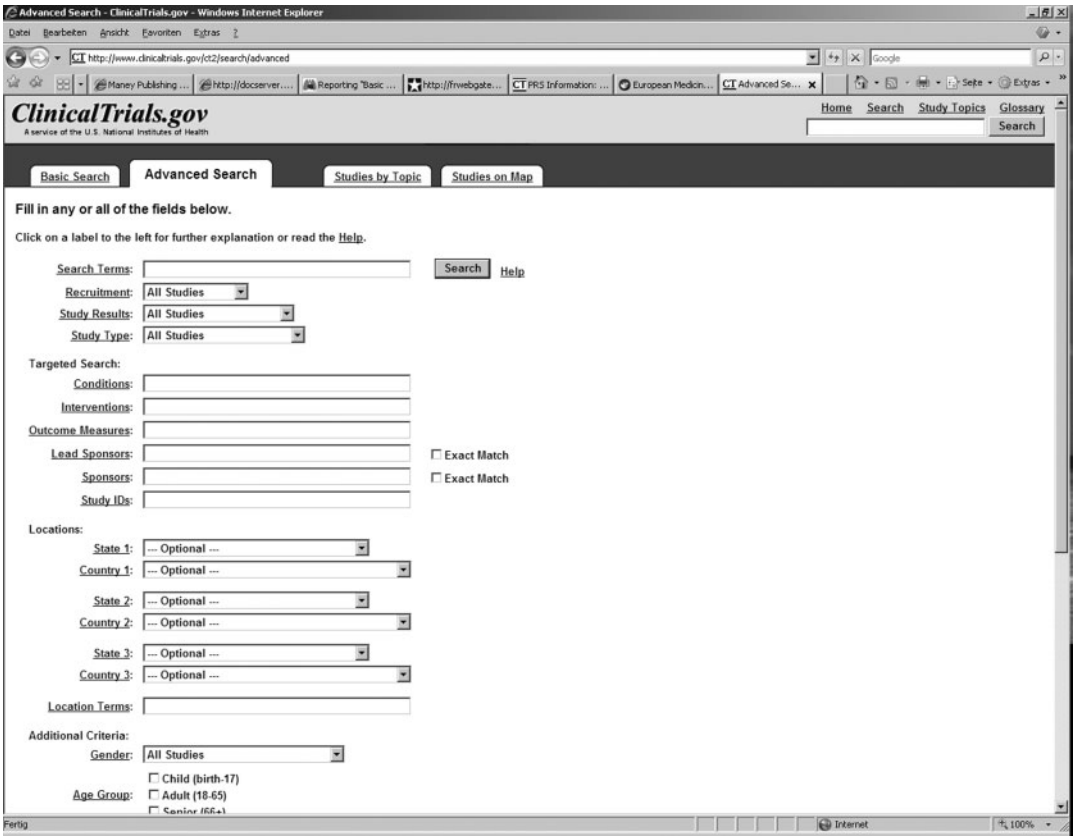


Figure 1 Advanced search interface of ClinicalTrials.gov

patients started, completed, and not completed are required, and the reasons for not completing the trial (or a part of the trial) must be given;

