

**UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF NEW JERSEY**

UNITED STATES OF AMERICA : Civil Action No.
 : 92-1744
Plaintiff, :
 :
v. :
 : OPINION
 :
BARR LABORATORIES INC., :
et al., :
 :
Defendants. :

APPEARANCES:

MICHAEL A. CHAGARES, ESQ.
BETH A. KASWAN, ESQ.
DEBORAH Y. YEOH, ESQ.
STEVEN I. FROOT, ESQ.
SPECIAL ASSISTANT U.S. ATTORNEYS
OFFICE OF THE U.S. ATTORNEY
970 Broad Street
Newark, NJ 07101
(Attorneys for Plaintiff)

THOMAS E. MOSELEY, ESQ.
LEE H. UDELSMAN, ESQ.
DE MARIA, ELLIS, HUNT,
SALSBERG & FRIEDMAN
744 Broad Street
Newark, NJ 07102

-and-

ERIC L. HIRSCHHORN, ESQ.
CATHY L. BURGESS, ESQ.
MICHAEL K. ATKINSON, ESQ.
WINSTON & STRAWN
1400 L Street, N.W.
Washington, D.C. 20005
(Attorneys for Defendant
Barr Laboratories, Inc.)

WILLIAM C. SLATTERY ESQ.
MEGAN E. GLOR, ESQ.
NORRIS, MCLAUGHLIN & MARCUS
721 Route 202-206
Box 1018
Somerville, NJ 08876-1018

-and-

ROBERT H. BECKER, ESQ.
ALAN H. KAPLAN, ESQ.
RICHARD S. MOREY, ESQ.
KLEINFELD, KAPLAN & BECKER
1140 19th Street, N.W.
Washington, D.C. 20036
(Attorneys for Defendants
Cohen, Price and Hamza)

CONTENTS

BACKGROUND	3
I. FINDINGS OF FACT	4
A. THE PARTIES	4
B. THE REGULATORY SCHEME	4
C. FDA INVESTIGATORY PRACTICE	5
D. DRUG TESTING OVERVIEW	5
1. Failures	5
2. Failure investigations	6
3. Outliers	8
4. Retesting	8
5. Resampling	10
6. Remixing	10
7. Averaging	10
8. Releasing a Batch for Distribution	11
9. Blend Testing	12
(a) Sample Size	12
(b) Site of Sampling	13
E. CRITICISMS OF BARR	13
1. Manufacturing Process Validation	14
(a) General Requirements	14
(b) Specific Process Problems	15
(1) Batch Failures	15
(2) Retrospective Validation	15
(a) Omitting Failing Results	15
(b) Omitting Failing Batches	16
(c) Insufficient Number of Batches	16
2. Failure Investigations	18
3. Release of "Failing" Batches	18
(a) and (b) Rejecting Failing Results As Laboratory Errors After Retesting	18
(c) Rejecting Failing Results After Resampling	19
4. Failure to Control Manufacturing Process Steps	20
(a) Barr Blend Testing	20
(1) Site of sampling	20
(2) Sampling Procedure	21
(3) Sample Size	21
(b) Mixing Time	21
(c) Particle Size Distribution	21
5. Failure to Validate Testing Methods	21
6. Cleaning Validation Deficiencies	23
7. Record-Keeping Deficiencies	27
8. Failure to File ANDA Supplements and Field Alerts	27
(a) ANDA Supplements	27
(b) Field Alerts	28
II. CONCLUSIONS OF LAW	29
A. JURISDICTION	29
B. STANDARD FOR INJUNCTIVE RELIEF	29
C. THE FEDERAL FOOD, DRUG AND COSMETIC ACT	29
D. EVALUATING BARR LABORATORIES	31
E. FASHIONING A REMEDY	31
(1) General CGMP violations	31
(2) Product Validation	32
(3) Product Recall	34
CONCLUSION	36

