January 2018


Aliya Cecilia Roginiel

Follow this and additional works at: https://elischolar.library.yale.edu/ymtdl

Recommended Citation
https://elischolar.library.yale.edu/ymtdl/3526

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

by
Aliya C. Roginiel

2019
EVIDENCE SUPPORTING FDA APPROVAL AND CMS NATIONAL COVERAGE DETERMINATIONS FOR NOVEL MEDICAL PRODUCTS, 2005 THROUGH 2016: A CROSS-SECTIONAL STUDY

Aliya C. Roginiel, Sanket S. Dhruva, Joseph S. Ross
Department of Internal Medicine, Yale University School of Medicine, New Haven, CT

Abstract:

Food and Drug Administration (FDA) and Centers for Medicare and Medicaid Services (CMS) rely on evidence from clinical trials when approving a therapeutic for marketing and insurance coverage in the US, respectively. No study has compared the quality and quantity of evidence examined by these agencies. Our purpose was to characterize evidence used by FDA and CMS to support marketing approval and coverage of novel therapeutics reviewed for CMS coverage from 2005 through 2016. We conducted a cross-sectional study of clinical trials described in FDA approval documents and CMS NCD memoranda. We compared the number of clinical trials used by each agency as well as the following characteristics of clinical trials: study size, randomization, double-blinding, and control arm. Twelve medical products met our inclusion criteria. FDA approvals were based on 22 pivotal trials. CMS NCDs were based on 27 original clinical trials; 14 clinical trials were used by both agencies. Between FDA pivotal and CMS original clinical trials, there was no significant difference in study size (P=.53), use of randomization (P=.75), double-blinding (P=.55), or control arm (P=.54). There was no statistically significant difference in median age between participants in trials reviewed by CMS versus those reviewed by FDA (62 vs 59 years, P=.26). The median time from FDA approval to publication of CMS NCD memorandum was 17 (interquartile range, 13-36) months. FDA approvals and CMS NCDs are based on a similar number and quality of trials, although trial participants are not reflective of the Medicare population, and the process of finalizing coverage determinations requires an additional 17 months.
Acknowledgements

First, I would like to express my sincere gratitude for the mentorship of my thesis advisor, Dr. Joseph S. Ross. He is one of the best physician-mentors whom I have had the privilege to work with; not only did he devote innumerable hours to helping me on this project, but he also inspired me to have a deep appreciation of the intersection between policy and public health. While rounding with Dr. Ross and his team during his time on service, I learned that he is also a phenomenal clinical teacher and a genuinely compassionate and highly skilled physician. He is the role model every medical student should follow. I would also like to thank Dr. Sanket Dhruva for lending his immense expertise and guidance on Medicare coverage policy to our work. His input was essential in finalizing our manuscript, which was published in the journal *Medicine* in October of 2018.

In addition, I want to thank Dean Angoff, Dean Belitsky, Dean Schwartz, Mr. Silverman, and the George Washington School of Medicine for making my transfer to the Yale School of Medicine possible. Without this transfer after my second year, I would not have had the tremendous support of my family during the majority of my medical studies. Of course, a great thank you to Dr. Forest and the Yale Office of Student Research for their support of this project through a Yale Student Research grant under NIH training award T35DK104689.

Finally, I want to thank my mother, Helen Roginie, and my husband, Alexander Svoronos, for their unending encouragement and positivity during this time. In particular, without my mother’s support, I never would have accomplished my dream of becoming a physician. I would also like to thank my late father, Dr. Ali Khodadoust, who despite many setbacks, encouraged me to work hard, to enjoy the journey, and accomplish the seemingly impossible.
# Table of Contents

## I. Introduction

- a. Overview of the FDA Approval Process
  - i. Medical devices
  - ii. Drugs
  - iii. Biologics
- b. Overview of the CMS Coverage Process
- c. Review of Relevant Policy Literature
- d. Aligning FDA and CMS Requirements through Parallel Review
- e. Statement of Purpose

## II. Methods

- a. Identifying CMS NCDs and FDA Approval Documents
- b. Selection of Final Study Sample
- c. Identification of Original Clinical Trials Review by CMS
- d. Identification of FDA Pivotal Efficacy Trials
- e. Data Extraction
- f. Statistical Analysis
  - i. Characterization of NCDs
  - ii. Characterization of trials evaluated by FDA and CMS

## III. Results

- a. Characteristics of NCDs
- b. Number of Trials Evaluated by the FDA and CMS
- c. Characteristics of Trials Evaluated by the FDA and CMS
- d. Time from FDA Approval to Initiation and Publication of a CMS NCD
IV. Discussion

a. Summary of Findings

b. Study Limitations

c. Implications for Regulatory Policy

d. Conclusions

V. References
INTRODUCTION

Marketing approval for medical devices, drugs, and biologics in the United States is granted by the Food and Drug Administration (FDA). Following approval by the FDA, a product requires third-party payer coverage for clinical adoption; in the US, the largest third-party payer is the Centers for Medicare and Medicaid Services (CMS), with Medicare alone covering 56.8 million Americans. (1) Both the FDA and the CMS rely on evidence from clinical trials when making their approval and coverage decisions, respectively. However, each agency has a specific goal in mind. The FDA must evaluate evidence to determine whether a product is “safe and effective”, whereas the CMS must use evidence to determine whether a product or procedure is “reasonable and necessary for the diagnosis or treatment of an illness or injury”. (2-4) Despite these overarching differences in the purpose of evidence review, we do not have a formal comparison on the type and quality of evidence reviewed by the FDA and the CMS to support their decisions. Characterizing clinical trial evidence used by the two agencies is therefore important for understanding how product approval and payer coverage processes are linked, and ultimately how millions of patients are able to receive beneficial medical treatment.

