

## **SECTION TWO**

### **DATA SOURCES**

This EA relies on a variety of data sources, including the Section 308 Pharmaceutical Survey conducted specifically for this regulatory development effort, the U.S. Department of Commerce, the U.S. Food and Drug Administration (FDA), Bureau of Labor Statistics (BLS), Dun & Bradstreet (D&B), Robert Morris Associates (RMA), the Pharmaceutical Research and Manufacturers of America (PhRMA), and various journal articles. Most of the analyses conducted in Sections Four through Ten make extensive use of the data collected from the Section 308 Pharmaceutical Survey. Other data sources were used primarily in the development of the industry profile in Section Three. Data gathered in the profile, however, provide the foundation for much of the analysis in later sections.

The following sections describe the three principal data sources for this EA: the Section 308 Pharmaceutical Survey, sources available through the U.S. Department of Commerce, and data on compliance costs of an air rule requiring Maximum Achievable Control Technology (MACT) to control air emissions. This MACT standards rule also will affect many of the same facilities and will be finalized at nearly the same time as the Final Pharmaceutical Industry Effluent Guidelines. Other data sources are described, as necessary, in Sections Three through Ten.

#### **2.1 THE SECTION 308 PHARMACEUTICAL SURVEY**

The Section 308 Pharmaceutical Survey obtained detailed technical and financial information from a sample of pharmaceutical establishments potentially affected by EPA's proposed effluent guidelines. EPA stratified the industry into five groups based on type of operation:

- A) Fermentation
- B) Biological and natural extraction
- C) Chemical synthesis

- D) Formulation and mixing/compounding
- E) Research

The stratification permitted EPA to census (i.e., survey all facilities) facilities within some subcategories and sample facilities within others. EPA took a census of all facilities that (1) manufacture active ingredients (subcategories A, B, C) and directly discharge process wastewater and (2) perform formulating and mixing/compounding (subcategory D) and directly discharge or directly and indirectly discharge process wastewater. EPA judged that a census of these facilities was necessary to achieve statistical accuracy because the overall universe was small, few facilities were in the same combination of subcategories, and each facility was expected to have wastewater generated by proprietary processes that would make their effluent significantly different from other facilities in the same subcategory. Overall, EPA conducted a census of 202 facilities in these four subcategories.<sup>1</sup>

EPA also censused subcategory D stand-alone facilities that use solvents and discharge indirectly, and subcategory D facilities with onsite research facilities (i.e., subcategory D/E) that use solvents, discharge indirectly, and have fewer than 19 employees or more than 747 employees. For subcategory D indirect discharging facilities with between 19 and 168 employees and between 169 and 747 employees, EPA used a sampling methodology. The sampling methodology stratified these facilities by flow rates and employee size using a linear regression between the log of the number of employees and log of the flow rate. Employee and flow rate data were available from EPA's *Development Document for Effluent Limitations Guidelines and Standards for the Pharmaceutical Point Source Category*.<sup>2</sup> Overall, EPA sampled 42 pharmaceutical facilities in subcategories D and D/E.<sup>3</sup> Survey results used throughout the EA are weighted according to the sampling plan. Subcategory D and D/E facilities with between 19 and 747 employees received a weight of approximately 2 (because only about half of these facilities were surveyed). (All subcategory D facilities are grouped with subcategory B facilities for the purpose of this analysis, which is discussed in Section Four.) All

---

<sup>1</sup> U.S. EPA, 1990. U.S. Environmental Protection Agency. *Supporting Statement for OMB Review: Detailed Questionnaire for the Pharmaceutical Manufacturing Industry*. Washington, DC: Office of Water Regulations and Standards.

<sup>2</sup> U.S. EPA, 1983. U.S. Environmental Protection Agency. *Development Document for Effluent Guidelines, New Source Performance Standards, and Pretreatment Standards for the Pharmaceutical Manufacturing Point Source Category*. Washington, DC: U.S. EPA.

<sup>3</sup> U.S. EPA, 1990. *Op. cit.*

other facilities received a weight of 1. The coefficient of variation in any particular strata (i.e., combination of subcategory and flow group) is not greater than 15 percent. Thus the total survey universe comprises 286 facilities, 202 of which were censused and 84 of which were sampled at approximately a 50 percent sampling rate.

EPA determined that no information was needed from three groups of pharmaceutical facilities:

- Facilities that do not discharge wastewater
- Facilities that do not use solvents and whose only source of process wastewater is from formulation and mixing/compounding
- Stand-alone research facilities

These facilities do not require effluent guidelines because their impact on water quality and POTW operations is considered to be negligible.

The survey data were used extensively in the development of BPT, BCT, BAT, NSPS, PSES, and PSNS regulations for the industry. Surveyed facilities provided technical information on pharmaceutical products; compound and chemical usage and disposition; waste minimization and pollution prevention activities; wastewater generation, collection, and conservation; wastewater treatment; steam stripping; and wastewater characteristics. The survey also collected financial data such as number of employees; ownership structure; discount rate; market value of land, buildings, and equipment; value of shipments; manufacturing costs; assets; liabilities; earnings; and net income. Financial data were collected at the facility, owner-company, and parent company levels.