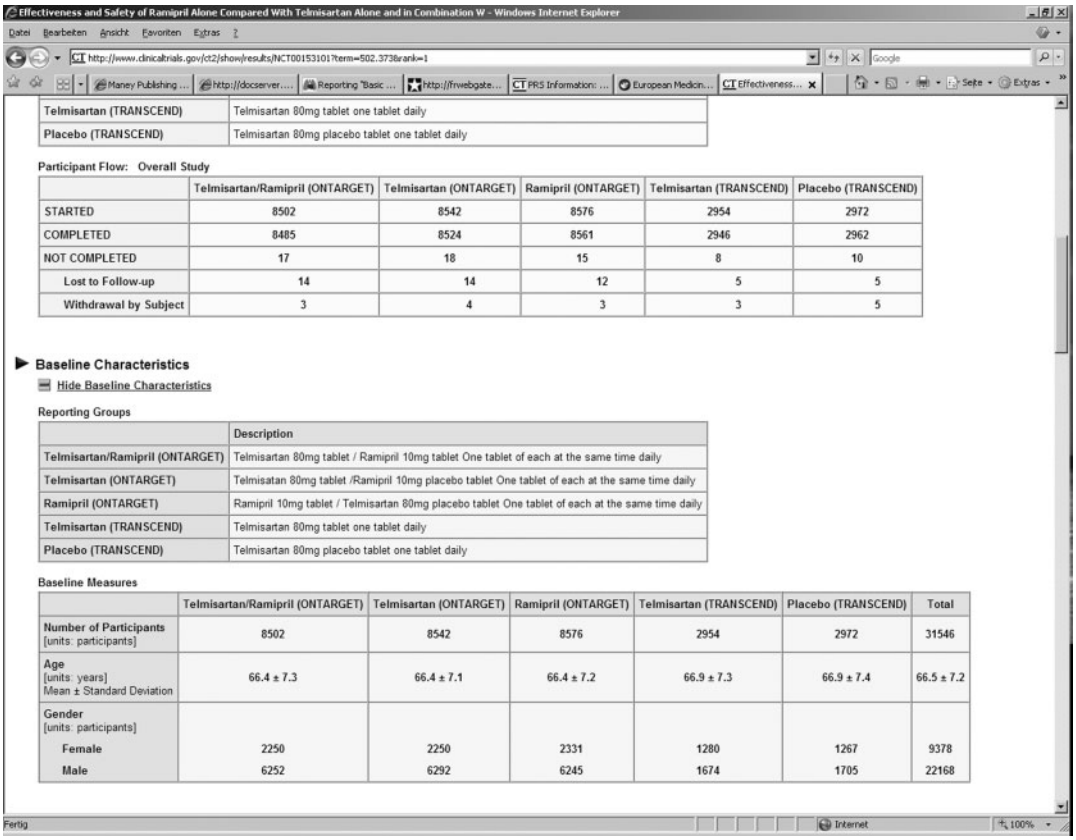
2. baseline measures — tables of the demographic and baseline data collected overall and for each arm of the clinical trial. As a default, certain required measures such as age and gender are necessary. Additional user-specified measures (e.g. baseline measures used in the statistical analyses) could be added;
3. outcome measures and statistical analyses — tables of values for each of the primary and secondary outcome measures for each arm of the clinical trial, including the results of scientifically appropriate tests of the statistical significance of such outcome measures. A measure name and description should be given for every outcome measure. Also a requirement is a time frame and whether or not the outcome measure is a safety issue. In addition to the reporting groups and measured values for the statistical analysis estimate, confidence interval and *P*-values could be given. Additional information about the analysis, such as null hypothesis, power calculation, definition of non-inferiority margin, and other key parameters, could be entered;
4. serious and frequent adverse events — a table of anticipated and unanticipated serious adverse events grouped by organ system, with number and frequency of such event in each arm of the clinical trial. Additionally, a table of anticipated and unanticipated adverse events that are not included in the serious adverse event table, which exceed a frequency of 5% within any arm of the clinical trial, is grouped by organ system, with number and frequency of such event in each arm of the clinical trial. For both adverse event sections, the total number of affected participants and for every preferred term in every organ system the number of participants at risk, the number of events and the number of participants affected are required.

A web-based Protocol Registration System on the website ClinicalTrials.gov enables the user to enter results directly via a form based interface. Entries are checked and incorrect values will be highlighted. This kind of manual data entry is only feasible for a very small number of trials with a reasonable number of necessary entries.