WOLIN, District Judge

Currently before the Court is plaintiff's application for a preliminary injunction directing defendants to suspend, recall or revamp numerous products in their current product line. Plaintiff filed this action in the United States District Court in the Southern District of New York on June 12, 1992, alleging that defendants violated the Federal Food, Drug and Cosmetic Act. In accordance with the first-filed rule, the case was transferred to the District of New Jersey on June 26, 1992, where it was consolidated with an action defendants brought against plaintiff on April 24, 1992, seeking relief from allegedly ad hoc drug regulations. On July 10, 1992, the Court filed case management and protective orders. Beginning on August 17, 1992, and continuing intermittently to October 12, 1992, through the testimony of inspectors, experts and employees, the parties presented exhaustive but conflicting views of defendants' business practices. At the Court's direction, on October 26, 1992, the parties submitted proposed findings of fact and conclusions of law.

BACKGROUND

Each day with confidence and hope millions of people in the United States and other countries reach for pills, powders, capsules and syrups to relieve or prevent an infinite number of physical and mental ailments. The weighty task of ensuring the integrity of these products, frequently unquestioned by most consumers, falls to the Food and Drug Administration, which monitors the practices of the drug industry through a system of approvals and investigations. Built into this maze of often ambiguous rules, however, is the recognition that drug manufacturers are businesses, which must follow efficient as well as effective procedures.

The current conflict surrounding these rules is best characterised as a confrontation between a humorless warden and his uncooperative prisoner. Exchanging heavy blows, the parties generated a record of more than twenty-three hundred pages of testimony, almost four hundred exhibits and numerous lengthy declarations. The plaintiff presented two witnesses, a government inspector, David Mulligan, and a regulatory expert, Dr. Robert Gerraughty. Defendants countered with a statistician, Dr. Sanford Bolton, their own regulatory expert, Dr. Christopher Rhodes, an analytical chemist, Dr. Norman Atwater, an expert in pharmaceutical biology, Dr. Murray Cooper, and Barr employees. These witnesses revealed an industry mired in uncertainty and conflict, guided by vague regulations which produce-tugs-of-war of varying intensity.

The divergent views presented to the Court reflect not only a difference of perspective, but also the changes made at Barr Laboratories since the first threat of this litigation. As a result, the record is a composite of two trials: the case that was and the case that is. (1789:6 (discussion between Court and Ms. Kaswan));¹ Barr Response ¶ 63. As such, the bases upon which some of the government's criticisms rest have disappeared during the course of this litigation. Wary of this timing element, the Court has reviewed the lengthy record and the parties' proposed findings with the dual desire to protect an unsuspecting public and to avoid unnecessarily burdensome rules and now makes the following findings of fact and conclusions of law.

¹ Because of frequency with which the Court references the hearing transcript, citations to the record will denote the page, line and speaker as follows: (page:line (speaker)).

I. FINDINGS OF FACT

A. The parties

1. Plaintiff, United States of America, brought this action on behalf of the Food and Drug Administration ("FDA"), an agency within the United States Department of Health and Human Services.
2. Defendant Barr Laboratories, Inc. ("Barr") is a manufacturer and distributor of drug products in the interstate and foreign commerce of the United States. Barr is incorporated under the laws of the State of New York and is doing business in Northvale, New Jersey and Pomona, New York.
3. Barr currently manufactures sixty drug products. Before this action was commenced, Barr voluntarily suspended the production and distribution of 115 drug products pending further order of the Court.
4. From 1970 until January 5, 1993, defendant Edwin A. Cohen was Barr's president and Chief Executive Officer and was in charge of the day-to-day operations of Barr. His current title is Chairman of the Board and Chief Executive Officer.
5. Defendant Gerald F. Price is Barr's Executive Vice president of Operations and is responsible for the performance of Barr's quality assurance department. He has held this position since 1990.
6. Defendant Ezzel-Din A. Hamza is Barr's Vice President of Technical Affairs and is responsible for regulatory affairs, research and development, and Barr's chemistry and microbiology laboratories. He has held this position since 1989.