Overview of the FDA Approval Process

The FDA, an agency within the U.S. Department of Health and Human Services, is responsible for protecting the public health by assuring the safety, quality, and security of human and veterinary drugs, vaccines and other biological products, and medical devices. (5) The FDA is also responsible for the safety and security of products such as cosmetics, dietary supplements and most of the U.S. food supply. (5) The agency is organized into four directorates and the Office of the Commissioner. (6) The four directorates include: Foods and Veterinary Medicine, Medical Products and Tobacco, Global Regulatory Operations, and Operations. (6) Within
Medical Products and Tobacco, the following divisions are responsible for approval of medical devices, drugs, and biologics, respectively: Center for Devices and Radiological Health (CDRH), Center for Drug Evaluation and Research (CDER), and Center for Biologics Evaluation and Research (CBER). Details on the approval process for each type of medical product are outlined below.

**Medical Devices**

In 1976, Congress granted the FDA the authority to regulate all medical devices. This authority was bestowed following the Dalkon Shield intrauterine device controversy, whereby use of the device led to hundreds of septic spontaneous abortions and multiple maternal deaths. With this law, the premarket approval (PMA) application was instituted. The PMA application required clinical testing with scientific evidence that provided “reasonable assurance that they device [was] safe and effective for its intended use…in the target population.” Today, similar to initial requirements, the PMA application is reserved for class III medical devices. Class III devices are those products considered to carry the highest risk and that perform life sustaining or supporting functions and include prosthetic heart valves and coronary stents. In brief, there are four steps necessary in order to review a PMA application: (1) an administrative and limited scientific review by FDA staff to determine completeness of the application; (2) an in-depth scientific, regulatory and quality system review; (3) a review and recommendation by the appropriate advisory committee; (4) final deliberations, documentation and notification the FDA decision. Valid scientific evidence must be presented, mostly including randomized controlled trials and single-arm studies with historical controls; although case histories are sometimes included in the review. Generally, the FDA prioritizes review of pivotal trials, or those clinical studies that “[gather] the final data on safety and effectiveness
needed for FDA review.” (12)

The FDA is required to complete its review of a PMA within 180 days of the manufacturer filing their application. (13) Once a PMA is approved, a summary of safety and effectiveness data (SSED) is available on-line to the public. (14) The goal of the SSED is to “summarize the key content of the PMA, such as Device Description, Preclinical Evidence, and Clinical Evidence, as well as FDA’s analysis of the scientific evidence that served as the basis for FDA’s [approval or non-approval] decision.” (15)

Following class III devices are moderate-risk or class II medical devices, with an intended use and safety profile similar to a predicate device. (10) Class II products, which in fact represent the majority of new medical devices, are approved through the 510(k) pathway. (10) Additional clinical data are usually not required if manufacturers can show “substantial equivalence” to a previously approved device. (16) Examples of moderate-risk devices include coronary guidewires, and contrast auto-injection systems. (10) The FDA provides a 510(k) summary on-line for class II medical devices, which includes information on the intended use of the product, and the technological characteristics of the device compared to a predicate device, along with additional legality data. (17)

Initially, there were numerous differences between the PMA and 510(k) approval process, including “limited authority for the FDA to withdraw clearance of a [moderate-risk] device that is found to be unsafe or ineffective” and lack of premarket inspections into device manufacturing. (8) Controversy arose when high-risk medical devices were approved using the 510(k) pathway, without reassurance that clinical trial data supported safety and efficacy of the product. (8) In response, the FDA issued a report in 2010 stating that products in the 510(k) pathway should require data from clinical trials and that they should be subject to premarket
inspections and postmarket studies. Since the report, changes have been made to increase postmarket surveillance and premarket inspection of moderate-risk devices; life-saving and sustaining products (i.e. high-risk devices) are subject to the specific PMA review process. Lastly, class I devices (e.g. adhesive bandages and hearing aides) are considered low-risk and are “non-life sustaining with [a] long history of safety and effectiveness”. Most class I devices do not require approval through the 510(k) pathway.

Drugs

The CDER is responsible for approving over-the-counter (OTC) and prescription drugs. If a manufacturer wants to submit a new drug for approval, they must submit a new drug application (NDA), which must support the drug’s safety and effectiveness for its intended use. In addition, manufacturers are required to report results of all clinical tests, the specific ingredients of the drug, results of animal studies, how the drug behaves physiologically, how it is manufactured and how it is processed and packaged. Before an NDA is submitted however, the drug must go through preclinical phase testing. The goal of preclinical phase testing is to use an animal model to achieve the desired clinical effect and ultimately predict outcomes of the treatment in humans. A second goal of preclinical studies is to identify and characterize toxicities of the drug to predict adverse events in patients. At this point, the manufacturer submits an investigational new drug (IND) application to the FDA and follows up with clinical trials.