Automatic Upload and Its Challenges

Data on ClincialTrials.gov are saved in XML and is, as such, possible to upload the complete result file to ClincialTrials.gov instead of entering the data manually. The contents of this XML file are defined by a data element definition, which defines and explains every possible entry of the result database. Currently this is still in draft. A production version of the XML schema for the results portion is available. Generally, in these results schema, the tag names and comments within the document relate to fields on the data entry screens and have links to the data element definitions document for full definitions. This schema can be changed without notice from the operators whenever necessary.[5]

Changes of the contents have happened and will continue to do so in the future. The adverse event



Effectiveness and Safety of Ramipril Alone Compared With Telmisartan Alone and in Combination W - Windows Internet Explorer

Telmisartan (TRANSCEND)	Telmisartan 80mg tablet one tablet daily
Placebo (TRANSCEND)	Telmisartan 80mg placebo tablet one tablet daily

Participant Flow: Overall Study

	Telmisartan/Ramipril (ONTARGET)	Telmisartan (ONTARGET)	Ramipril (ONTARGET)	Telmisartan (TRANSCEND)	Placebo (TRANSCEND)
STARTED	8502	8542	8576	2954	2972
COMPLETED	8485	8524	8561	2946	2962
NOT COMPLETED	17	18	15	8	10
Lost to Follow-up	14	14	12	5	5
Withdrawal by Subject	3	4	3	3	5

Baseline Characteristics

Hide Baseline Characteristics

Reporting Groups

	Description
Telmisartan/Ramipril (ONTARGET)	Telmisartan 80mg tablet / Ramipril 10mg tablet One tablet of each at the same time daily
Telmisartan (ONTARGET)	Telmisatan 80mg tablet /Ramipril 10mg placebo tablet One tablet of each at the same time daily
Ramipril (ONTARGET)	Ramipril 10mg tablet / Telmisartan 80mg placebo tablet One tablet of each at the same time daily
Telmisartan (TRANSCEND)	Telmisartan 80mg tablet one tablet daily
Placebo (TRANSCEND)	Telmisartan 80mg placebo tablet one tablet daily

Baseline Measures

	Telmisartan/Ramipril (ONTARGET)	Telmisartan (ONTARGET)	Ramipril (ONTARGET)	Telmisartan (TRANSCEND)	Placebo (TRANSCEND)	Total
Number of Participants [units: participants]	8502	8542	8576	2954	2972	31546
Age [units: years] Mean ± Standard Deviation	66.4 ± 7.3	66.4 ± 7.1	66.4 ± 7.2	66.9 ± 7.3	66.9 ± 7.4	66.5 ± 7.2
Gender [units: participants]						
Female	2250	2250	2331	1280	1267	9378
Male	6252	6292	6245	1674	1705	22168

Figure 2 Example results (participant flow, baseline characteristics) of a study on ClinicalTrials.gov

reporting only became mandatory after September 2009; about every 6 months changes and item additions are introduced. An expansion of results by rulemaking is foreseen in September 2010. As it is planned by the European Medicines Agency (EMA) to require also results disclosure on ClinicalTrials.gov, more changes will be needed.

Collecting and Uploading Results

In order to promptly comply with the regulation, Boehringer Ingelheim introduced a mainly manually driven process. After identifying the trials, for which results had to be published, the trial team will be contacted and an Excel-based template will be delivered. The trial team will have to fill in this Excel template by a pre-specified deadline. After a first formal quality check, the information from the Excel template will be entered into the preview web form from ClinicalTrials.gov. Finally this file will be reviewed, and then uploaded to ClinicalTrials.gov.

This process was only introduced in order to have a quick — and easy to implement — approach to adhere to the new legal requirements (Fig. 3). The process was very labor- and people-intensive because all of the steps during identifying, collecting, and uploading the necessary information involved many manual actions, which had also to be checked and reviewed regularly. Because manual data entry causes errors very often, most of those involved in the process had to deal with the same study more than once to correct errors and validate entries again.

Shortly afterwards the development of a new process was initiated, which avoided manual interactions as much as possible. This new process will be integrated in the currently established processes of initiation, the conduct, analysis, reporting, and publishing of a trial.