B. The Regulatory Scheme

7. Under the Federal Food, Drug and Cosmetic Act (the "Act"), a drug is adulterated if "the methods used in, or the facilities or controls used for, its manufacture, processing, packing or holding do not conform to or are not operated or administered in conformity with current good manufacturing practice to assure that such drug meets the requirements of this chapter as to safety and has the identity and strength, and meets the quality and purity characteristics, which it purports or is represented to possess." 21 U.S.C. § 351(a)(2)(B).
8. Current Good Manufacturing Practice ("CGMP"), explained in greater but by no means sufficient, detail in regulations promulgated by the FDA, see 21 C.F.R. Parts 210 and 211, sets the minimum standards for drug manufacturers. Designed as a quality control measure to prevent super- and sub-potency, product mix-ups, contamination, and mislabeling, (870:20; 872:3 (Gerraghty)), the CGMP regulations outline general rules for all aspects of drug manufacture including buildings and facilities, personnel, equipment, drug components and containers, production, packaging and labeling, and record-keeping. Failure to comply with CGMP regulations renders any resulting drug product "adulterated" and the drug product and its producer subject to regulatory action. 21 C.F.R. § 210.1(b).
9. Congress recognizes the United States Pharmacopeia ("USP"), a nonprofit corporation which develops drug product standards with the help of professionals from academia, the medical community, the pharmaceutical industry and the FDA, as an official compendium. (1503:11 (Rhodes)). The USP supplements the CGMP provisions, specifying both the types of tests firms must perform and the range of acceptable results these tests must generate for releasing drug products. As with the CGMP regulations, the USP contains minimum testing requirements, (910:9 (Gerraghty)). These criteria, however, rest on the assumption that the products to which they apply satisfy CGMP. (911:11 (Gerraghty)).
10. Firms outline their chosen standards and procedures in new drug applications (NDAs) or abbreviated new drug applications (ANDAs), which are submitted to the FDA for approval. The ANDA guides a product's manufacture and release, but does not supersede the overarching CGMP requirements (448:24 (Mulligan)).
11. Thus, the CGMP regulations provide the yardstick with which FDA investigators, and the Court in the instant action, measure firm behaviour. Ironically, the regulations themselves, whose broad and

sometimes vague instructions allow conflicting, but plausible, views of the precise requirements, transform what might be a routine evaluation into an arduous task.

12. To the extent that the regulations create ambiguities, the industry can turn for guidance to literature from seminars and pharmaceutical firms, textbooks, reference books and FDA letters to manufacturers, (865:11 (Gerraghty)), or employ scientific judgement where appropriate. The Court, however, cannot rely on industry practice to determine whether an individual firm meets the statutory requirements, since industry standards themselves must be reasonable and consistent with the spirit and intent of the CGMP regulations. (1258:24 (Gerraghty)).

C. FDA Investigatory Practice

13. FDA investigators conduct both pre-approval and general compliance inspections. During a pre-approval inspection, investigators review pending ANDA applications, (414:7 (Mulligan)), while inspectors conduct compliance investigations to determine whether the firm is following CGMP. After a compliance investigation, FDA inspectors must issue a Form 483 Notice of Inspectional Observations ("Form 483"), in which they record their observations about the firm's more serious CGMP violations.
14. FDA investigators conducted a general inspection of Barr's Northvale facility during August and September 1989 as well as separate general inspections of Barr's Northvale and Pomona facilities from May to September 1991. After each inspection, the investigators issued Forms 483.
15. The 1989 Form 483, which contained six general observations, cited Barr for unvalidated manufacturing and cleaning processes, the lack of failure investigations, incomplete annual reviews and failure to explain retesting. See Exhibit 51. In the 1991 Pomona inspection, the FDA criticized nineteen products extensively and made general comments about Barr's equipment and complaint logs, stability programs, raw material controls and documentation procedures. See Exhibit 3. The Northvale inspection that same year also censured individual products as well as failure investigations, validation and product mix-ups. See Exhibit 1.
16. In February 1992, FDA investigators returned to the Barr facilities in Northvale and Pomona. The resultant 1992 Forms 483 listed seventy-five and forty-seven criticisms for Northvale and Pomona, respectively, many of which also had been recorded during prior visits to Barr, in the areas of process validation, failure investigations and laboratory practices. See Exhibits 2, 4.