All surveyed facilities were given the option to legally certify that the facility would incur no significant economic impact as a result of the effluent guidelines. These facilities gave up their right to challenge aspects of the Final Pharmaceutical Industry Effluent Guidelines based on economic achievability so long as the cost of compliance with the effluent guidelines ultimately promulgated by EPA does not exceed the compliance cost estimated in the survey. Certifying facilities were excused from completing the bulk of the financial questionnaire. Sixty-five of the 244 surveyed facilities certified no significant economic impact and thus did not provide financial data.

## 2.2. U.S. DEPARTMENT OF COMMERCE SUPPLEMENTAL DATA

The EA supplements financial data collected in the Section 308 Pharmaceutical Survey with data from the U.S. Department of Commerce. Commerce divides the pharmaceutical industry into four 4-digit Standard Industrial Classifications (SICs):

- ***SIC 2833 Medicinal and Botanical.*** Establishment primarily engaged in: (1) manufacturing bulk organic and inorganic medicinal chemicals and their derivatives and (2) processing bulk botanical drugs and herbs.
- ***SIC 2834 Pharmaceutical Preparations.*** Establishments primarily engaged in manufacturing, fabricating, or processing drugs in pharmaceutical preparations for human or veterinary use. The greater part of the products of these establishments are finished in the form intended for final consumption, such as tablets, capsules, liquids, etc. These pharmaceutical preparations are promoted to the medical profession (prescription drugs) and the general public (over-the-counter [OTC]).
- ***SIC 2835 In Vitro and In Vivo Diagnostic Substances.*** Establishments engaged in the manufacturing of chemical, biological, and radioactive substances used in diagnosing or monitoring human and animal health by identifying and measuring normal and abnormal constituents of body fluids or tissues.
- ***SIC 2836 Biological Products, Except Diagnostic Substances.*** Establishments engaged primarily in the production of bacterial and virus vaccines, toxoid, and analogous products, serums, plasmas, and other blood derivatives for human and veterinary use.

Commerce collects a wide range of data at the 4-digit SIC level including number of establishments, number of employees, volume of shipments, exports, imports, value added, apparent consumption, manufacturing costs, and other data. Commerce further segments the pharmaceutical industry into 14 five-digit and hundreds of seven-digit SIC codes. Comprehensive financial data at the five- and seven-digit levels, however, is available only under SIC 2834 Pharmaceutical Preparations. Commerce data are reported in publications such as the *Census of Manufactures*, *County Business Patterns*, and *U.S. Industrial Outlook*. The EA uses the most current available data from these sources in the development of the industry profile in Section Three.

Numerous other data sources employed by EPA in the EA are organized by SIC code. For example, price indices generated by BLS and financial ratio data reported by D&B and RMA are organized by SIC code.

A major difficulty with using data organized by SIC is its inability to capture all establishments engaged in the production of pharmaceuticals. Commerce classifies facilities by their primary line of business. Thus, only establishments that garner at least 50 percent of their revenues from pharmaceutical-related business are classified in the four pharmaceutical SIC codes. Facilities that manufacture pharmaceuticals but list some other line of business (e.g., chemical production) as their primary SIC are not captured in the four pharmaceutical SICs. Thus, Commerce data do not provide a complete picture of the U.S. pharmaceutical industry.

The Section 308 Pharmaceutical Survey data cover only a subset of the pharmaceutical industry. The five categories used to segment the pharmaceutical industry in the survey do not correspond with the four pharmaceutical SICs. Moreover, surveyed facilities were not asked to report their SIC. Thus, no direct comparison can be made between Commerce and survey data.

### **2.3 MACT STANDARDS COST DATA**

EPA's Office of Water, Engineering and Analysis Division, received cost data from EPA's Office of Air Quality Planning and Standards (OAQPS). These data included capital and operating costs for 98 facilities to install and operate equipment to meet MACT air quality standards. Of these 98 facilities, 71 will incur both MACT standards and effluent guidelines costs. Because the two rules (effluent guidelines and MACT standards) will be finalized in 1998, EPA considers the effect of MACT standards costs on these 71 facilities in this EA. MACT standards costs include costs for six components: equipment leaks, dedicated process vents, nondedicated process vents, storage tanks, partially soluble wastewater, and soluble wastewater. The last two MACT standards cost components are considered wastewater emission control costs; and the entire group of costs are considered total MACT standards costs. EPA has developed three baselines for assessing impacts of the Final Pharmaceutical Industry Effluent Guidelines. Baseline 1 uses just the Section 308 Pharmaceutical Survey data to establish current conditions. EPA incorporates the wastewater emission cost portion of the MACT standards costs into Baseline 1 to create a Baseline 2, and incorporates total MACT standards costs into Baseline 1 to create Baseline 3 (see Sections Five and Six for more details). The impacts of the Final Pharmaceutical Industry Effluent Guidelines are judged against all three baselines. Appendix B presents the costs as received from OAQPS and used in creating Baselines 2 and 3.