In general, clinical studies are classified as either Phase I, Phase II or Phase III. Phase I trials provide information on a drug’s pharmacokinetic profile, side effects, tolerability, toxicity and dosing schedule; both Phase II and Phase III trials are designed to study the risks of benefits
of the drug and are much larger studies than Phase I trials. (23) Similar to the review process for medical devices, the FDA classifies certain studies as pivotal trials that are often used as the basis for the FDA’s approval decision. (3) In general, the FDA requires “adequate and well-controlled investigations” and recommends that manufacturers submit data from at least two clinical trials. (3) Finally, an NDA is submitted to the FDA that contains results from all preclinical and clinical testing. (21) In general, the FDA is given a time period of 12-months to review an NDA and 6-months if an application falls under priority review (e.g. drugs that are considered to have significant advances over current therapies). (21) Once approved, drug approval summaries are made available on-line by the FDA. (24) These summaries contain comprehensive information on topics ranging from drug labeling, medical reviews, and pharmacology reviews. (24) However, it is important to note that any clinical trial data considered confidential or commercial may not be included for public review. (23)

**Biologics**

Biologics are “active pharmaceutical ingredients derived from living organisms that cannot reasonably be synthesized by chemical means.” (25) Insulin, the first biologic medicine, was originally derived from extraction of the hormone from pancreatic tissue of cattle and pigs. (25) Later, insulin was produced using a recombinant DNA technology. (25) In brief, recombinant DNA technology involves synthesizing DNA fragments that contain a particular gene, inserting them into vectors, and inserting those vectors into a host organism such as a bacterium. (26) The hosts are then grown to produce multiple clones of the DNA fragment. (26) Today, there are several distinct classes of biologics including monoclonal antibodies (mAbs), receptor modulators, and enzyme modulators. (27)
Monoclonal antibodies are monovalent antibodies that bind to the same antigen-binding site and are manufactured from a single clone of B-lymphocytes. (28) Examples include rituximab used for non-hodgkin’s lymphoma and infliximab for rheumatoid arthritis. (29, 30) Receptor modulators, such as tamoxifen for the prevention and treatment of breast cancer, bind to and affect a variety of receptor subtypes. (31) Enzyme modulators either induce or inhibit enzyme activity and include the drug alteplase, a potent thrombolytic. (27)

In order for a manufacturer to market their biologic, they must have an approved investigational new drug application, followed by adequate clinical trials and eventual submission of an NDA or Biologics License Application (BLA). (32) Like applications for devices and drugs, the FDA generally bases its approval decision on a few key pivotal trials. Following FDA approval, third-party payers such as CMS conduct their own review to determine whether a product should receive coverage.

**Overview of the Medicare Coverage Process**

The majority of Medicare coverage decisions are made by contracted insurers at the state or region-wide level. (33, 34) These local insurers issue local medical review policies (LMRPs) that are only effective within their coverage area. (33) However, the central CMS agency will issue between ten to fifteen national coverage decisions (NCDs) per year for those products that are deemed particularly controversial or are projected to have a major impact on the Medicare program. (34) NCDs can be requested either by an outside organization or by CMS directly. CMS will generally make a request for an NCD under the circumstances such as: conflicting local coverage decisions, new evidence of a technology that promises large clinical impacts, or if there is significant uncertainty about a product’s risks and benefits. (34)
The content of an NCD includes the final decision of CMS along with conditions for coverage if approved, background of the condition the product is intended to treat, history of Medicare coverage, timeline of activities leading to the coverage decision, FDA status of the product, a review of evidence used to support the final coverage decision, links to an external technology assessment if commissioned, results of an internal technology assessment, literature review, whether CMS requested a Medicare Evidence Development and Coverage Advisory Committee (MEDCAC) meeting, and results from that meeting if applicable.(35) Clinical guidelines, public comments, and professional society position statements may also be included.(35) In particular, MEDCAC statements and external technology assessments are unique aspects of the national CMS coverage review process.

The purpose of the MEDCAC is to “provide external assistance in judging the strength of evidence for NCDs”.(34) Their recommendations are used as a supplement and are often based on a review of medical literature, technology assessments, and public testimony.(36) The CMS selects up to 100 committee members from a wide range of backgrounds including patient advocacy and clinical and administrative medicine to serve on the MEDCAC.(36) For each meeting, a maximum of fifteen members with relevant experience are expected to provide recommendations on a specific product.(36) The CMS may also request an external technology assessment when, among other reasons, the total quantity of evidence to review is too extensive, the review requires unique technical or clinical expertise, or when the topic was referred to MEDCAC for additional recommendations.(37) The CMS will generally contract with the Agency for Healthcare Research and Quality (AHRQ) for an external technology assessment and the completed report includes a summary of the literature search strategy, along with the final technology assessment.(37)
NCDs can apply to medical devices, surgical procedures, diagnostic imaging, drugs, and/or drug delivery systems. Furthermore, the final coverage decision can be one of the following: full coverage, coverage with restrictions, referral to local contractors, or noncoverage; once a decision is made, all local Medicare contractors are bound to that particular NCD.

**Review of Relevant Policy Literature**

There are several pertinent studies that have evaluated the strength of evidence reviewed by FDA and CMS independently. Downing et al. reviewed clinical trial evidence used to support FDA approval of novel therapeutics. Similar to our study, they characterized pivotal trials used to support the final approval decision and evaluated implementation of randomization and blinding, features that are associated with higher-quality studies. They also studied the type of comparator, clinical trial endpoint, trial duration and completion rate. The quality of evidence used to support FDA approval varied widely across several indications; specifically, the choice of comparator and trial end point as well as the size and duration of trials varied greatly. However, approval for each product indication was supported by at least one randomized controlled double-blinded trial.

A study by Rathi et al. also evaluated clinical trial evidence supporting FDA approval of medical products, with a focus on otolaryngologic prescription drugs. Similar to Downing’s study, the authors found that most product indications were supported by at least one randomized controlled study. In addition, they found that almost half of indications were approved using studies that included a surrogate marker (e.g. “a biomarker expected to predict clinical benefit, such as progression-free survival”) as a primary endpoint.