D. Drug Testing Overview

17. Testing lies at the heart of a drug manufacturer's successful operation. Through testing companies validate their processes and ensure the quality of batches for release. As the Forms 483 suggest, much of this current litigation stems from allegedly defective testing practices. With the mechanics of test-taking left undefined by the regulations, before discussing the specific allegations against Barr, the Court will outline the CGMP-required parameters which will guide its evaluation.

1. Failures²

18. In the government's view, a batch failure occurs each time an individual test result does not meet the specifications outlined in the USP or the firm's ANDA. (132:22 to 133:3 (Mulligan)). In contrast, Barr does not classify initial out-of-specification results as batch failures. Instead, only after confirming out-of-specification results with additional testing, (1394:1 (Bolton); 1978:18 (Cooper)), pursuant to the firm's predetermined testing procedure, (1979:7 (Cooper)), would Barr conclude that a batch failed.
19. Out-of-specification results obtained in the laboratory fall into three general categories: (1) laboratory error; (2) non-process-related or operator error; and (3) process-related or manufacturing error. (227:23 (Mulligan))

² What's in a name? Would a "failure" by any other name smell sweet? The Court will use the term "out-of-specification result" when referring to an individual test value that does not meet predetermined specifications. Investigations that follow such results are labeled "failure investigations." The Court has attempted to avoid the use of "failure" when discussing non-passing results.

20. Laboratory error can result from an analyst's mistake or malfunctioning laboratory equipment. (923:18 (Gerraughty)). Examples of analyst error include mistakes in calculations, the use of incorrect standards for comparison, and simple mis-measurement. (923:17 (Gerraughty)). Those human and mechanical errors which occur during the manufacturing process cause non-process-related errors. For example, manufacturing equipment may malfunction or an operator may fail to add the proper amount of an active ingredient. In contrast, process-related problems, such as an incorrect mixing time, occur even though the workers and machines function properly.
21. While each type of problem is a matter of great concern which requires some form of corrective action, only nonprocess-related and process-related errors properly are labelled failures. As Inspector Mulligan acknowledged, all failures are not alike. (133:20 (Mulligan) (noting-if "reason and logic or science" indicate that out-of-specification result is anomaly, batch need not fail)). An out-of-specification result identified as a laboratory error by a failure investigation or an outlier test, *see* Ex. 7 at 1503; (202:21, 798:20 (Mulligan)), or overcome by re-testing is not a failure. (225:7-14 (Mulligan)). Thus, the Court is unwilling to adopt the government's view of failure.

2. Failure investigations

22. Only with an investigation will a firm be able to identify the cause of an out-of-specification result. CGMP requires a thorough investigation following:

[a]ny unexplained discrepancy (including a percentage of theoretical yield exceeding the maximum or minimum percentages established in master production and control records) or the failure of a batch or any of its components to meet any of its specifications . (which) shall extend to other batches of the same drug product and other drug products that may have been associated with the specific failure or discrepancy. A written record of the investigation shall be made and shall include the conclusions and follow-up.

- 21 C.F.R. § 211.192; *see* (924:19 (Gerraughty) (violation of CGMP to discard out-of-specification results and pass batch on retesting alone)).
23. The government argues that an adequate failure investigation requires a timely, thorough, and well-documented review of the problem, which yields a written record containing: (1) the reason for the investigation; (2) a summation of process sequences that may have caused the problem; (3) the corrective actions necessary to save the batch and prevent recurrence; (4) list of other batches and other products possibly affected along with their investigation results; and finally, (5) comments and signatures of production and quality control personnel regarding approval of any material reprocessed after additional testing. Government Findings ¶ 59. Barr advocates a sliding-scale approach, claiming that the nature of the failure should govern the intensity of the investigation.
24. In accordance with the CGMP-required for-cause failure investigation, (1027:13 (Mulligan)), the goal of every such inspection is to determine into which of the three failure categories the problem falls. The degree of inquiry required successfully to complete this task may vary with the object under investigation. As a result, a full investigation of the type the Government outlined always will not be necessary.