The newly developed process to collect, save, and upload results information to ClinicalTrials.gov assigns the main responsibilities for disclosure data to the relevant study statistician and programmer. Final

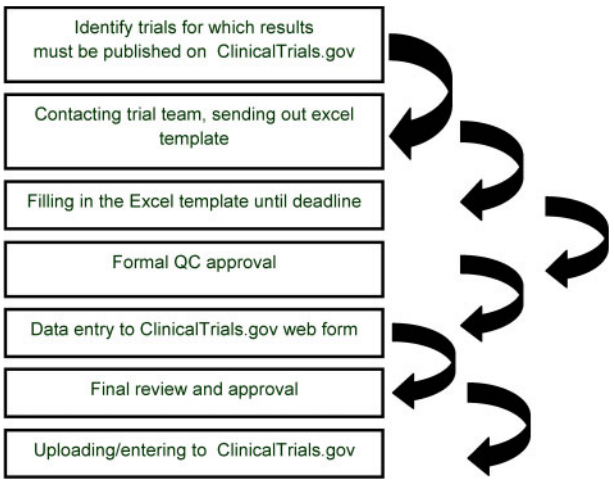


Figure 3 First (temporary process): description for collecting and entering data to ClinicalTrials.gov

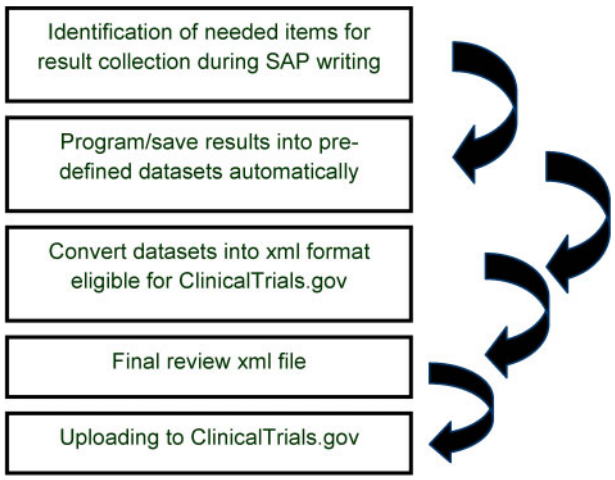


Figure 4 Final planned process description for collecting and entering data to ClinicalTrials.gov

approval is still given by the appropriate — a medical professional. The basic idea is to save the information needed for ClinicalTrials.gov at a point during a clinical trial, at which these data are produced anyway. This would be when the tables for the Clinical Study Report are created. The programmers could then save the relevant information when the necessary tables for the Study Report are produced. Once this is achieved, no manual entry of study results would be needed, and the correctness of the result data could be established during the general program validation of study programs. The underlying specifications, for which tables of the study report are needed for results disclosure, and which items from these tables are needed, are defined by the study statistician (with support of other involved functions, if necessary).

As described in Fig. 4, after identification of a trial eligible for ClinicalTrials.gov disclosure, the first step is to identify and specify all needed items for the result collection. This task takes place during the writing of the Statistical Analysis Plan (SAP), in which the detailed structure, order, and layout of all the tables to be included in the Clinical Study Report, will be defined. It is important to define uniquely which information from which table is needed for disclosure. Based on these specifications, the results needed for disclosure will be written to a pre-defined dataset, where all needed disclosure data will be accumulated. The writing of this Disclosure Data Set (DDS) could be inserted into the standard programs used for table programming. It will be done by calling a validated SAS® macro which ensures that all the information in the DDS is identical to the corresponding information in the study report table, if the parameters given in the macro call are correct. For validation a careful code review of the macro call is essential. After finalization of the programs for the study report tables, the DDS is also ready. Manual interaction in writing data into this dataset must not happen. The final output dataset for ClinicalTrials.gov will always be available at the same time as the final tables for the Clinical Study Report are finalized. The final DDS will then be converted into the XML format required for ClinicalTrials.gov. For this step, an external program is used, but it is planned to do this also with SAS in the future.