Rathi et al. also studied clinical trial evidence used to support FDA approval of high-risk
Otolaryngology medical devices. Results showed that fewer than half of trials used to support medical device approval were randomized, blinded or controlled. Authors mentioned that in prior studies of cardiovascular, orthopedic, and neurosurgical devices, those products approved on the basis of limited evidence were more likely to be unsafe. Another study by Hwang et al. evaluated pivotal clinical trials used to approve novel ophthalmic drugs and medical devices. Over 90% of ophthalmic drugs were approved based on randomized and blinded trials; however, fewer trials used to support ophthalmic device approval were randomized compared to trials for ophthalmic drugs. Several studies have also explored clinical evidence used to support Medicare NCDs; research published by Neumann et al. is particularly relevant.

In 2008, Neumann et al. published an analysis of Medicare NCDs issued from 1999 through 2007. For most NCDs, CMS considered the available supporting evidence to be either of fair or poor quality with higher strength of evidence for drugs than for other products, similar to results from Hwang et al. Evidence was classified as “good” if trials were well-designed and were conducted in representative populations. Authors also studied mean CMS review time, with an average duration of 243±65 days from the date when a review was initiated by CMS to the date the final decision memorandum was published. Results from a prior study by Neumann et al. examining NCDs published from 1998 to 2003 yielded similar conclusions; most often, technologies demonstrated a moderate net-benefit, but their evidence base was generally considered fair. In some cases, there was insufficient evidence to determine whether there was a net-benefit. However, in both studies, products with good evidentiary support were more likely to be covered.

Dhruva et al. focused on the generalizability of clinical data used to support NCDs to the Medicare beneficiary population. Their meta-analysis of over 40,000 patients from 141 trials
showed that characteristics of trial patients differed markedly from the Medicare population; trial participants were more likely to be younger, male, and residents outside the U.S.\(^\text{(14)}\) This is particularly concerning as the Medicare population is mostly female, and most are above 65 years of age (with the exception of qualifying patients <65 years of age); reported outcomes and adverse events of such high-risk cardiovascular devices may therefore not translate to the intended population.\(^\text{(14)}\) Many of the findings from these aforementioned studies were the inspiration for this thesis research.

**Aligning FDA and CMS Requirements through Parallel Review**

The FDA-CMS parallel review program was initially proposed in 2010 as a way to streamline the process from FDA approval to approval for CMS coverage.\(^\text{(46)}\) In other words, product manufacturers who are planning to seek Medicare coverage for their technology could ask CMS to start the review time in conjunction with later stages of FDA review.\(^\text{(46)}\) Furthermore, parallel review attempts to reconcile the goals of evidence review by the FDA and the CMS.\(^\text{(46)}\) For example, clinical trials geared towards FDA approval are often based on rigid protocols with “narrow patient populations” that are designed to demonstrate a product’s safety and efficacy.\(^\text{(46)}\) However, these populations are often not reflective of real-world patients. On the other hand, the CMS requires evidence that will help them determine whether or not the product will be reasonable and necessary for the broader population of Medicare beneficiaries.\(^\text{(46)}\) Therefore, the parallel review program has the potential to both fast-track the adoption of beneficial technologies for Medicare beneficiaries, and to significantly improve the evidence base upon which approval and coverage decisions are made.

One such success example of the parallel review program is the diagnostic test
Cologuard. In brief, cologuard is used as a stool DNA screening test to detect biomarkers associated with colon cancer.(47) In a multi-center cross-sectional study, cologuard detected 69.2% of advanced pre-cancerous lesions compared with 46.2% lesion detection with the fecal occult blood test FIT.(48) Through the parallel review pathway, time from PMA submission to the FDA to publication of the final CMS NCD was 489 days compared with 612 days for similar products going through the standard FDA approval and CMS coverage process.(47) Therefore, coverage and reimbursement for Cologuard was available much sooner for Medicare beneficiaries than it would have been without the parallel review program. Parallel review was eventually made permanent in 2016, although it is currently restricted to medical devices.(49)

STATEMENT OF PURPOSE

The purpose of this study was to address a major gap in our understanding of how the FDA and the CMS, two of the most influential regulatory agencies for healthcare in the U.S., compare in their use of evidence to approve and cover medical products that will ultimately be used by millions of Medicare beneficiaries. To achieve this, we characterized evidence used to support both FDA approval decisions and CMS coverage decisions for novel medical products, including medical devices, drugs, and biologics. More specifically, we analyzed the proportion of clinical trials that had the following characteristics associated with good-quality evidence: use of randomization, blinding, inclusion of a control group, inclusion of either a placebo or active comparator, and use of a clinical outcome as a primary measure of efficacy. Then we compared clinical trial characteristics between the two agencies to evaluate whether there were any discrepancies in the type, quality, and number of clinical trials used to approve and cover a medical product.
In addition, we examined the length of time from FDA approval to initiation of an NCD and the time from FDA approval to publication of the final NCD. This was important for us to point out in our study as delays in time from FDA approval to NCD initiation and publication ultimately delay availability of innovative medical technology. Lastly, we described trial participant age and gender for all original trials reviewed in the CMS coverage review process, and analyzed results in the context of average age and gender of Medicare beneficiaries. Given that prior studies have shown that CMS relies mostly on fair to poor quality evidence for review, and that the demographic of trial participants was found to be inconsistent with that of the intended Medicare population, we hypothesized that clinical trials evaluated by the FDA will be more robust than trials evaluated by the CMS. We further hypothesized that neither patients in trials evaluated by the FDA nor patients in trials evaluated by the CMS would be reflective of Medicare beneficiaries.