25. The issue of failure investigations first arises when testing produces a single out-of-specification result. Before proceeding to retest, (924:11 (Gerraughty)), this unhappy occurrence must be met with a step-by-step review of the suspect laboratory tests. (Id.; 1975:12 (Cooper); 1517:1 (Rhodes)). Specifically, the analyst who performed the test must report the problem to a supervisor, and the two technicians must conduct an informal laboratory inspection, reviewing the notebook which contained the out-of-specification result, discussing the testing procedure along with any required calculations and examining the instrument used. (295:20 (Mulligan); 1831:8 (Atwater)). Thus, the Court requires that a single out-of-specification result be met with more than "a laboratory investigation consisting principally of retesting." Barr Response to Government Findings ¶¶ 58-59.
26. Such an investigation, along with any conclusions reached, must be preserved with written documentation that enumerates each step of the review, (1519:10 (Rhodes)), in the form of a "computer generated flow sheet" (1517:19 (Rhodes)), or a check-list. (1974:25 (Cooper)). This writing should be preserved in an investigation or failure report, (1517:1 (Rhodes)), and placed in a central file. (1519:19 (Rhodes)). In order to enhance this review process, each analyst conducting a test should follow a written procedure, checking off each step as it is completed. (1974:25 (Cooper)).
27. Any easily identified analyst errors, such as calculation mistakes, should be specified with particularity and supported by evidence. (925:20 (Gerraughty)). In some instances, however, "the subtle influences which can result in test variability are not apparent when such an assay or test investigation is carried out." (1954:13 (Cooper)). Thus, because it can be difficult to pin down the exact cause of a problem (1833:13 (Atwater)), it is unrealistic to expect that the cause of analyst error always will be determined and documented. (1416:22 (Bolton)).
28. In recognizing the existence of less readily identifiable mistakes and the influence of variables unrelated to the purity or potency of the drug under scrutiny, the Court does not intend to create a means for firms to avoid the performance and documentation of an informal laboratory investigation. The inability to identify an error's cause with confidence affects retesting procedures, see ¶¶ 38-39 *infra*, but does not affect the failure inquiry required for initial out-of-specification results.
29. Other problems more serious than single out-of-specification results, from multiple out-of-specification results to product mix-ups and contamination, require full-scale inquiries involving quality control and assurance personnel in addition to laboratory workers in order to identify the exact process or non-process-related errors.
30. Extending beyond the laboratory and often labeled formal investigations, (1563:25 (Rhodes)), these inquiries should follow the outline the Government provided, with firms paying particular attention to any necessary corrective action, whether reprimanding, retraining or firing employees, remixing batches or adjusting processes. Thus, in the failure report firms must: (1) state the reason for the investigation; (2) provide a summation of the process sequences that may have caused the problem; (3) outline the corrective actions necessary to save the batch and prevent a similar recurrence; (4) list other batches and products possibly affected, the results of their investigations, and any required corrective action; and finally, (5) preserve the comments and signatures of all production and quality control personnel who conducted the investigation and approved any reprocessed material after additional testing.
31. The outcome of the failure investigation will determine whether additional batches of the same product and related products also require remedial measures. (1709:10 (Rhodes)). process-related errors suggest the need to examine other batches of the problem product as well as other products made according to similar procedures. Addressing non-process related errors requires an examination of other batches or products the trouble-making employee or machine may have handled.
32. Thus the elements of a "thorough" investigation necessarily will vary with the nature of the problem identified. However, all failure investigations must be performed promptly, within thirty business days of the problem's occurrence, and recorded in written investigation or failure reports.