Responsibilities and Challenges during Results Collection

The process of results collection for disclosure on ClinicalTrials.gov is fully embedded in the well established process of writing the SAP, and programming of the tables for the Clinical Study Report. The main responsibilities are taken by statisticians and programmers; support is needed from the Medical. A first important step is the definition and specification of tables (and items in these tables) for disclosure information. Some results, which should be reported on ClinicalTrials.gov, are not generally part of the report tables, and were not displayed in tables of the study report in the past, e.g. non-serious adverse events were not displayed alone (excluding serious adverse events) as required by ClinicalTrials.gov. Therefore, the standard display templates for the Clinical Study Report have been adapted to ensure that all of the information which must be disclosed can also be found in a study report table.

The specification of the items needed for disclosure is done in an easy, but also unique way. The additional specifications during writing the SAP by the study statistician should not add too much additional workload. At the same time, the specifications are defined in such a way that the programmer easily knows what is needed for the disclosure dataset. It is feasible to define the needed items by marking them in a mock table template (if it is obvious which marked item corresponds to which ClinicalTrials.gov entry). Another possibility is to define the needed items by well-defined row and column labels used in the table anyway. The important issue for the specification step is that statistician and programmer find a modus operandi which both can easily understand and operate.

During programming Clinical Study Report table production, all data needed for disclosure on ClinicalTrials.gov are saved in a DDS. This dataset has a structure very similar to the final xml structure required by ClinicalTrials.gov. In addition to an identification variable, the dataset consists of one variable for text entries (either recommended text or free descriptive text) and one (character) variable for each treatment investigated in the study. Using a validated SAS macro, the needed items from a table can be copied into the DDS. This should occur directly before the study table is written to the output

file in order to ensure that exactly the same data are used for the table and the DDS. The programmer only has to take care of the correct usage of the macro call according to the specifications given by the statistician. No review of the real numbers (or texts) saved in the DDS is mandatory, but exemplary checks of the saved can just be done. The macro calls are easy to implement into table creation programs so the additional workload for the programmer is minimal.

Because the disclosure dataset is created during the general study report table creation, there is always a final version of the DDS available at the time when the tables for the Clinical Study Report are final. Draft versions of the DDS (and therefore of the disclosure information) are also available before, e.g. when draft tables (before or after unblinding of the data) are presented. In this way, it is easily possible also for Medical to look at the disclosure data at an early stage and they can understand and review them easily. It is obvious that the whole process is integrated in current standardized processes as much as possible. The additional workload, both for the statistician and the programmer, is low, especially in those cases where standardized table templates are used to create the study report tables. For these table templates, there are also templates/examples provided as to how a corresponding DDS specification should look like. Therefore, the process for collecting disclosure information encourages also all involved functions to adhere to existing standards.

Challenges

The first experiences with the new process of collecting disclosure information for ClinicalTrials.gov were positive. If standard table templates and programs were used, the feedback from statisticians, programmers, and medical people underlined the early creation of disclosure information and the successful integration of the new disclosure requirements into the current processes.

But there are nevertheless challenges for the involved functions. Mostly questions or problems occur if the study design or the study outcome measurements cannot be reported in the basic format that ClinicalTrials.gov allows.

Participant Flow

The participant flow accommodates a range of study designs and allows for the description of key events following study enrolment, but before group assignment. At a minimum, the total number of participants starting and completing the overall study must be reported for each study arm. If milestones are defined, the number of participants achieving each milestone must be less than or equal to the number starting the period and achieving previous milestones.[6] Sometimes, even in complex study designs, it is difficult to comply with these rules and additional discussion as to how the study flow could be disclosed in a correct, compliant, and understandable way is necessary.