METHODS

Identifying CMS NCDs and FDA Approval Documents

The Medicare Coverage Database (https://www.cms.gov/medicare-coverage-database/) offers public access to local and NCD memoranda and other documents germane to the Medicare coverage determination process. NCDs are published memoranda that include the following information: the coverage decision (covered, covered with evidence development, or not covered), a concise background of the disease intended to be diagnosed or treated, the history of Medicare coverage, a timeline of recent activities, FDA status of the medical product, general principles of Medicare’s evidence review, the evidence base evaluated by CMS through both
internal and external technology assessments, reports of MEDCAC meetings, and professional society position statements or guidelines that informed the final coverage decision. In December 2016 and January 2017 we downloaded all NCD memoranda and external technology assessments referenced in the memoranda.

All FDA approval documents were available on-line. For moderate-risk devices, 510(k) summaries were downloaded from the FDA website. Links to FDA Summaries of Safety and Effectiveness Data (SSEDs) for all high-risk medical devices were included in the corresponding Medicare determination memoranda. The SSED is a document whose goal is to provide a reasoned, objective, and balanced critique of the scientific evidence which served as the basis of the decision to approve or deny the premarket approval (PMA) [application].(15) Approval of a PMA application is based on a determination by FDA that the PMA contains sufficient valid scientific evidence to assure that [a high-risk device] is safe and effective for its intended use(s).(15) Approval documents for pharmaceuticals and biologics were obtained through the Drugs@FDA on-line database and the FDA site for approved products under Vaccines, Blood, and Biologics, respectively, in December 2016 and January 2017.

Selection of Final Study Sample

We included all NCD memoranda published from January 1, 2005, through December 31, 2016, that pertained to high-risk medical devices, moderate-risk medical devices, pharmaceuticals, or biologics, including both NCDs that resulted in coverage and those that did not. We excluded NCDs for diagnostic technologies, non-medical services, or surgical procedures that did not include discussion of a specific therapeutic product (Figure 1).
Identification of Original Clinical Trials Review by CMS

Each NCD was reviewed by 1 investigator (ACR) who identified original clinical trials used by CMS to support its NCD. Included trial evidence was summarized within the memoranda under the ‘Internal Technology Assessment’ section, which summarizes CMS’ evidence assessment. Original trials (e.g., excluding follow-up studies, interim-analyses, and pooled studies) reviewed by CMS were compared with pivotal trials used to support FDA approval. Uncertainty about trial inclusion was resolved by consensus between 2 investigators (ACR and JSR). We also recorded those additional studies that were based either on the original clinical trials reviewed by CMS or on FDA pivotal trials. These studies were not included in the sample of original clinical trials and included follow-up studies, pooled analyses, or interim analyses.

*These products were listed in the context of a therapeutic procedure, but were not the focus of the NCD.

NCD = National Coverage Determination
Identification of FDA Pivotal Efficacy Trials

One investigator (ACR) identified trials labeled “pivotal” or “primary efficacy and safety” in FDA Medical Reviews for each included pharmaceutical or biologic and in SSEDs for high-risk devices. Of note, some pivotal trials in SSEDs were identified with a preceding statement such as “data from these clinical studies were the basis of the PMA approval decision.” Trials with narrative summaries within 510(k) documents for moderate-risk devices were classified as pivotal trials. Any uncertainty on identification of a pivotal trial was resolved by consensus among all study investigators.

Data Extraction

For each NCD, the following data were extracted from the decision memorandum: date of NCD initiation, date of NCD posting by CMS, individual or agency requesting the NCD, number and type of products specified for coverage, whether MEDCAC was convened, whether an external technology assessment was requested (either from the Agency for Healthcare Research and Quality [AHRQ] or another external health technology assessment body), whether the referenced external technology assessment discussed cost-effectiveness, and the final coverage decision (covered, covered with evidence development [CED], or not covered). Links to NCD decision memoranda are provided on the CMS website for each NCD. The FDA approval date was determined from the NCD decision memoranda or from primary FDA approval documents.

We collected the following data to characterize all original trials evaluated by CMS and FDA: study size, use of randomization, use of double-blinding, inclusion of a control arm (placebo or active control), age of trial participants, proportion of female trial participants,
proportion of trials incorporating clinical outcome(s) as a primary efficacy measure, proportion of trials including a US study location, and proportion of trials with multiple enrollment sites. Four of these variables (study size, use of randomization, double-blinding, and inclusion of a control arm) were considered primary trial characteristics based on the Cochrane Study Quality Guide and prior studies examining quality of evidence reviewed in CMS NCDs.(34, 45, 50)

**Statistical Analysis**

*Characterization of NCDs*

Descriptive statistics were used to evaluate mean CMS review time (from initiation of NCD request to publication of decision memorandum); proportion of NCDs initiated internally by CMS; proportion of memoranda that referenced an advisory report from MEDCAC; proportion of memoranda that referenced an external technology assessment; proportion of external technology assessments that included a discussion of cost-effectiveness; median number of public comments in each of the 2 public comment periods (an initial 30-day period after initiation of the NCD request, and a second 30-day period after publication of the proposed decision memorandum); and proportion of NCDs referencing professional society statements, clinical practice guidelines, FDA approval documents, or expert opinion. We also determined the proportion of CMS evaluations leading to a positive coverage decision, including CED. All calculations were performed in Microsoft Excel v. 15.31 by 1 investigator (ACR) with verification by JSR and SSD.