3. Outliers

33. The outlier test provides an alternative means of invalidating an initial out-of-specification result. If the failure investigation of an initial out-of-specification result proves inconclusive, firms searching for a better explanation can utilize this method.
34. Significant limits accompany the outlier test, however. The USP specifically warns against using outlier tests too often, and thus, as a general rule,³ firms must be careful not to reject results frequently on this basis. (2036:9 (Baldassarre))
35. In addition, the utility of the outlier test varies with the type of assay performed.⁴ The USP expressly allows firms to apply this test to biological and antibiotic assays, see Ex. 7 at 1503; (202:21, 798:20 (Mulligan)), but is silent on its use with chemical tests. (1239:8 (Mulligan)). Although some experts advocated use of the outlier method for chemical assays, (1395:1 (Bolton)), other testimony suggested that firms generally do not rely on outlier tests to invalidate chemical test results. (1238:21 (Mulligan)). In the Court's view the silence of the USP with respect to chemical testing and outliers is prohibitory.
36. The substantial innate variability of microbiological assays, (1916:1 (Cooper)), supports this distinction. (1238:11 (Mulligan)). Chemical assays are considerably more precise than biological and microbiological assays, (1915:12 (Cooper)), since only the latter testing is "subject to whims of microorganisms." (1915:18 (Cooper)).

4. Retesting

37. In addition to triggering a failure investigation, out-of-specification results also generate the need for retesting. A retest is defined as additional testing on the same sample, and thus it necessarily follows an initial test. An analyst performing a retest takes the second aliquot from either: (1) the sample that was the source of the first aliquot (1022:5 (Bolton), 1533:2 (Rhodes)); or (2) the larger sample previously collected for laboratory purposes.⁵ (1534:8 (Rhodes)). These procedures are equivalent. (1839:17, 1840:16 (Atwater)). Thus, whether retesting is performed at the finished product or blend stage, such testing should be performed from the same bottle of tablets or capsules and the same drum or mixer, respectively. (661:2 (Mulligan); 1423:3 (Bolton)).

³ The Court chooses not to rely on the phrase "rule of thumb." Although apparently a common and useful phrase, as evidenced by the frequency with which it appears in the hearing transcript, this expression is derived from the historical common law right of the husband to beat his wife with a switch, provided it was "no thicker than his thumb." Caitlin E. Borgman, Note, Battered Women's Substantive Due Process Claims: Can Orders of Protection Defeat DeShaney?, 65 N.Y.U. L. Rev. 1280, 1281 n.3.

⁴ Outlier tests should not be used to invalidate content uniformity test results, since variability among test values may indicate nonuniformity in the blend. See Hammerstrom Declaration, ¶¶ 10-12.

⁵ Although there is a possibility that any two 20-tablet grinds taken from the same laboratory sample will not be identical (1676:10 (Rhodes)), this risk is small and this procedure properly is termed a retest and not a resample. (1677:1 (Rhodes)). Any variation in the batch should be detected through content uniformity testing. (1677:14 (Rhodes)).

38. Retesting is proper only after a failure investigation is underway, (235:1, 661:2 (Mulligan)), since the outcome of the failure investigation itself, in part, determines when retesting is appropriate.⁶ Retesting is necessary if a failure investigation indicates that analyst error caused an initial out-of-specification result. (144:1, 1023:18 (Mulligan)). A retest is similarly acceptable when review of the analyst's work is inconclusive. (664:12 (Mulligan)). In these instances, retesting substitutes for or supplements the original tests which have been rejected or questioned, respectively. In the case of nonprocess and process-related errors, however, retesting is suspect.⁷ (1024:10 (Gerraughty)). Because the initial tests are genuine, in these circumstances, additional testing alone cannot infuse the product with quality. (1024:15 (Gerraughty)).
39. As a general matter, the amount of retesting required also varies with the problem identified. Out-of-specification results attributed to analyst error require limited retesting. (1529:19 (Rhodes)). Here, retesting merely supplants the first round of initial tests. More extensive retesting should follow an inconclusive failure investigation, since firms need to determine whether the out-of-specification result is a mere anomaly or a reason to reject the batch. (1390:1 (Bolton))⁸
40. The USP contemplates retesting for quality control purposes (1937:14 (Cooper)), although it does not prescribe or recommend the number of individual tests that must be performed in order to reach a definitive conclusion about the quality of a product. (1929:1 (Cooper)) Thus, the number of retests performed before a firm concludes that an unexplained out-of-specification result is invalid or that a product is unacceptable is a matter of scientific judgment. (488:7, 816:9 (Mulligan); 1959:3 (Cooper)). Yet the goal of retesting is clear; firms must do enough testing to isolate the out-of-specification result in order to reach the point at which the additional testing overcomes the out-of-specification result. (805:4 (Mulligan)).⁹
41. Nevertheless, retesting cannot continue ad infinitum. Because such a practice is not scientifically valid, a firm's predetermined testing procedure should contain a point at which testing ends and the product is evaluated. At this time, if the results are not satisfactory, the batch must be rejected. (1958:7 (Cooper)).
42. When evaluating retest results, it is important to consider them in the context of the overall record of the product. Relevant to this review are the history of the product,¹⁰ the type of tests performed,¹¹ and any results obtained for the batch at other stages of testing.¹² As such, retesting determinations will vary on a case-by-case basis, a necessary corollary of which is that an inflexible retesting rule, designed to be applied in every circumstance, is inappropriate. (1476:9 (Bolton)).