Outcomes

It is legally required to report data for all prespecified primary and secondary outcomes. Precise descriptions of outcome measures are considered as necessary to interpret the data. At a minimum, the following outcome information is required: an informative outcome measure title (what was measured); a specific description (how it was measured); the time frame for assessment (when it was measured); measure type (e.g. number, median); a measure of dispersion or precision for continuous measures (e.g. SD, 95% confidence interval, interquartile range); and the unit of measure (what the reported values represent). For outcome measures, it is very important to describe the measurements very clearly; the usage of any abbreviations is only allowed if they have been defined before.[6] A lack of precision in the wording, inconsistencies in units and numbers, and insufficient details are the most common problems with entries on ClinicalTrials.gov. Because of the involved functions, statistician, programmer, and medical people, such problems generally could be solved in early discussions.

Ad Hoc Updates

Another issue is the regular updates to the ClinicalTrials.gov website. More often additions and changes are implemented, which directly affects the structure of the XML file for upload to ClinicalTrials.gov and, therefore, also affects the specification, structure, and creation of the DDS, e.g. in September 2009 new elements for the adverse events disclosure were introduced; starting that January 2010 confidence intervals could be specified as one-sided or two-sided. Any additions or changes have to be implemented first into the structure of the DDS, so that the SAS macro for copying data into the DDS could use this new item. Then it should be discussed and defined, if and how the addition would have an influence on the current planned disclosure data. Finally, changes must be entered and submitted. Changes by rulemaking are foreseen for September 2010 and the EMA will also require results disclosure likely 2011. Therefore, major updates could happen to ClinicalTrials.gov, which would affect the established process.

Conclusions

- As results disclosure is now legally required for a wide range of clinical studies a standardized process was developed to handle the corresponding requirements.

- Manual data entry must be avoided as much as possible, because it is very error-prone and additional validation steps are needed.
- The newly established process is fully integrated in the existing workflow for creating study report tables. Responsibilities are mainly assumed by the statistician and programmer with the support of the medical people.
- The final disclosure information is prepared at the same time as the study report tables are finalized. The more standardized tables and programs are used, the less is the additional workload for statistician and programmer.

The interpretation of the ClinicalTrials.gov requirements and the regular updates of the website are the major challenges during defining and collecting disclosure information. The more studies are disclosed and the more stable the current version of ClinicalTrials.gov is, the easier it will be to disclose the required results data without a huge additional workload.

Acknowledgements

The author expresses sincere thanks to the following individuals for their contributions to this article: Erika Richter, Paul Kearney, and Oktay Yildirim (all Boehringer Ingelheim).

References

1 US Food and Drug Administration. FDAAA, US Public Law No. 110-85 § 801 [document on the Internet]. Silver Spring, MD: US Food and Drug Administration; 2007 [cited 2010 Jul 30]. Available from: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ085.110.pdf

2 US National Library of Medicine. ClincialTrials.gov — about the ClinicalTrials.gov database [document on the Internet]. Bethesda, MD: US National Library of Medicine [updated 2009 Aug 4; cited 2010 Jul 30]. Available from: http://www.clinicaltrials.gov/ct2/info/results

3 US National Library of Medicine. ClincialTrials.gov protocol registration system — PRS and U.S. Public Law 110-85 [document on the Internet]. Bethesda, MD: US National Library of Medicine [updated 2010 May 17; cited 2010 Jul 30]. Available from: http://prsinfo.clinicaltrials.gov/fdaaa.html

4 US National Library of Medicine. ClinicalTrials.gov — advanced search [document on the Internet]. Bethesda, MD: US National Library of Medicine [cited 2010 Jul 30]. Available from: http://www.clinicaltrials.gov/ct2/search/advanced

5 US National Library of Medicine. ClinicalTrials.gov 'basic results' data element definitions (DRAFT) [document on the Internet]. Bethesda, MD: US National Library of Medicine; February 2010 [cited 2010 Jul 30]. Available from: http://prsinfo.clinicaltrials.gov/results_definitions.html

6 Tse T, Williams RT, Zarin DA. Reporting 'basic results' in ClinicalTrials.gov. Chest 2009; **136**: 295–303.