*Characterization of trials evaluated by FDA and CMS*

For original clinical trials, we used descriptive statistics to evaluate the aforementioned
trial characteristics. The proportion of secondary analyses (i.e., interim-studies, follow-up analyses, pooled studies) reporting data on patients with an average age of at least 65 years was also determined. We used the Fisher Exact Test to compare the following characteristics among trials reviewed by FDA and CMS: use of randomization; double-blinding; inclusion of a control arm (active or placebo); inclusion of a US study location; enrollment from multiple study sites; and inclusion of a clinical outcome in primary efficacy analyses. The Wilcoxon Signed-Rank Test for paired samples was used to compare the total number of trials evaluated by each agency, the age of participants across original trials, the proportion of female trial participants, and the sample size across original clinical trials. Of note, the proportion of female trial participants was only evaluated for 10 of the 12 products because 2 products pertained to the treatment of prostate cancer (PROVENGE and Plenaxis). All tests were 2-tailed and used a type I error rate of 0.05. Analyses were conducted using Excel v. 15.31 by 1 investigator (ACR.), with subsequent verification by JSR and SSD. Because this study made use of publicly-available documents and materials and did not use human subject data, it was exempt from review by the Yale Institutional Review Board.

RESULTS

Characteristics of NCDs

From 2005 through 2016, 11 CMS NCDs covering 12 products (1 NCD included coverage for 2 FDA approved products) met our inclusion criteria. These included 3 pharmaceuticals, 1 biologic, 6 high-risk medical devices, and 2 moderate-risk devices. CMS initiated 4 NCD requests, 3 were requested by product manufacturers, 2 by medical societies, 1 by a Medicare beneficiary, and 1 by a physician.
Median CMS review time was 263 (IQR, 248–272) days. Only 1 (9.1%) NCD included MEDCAC review. Five (45%) NCDs reviewed an external technology assessment, 2 of which included information on cost-effectiveness (Table 1). Eight (73%) NCDs considered professional society statements, and 7 (64%) referenced clinical practice guidelines. Nine (82%) NCDs specifically stated that FDA approval documents were taken into consideration during the coverage determination process. Six NCDs (54%) referenced expert opinion.

The median number of public comments during the first period for all NCDs was 43 (IQR 19-139), and for the second comment period was 76 (IQR 38-164).

A total of 8 (73%) NCDs were positive coverage determinations, with 3 (38%) of those requiring coverage with evidence development. The remaining 3 NCDs were determinations by CMS to not provide coverage.

Number of Trials Evaluated by the FDA and CMS

CMS NCDs for the 12 medical products were based on 27 original clinical trials, and 11 secondary analyses of the original clinical trials. FDA approval of these same 12 products was based on review of 22 pivotal trials. Fourteen (52%) of the 27 original clinical trials evaluated by CMS were the same pivotal trials examined by the FDA (Table 1). Eight (4%) of the pivotal trials evaluated by FDA were not included in CMS NCDs. Of these 8 trials, 4 were for a pharmaceutical (Natrecor), 3 for 1 high-risk medical device (SCULPTRA), and 1 for a different high-risk medical device (MitraClip). The NCD for Natrecor did include a brief result summary from the FDA pivotal trials, along with reference to a meta-analysis that included data from clinical trials used to support FDA approval. The NCD for dermal injections of facial lipodystrophy syndrome, which included SCULPTRA, included results from a 24-month
extension of 1 of the original FDA pivotal trials. Of the remaining 2 trials supporting
SCULPTRA’s FDA approval, 1 was a randomized, controlled trial conducted in the US and the
other was a single-arm trial conducted in the UK. Neither was specifically mentioned in the CMS
NCD. The NCD for transcatheter mitral valve repair (MitraClip) includes a discussion of the
FDA pivotal trial in CMS’ internal technology assessment.

There was no significant difference in the median number of clinical trials evaluated by
the FDA and CMS for all products (FDA 1 [IQR, 1-2] versus CMS 2 [IQR, 1-2]; P = .59).
Table 1. Characteristics of trials used by the FDA for approval decisions and CMS for national coverage determinations

<table>
<thead>
<tr>
<th>FDA Approval Date</th>
<th>CMS Coverage Decision Date</th>
<th>Median Time Elapsed (months)</th>
<th>Interquartile Range (months)</th>
<th>NCD Request</th>
<th>FDA Approval Date</th>
<th>CMS Coverage Decision Date</th>
<th>Median Time Elapsed (months)</th>
<th>Interquartile Range (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>01/01/2000</td>
<td>01/01/2000</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Characteristics of Trials Evaluated by the FDA and CMS

There were no significant differences in the characteristics of trials used to support FDA approval and CMS NCDs. Specifically, the following characteristics in original and pivotal clinical trials were similar in both groups: randomization (FDA: 15/22 [68%] versus CMS: 20/27 [74%]; P = .75), use of either placebo or active control (FDA: 14/22 [64%] versus CMS: 20/27 [74%]; P=.54), double-blinding (FDA: 6/22 [27%] versus CMS: 10/27 [37%]; P = .55), and median number of enrolled study participants (FDA: 279 [IQR, 120–345] versus CMS: 291 [IQR, 174–328]; P = .53) (Table 2).

The number of CMS clinical trials including a US study location was approximately equal to the number of FDA clinical trials including a US study location (FDA: 18/19 [95%] versus CMS: 22/27 [81%]; P = .63). Of note, study location was unavailable for 3 FDA pivotal trials. There was no significant difference in the number of multi-center clinical trials evaluated by FDA and CMS (FDA: 17/22 [77%] versus CMS: 22/27 [81%]; P = .74). There was no significant difference in the use of a clinical outcome measure as a primary efficacy endpoint (FDA: 18/22 [82%] versus CMS: 22/27 [81%]; P=1.00).

There was no significant difference in overall trial participant age (FDA: 59 years [IQR, 45–73] versus CMS: 62 years [IQR, 48–72]; P = .26). Of the original clinical trials reviewed by CMS, 4 (15%) included a subgroup analysis for patients older than 65 years of age. Of the 11 secondary analyses reviewed by CMS based on the original trials, 4 (36%) included subgroup data on patients older than 65 years. There was no significant difference in the proportion of female participants in FDA and CMS clinical trials (FDA: 29% [IQR, 20–44] versus CMS: 32% [24–29]; P= 1.00).
<table>
<thead>
<tr>
<th>CMS*</th>
<th>FDA†</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
<th>Interquartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
<td>0.66 (0.48-0.94)</td>
</tr>
</tbody>
</table>

*Based on one trial. †Interquartile range. ‡Sample size reported as median value averaged over all original clinical trials. CMS original trials are the same as FDA pivotal trials.
Time from FDA Approval Initiation and Publication of a CMS NCD

The median time from FDA approval to initiation of an NCD was 9 (IQR, 4–29) months. The median time from FDA approval to publication of a CMS NCD memorandum was 17 (IQR, 13–36) months.

DISCUSSION

Summary of Findings

We characterized clinical trials used to support FDA approval and CMS NCDs for novel medical products (pharmaceuticals, biologics, moderate and high-risk medical devices) that were issued NCDs from 2005 through 2016, focusing on characteristics reflective of high-quality evidence. There were no significant differences in use of randomization, double-blinding, control arm, or study size between original clinical trials evaluated by CMS and pivotal trials reviewed by FDA. Over half of original clinical trials reviewed by CMS were the same pivotal trials evaluated by FDA.

Participant demographics in trials reviewed by FDA and CMS were similar. Females constituted fewer than 50% of CMS trial participants, which is not significantly different from the median percent of female participants in FDA pivotal trials. In addition, neither CMS nor FDA trials had a median participant age above 65 years. However, CMS did include additional studies in its evaluation that frequently incorporated sub-group analyses of the original clinical trials for patients older than 65 years.

Study Limitations

Limitations of our study include a small number of total reviewed NCDs; only 10 to 15
are issued annually by CMS, (34) whereas the majority of coverage decisions are made by regional contractors, whose determinations are not made publicly available. Such a small sample size of NCDs possibly did not give us the statistical power to detect differences between trial characteristics in studies evaluated by FDA and CMS. However, since NCDs are issued for the most controversial technologies likely to have a significant impact on the Medicare population and apply to all contractors, they are among the most important to study. We should note that the recently enacted 21st Century Cures Act institutes new requirements mandating greater transparency of these regional coverage decisions, including public availability of documents summarizing evidence that supported development of a local coverage determination.(51) As local coverage determinations become available, it would be useful to compare and contrast evidence used to support local versus national coverage determinations. In addition, it would be interesting for future researchers to gather additional data from NCDs as they accrue overtime and conduct additional subgroup analyses on trial characteristics, such as time of follow-up and use of an active or placebo comparison group.

Another potential limitation of our study is our focus on comparing pivotal clinical trials used to support FDA approval of a medical product, to original trials evaluated by CMS for product coverage. CMS often included multiple other studies in their review, ranging from systematic reviews and meta-analyses to case reports. There was no information in the NCD memoranda that definitively stated the value of each study to their review process. We can say that CMS was likely to prioritize original clinical trials, as these trials generally had qualities associated with a stronger evidence base (e.g. randomization and blinding, as outlined in the NCD memoranda). The strength of evidence is important in the CMS coverage evaluation process, and it was our prediction that original clinical trials were nearly equivalent to FDA pivotal clinical
trials. However, future research may benefit from a more comprehensive review of all CMS studies.

Additionally, CMS and FDA documents occasionally had missing data. For example, information on study site was unavailable for 3 FDA pivotal trials. This may have prevented us from making accurate conclusions on differences in trial location, although this was not a main outcome measure in our study.

**Implications for Regulatory Policy**

Our finding that clinical trial characteristics are similar between the two agencies is of particular interest given differences in mandates for FDA approval and CMS coverage. FDA’s goal is to ensure that pivotal trials demonstrate safety and efficacy of the medical product, whereas CMS must ensure that trials show that a product is reasonable and necessary for diagnosis or treatment of an illness or injury in the Medicare population.(52) It is, therefore, reasonable to expect differences in trial characteristics. For example, FDA might be more likely to rely on a greater proportion of randomized, double-blind, placebo-controlled trials. CMS would be expected to prioritize trials that generalize to Medicare beneficiaries, including larger proportions of female participants over 65 years of age and that utilize active controls to determine if there is a benefit to covering the newer therapeutic over available alternatives.

However, our results show that FDA and CMS not only rely on many of the same trials but even when other clinical trials are considered in Medicare’s review, there is no overall difference in the quality of trials evaluated by each agency. Variation in the objectives of evidence review by FDA and CMS does not reflect variation in the trials—and their characteristics—evaluated of product approval/coverage.
Regarding demographics of trial participants, our finding that the median age of patients was under 65 for studies reviewed by both agencies was concerning, along with our finding that approximately two-thirds of trial participants were male. This may have a significant impact for regulatory policy going forward, as more than half (54%) of Medicare beneficiaries are women (53) and are greater than 65 years-of-age. These findings call into question the generalizability of those studies that are crucial for FDA approval and CMS coverage. In the future, it may become necessary for both agencies to seriously consider instituting requirements for manufacturers that they submit trials that are reflective of the Medicare population, particularly if those manufacturers are later seeking CMS coverage. Poor generalizability leads to poor external validity, which means that outcomes reported in clinical trials may not necessarily be reflective of real-world outcomes once the device is used in its intended patient population. We should note that there were several sub-analyses of patients older than 65-years-of-age among CMS original trials; however, this is not ideal and should not be used to replace high-quality studies that can be designed with the Medicare population in mind from trial initiation. Knowing the data specifications requested by CMS in advance (e.g. median age of trial participants should be >65 years, and >50% should be female) might potentially help reduce the time lag from FDA approval to NCD initiation. This would be of great interest for future studies as more products are routed through the parallel review program that was initially piloted and established for medical devices in 2016 (54).

Parallel review is an FDA–CMS collaboration designed to reduce the time between FDA review and CMS evaluation and to help manufacturers understand and address the data requirements of both agencies. (52) Through the program, manufacturers can request initiation of a CMS NCD while the product is still under FDA review. (52) For example, CMS initiated
their coverage review for the Edwards SAPIEN transcatheter heart valve more than 1 month (9/28/11) before FDA approved the device (11/02/11). Whereas the SAPIEN transcatheter heart valve was not officially part of the parallel review program, this was 1 instance in which the manufacturer engaged both FDA and CMS before starting its pivotal clinical trial. Early review obviated the median 9-month delay from FDA approval to initiation of a CMS NCD that we found. Therefore, parallel review may substantially decrease the delay in FDA approval to CMS coverage for medical devices. Yet parallel review is only instituted for medical devices at this time and does not apply to drugs or biologics.

One possible reason for the lack of expansion of the parallel review program is that CMS issues so few NCDs for drugs and biologics that, presumably, incorporating these products into the parallel review pathways is unlikely to have a significant impact. Furthermore, the pharmaceutical industry is not necessarily interested in pursuing a parallel review track. These are only two reasons as to why the parallel review program has not fully taken off even though the first pilot program was initiated over 8 years ago. Additional reasons include limited scope of the parallel review program; currently, the pathway is only for medical devices that require pre-market approval or must be approved under a de novo classification, meaning that it cannot be approved through the 510(k) pathway. This requirement severely limits the number of medical products that are eligible for route through the parallel review program. For example, in 2014 there were over 3,000 510(k) device approvals through the FDA compared to only 42 PMAs. Furthermore, if a manufacturer wants to go through the parallel review program, they are required to submit their product for approval through an NCD and if that product is not approved through an NCD, it will not be available for coverage anywhere in the U.S. through any regional contractor. As a final comment, there is also a limit of how many
products can go through the parallel review pathway each year, (57) effectively creating a bottleneck for manufacturers who despite several limitations, may still want to pursue the parallel review program.

Possible ways to improve the parallel review program are first to expand inclusion criteria to include drugs and biologics, along with moderate-risk medical devices. Given that most Medicare coverage decisions are made by local contractors, it would also be prudent for CMS to expand the parallel review program to cover LCDs in addition to NCDs. Our study has shown that there is poor external validity in studies used to support FDA approval and CMS coverage, leading us to suggest that manufacturers could benefit considerably from greater involvement in a parallel review program that focuses on evaluation of studies that are generalizable to Medicare beneficiaries. However, other factors must be considered as well when merging FDA and CMS review processes.

CMS, unlike FDA, holds 2 30-day public comment periods for each NCD and its decision memorandum includes a summary of, and response to, comments received. Public comments are designed to enhance the “quality of agency decision making” and can encompass contributions ranging from medical society recommendations to patient narratives.(58) In order to streamline the parallel review program, perhaps the public comment period could be initiated within a specified time period following FDA approval.

Furthermore, CMS considers evidence-based guidelines and the opinions of members of the medical or scientific community (labeled “expert opinion” in NCD memoranda) and may consider MEDCAC reports. Since many of these additional reports and guidelines reflect direct patient and physician experiences with the medical product once it is on the market, it will be difficult to merge these processes through parallel review. An additional barrier to the parallel
review program may involve the relationship between manufacturers and the national CMS agency. Because of specific requirements for CPT coding, manufacturers may receive lower reimbursement rates from CMS once an NCD is issued. In addition, anecdotally, it has been suggested that manufacturers may have the potential to exert more influence in the local coverage determination (LCD) process, although this has not been systematically studied.

Despite these limitations, expansion of the parallel review program would be ideal for expediting the coverage process for beneficial medical devices, drugs and biologics. However, several limitations must first be addressed and manufacturers should be given more incentives for participating in the program.

Conclusions

In conclusion, FDA approval and CMS NCDs of novel therapeutics often rely on the same clinical trial evidence and on trials of similar quality. However, the process of finalizing coverage determination requires an additional 17 months. FDA and CMS should continue to work together to ensure timely coverage decisions after FDA approval, perhaps by encouraging manufacturers to include larger proportions of older and female participants in their trials supporting FDA approval.
References