⁶ The line between a failure investigation and retesting is difficult, if not impossible, to draw. Although a single out-of-specification result alerts the firm to a possible problem, and the need to investigate, one test does not allow conclusive judgments about the batch. (1394:13 (Bolton)). Thus, repeat testing on the same sample is an appropriate part of the failure investigation. (1261:18-1262:2 (Mulligan); 1419:3 (Bolton))

⁷ The Court acknowledges that some retesting may precede a finding of nonprocess or process-based errors. Once this determination is made, however, additional retesting for purposes of testing a product into compliance is not acceptable.

⁸ For example, in the case of content uniformity testing, designed to detect variability in the blend or tablets, failing and nonfailing; results are not inherently inconsistent, (208:9 (Mulligan)), and passing results on limited retesting do not rule out the possibility that the batch is not uniform.

⁹ Such a conclusion cannot be based on 3 of 4 or 5 of 6 passing results, but possibly 7 of 8. (804:17-25 (Mulligan)).

¹⁰ As part of the investigation, Barr should consider the record of previous batches, (1711:4 (Rhodes))¹ since similar or related failures on different batches would be a cause of concern (1714:8 (Rhodes)). Factors deserving particular focus include the number of passing and nonpassing results, as well as the number of batches made that did not have any problems. *Id.*

¹¹ For example, the USP makes it clear that additional testing is necessary in the case of biological assays, (1928:21, 1929:8 (Cooper) (referring to USP Ex. 219, page 1493, test monograph 81)), where two or more independent assays are required for a reliable potency estimate. (1931:3 (Cooper)).

¹² Thus, for example, assay data at the finished product stage must be viewed in light of all other data, including blend results and finished product content uniformity data, (1678:20 (Rhodes)), but firms cannot disregard failing assay data simply because the content uniformity results are acceptable. (1679:1 (Rhodes)).

5. Resampling

43. Resampling, in contrast, is a more controversial practice. Typically resampling occurs after the initial test and the retests have produced out-of-specification results, thereby indicating a more serious problem. When performing a resample, an analyst leaves the laboratory and takes a new sample from the universe of the batch. (1534:13 (Rhodes)).
44. Resampling is appropriate where provided by the USP (1044:21 (Gerraughty))¹ as in cases of content uniformity and dissolution testing. (1045:3 (Gerraughty)). Similarly, in the limited circumstances in which a failure investigation suggests that the original sample is unrepresentative, resampling is acceptable. (1024:19 (Gerraughty)). Evidence, not mere suspicion, must support a resample designed to rule out preparation error in the first sample¹³ (1955:6, 1956:1 (Cooper)).
45. Absent these limited exceptions outlined above,¹ however, firms cannot rely on resampling to release a product that has failed testing and retesting. (1045:19 (Gerraughty)).

6. Remixing¹⁴

46. The need for remixing arises during the blend stage when testing reveals problems with content uniformity. The regulations themselves allow reworking, which essentially is remixing. (906:11 (Gerraughty)). As evidenced by the Generic Drug Office directive and the consent agreement between the FDA and Eli Lilly, remixing is allowed in this circumstance. (558:2 (Mulligan)).
47. The instance of remixing, however, is directly related to its propriety. Occasional remixing is acceptable, (674:2 (Mulligan)), but frequent or wholesale remixing is not. The need to remix often provides a clear indication that the process is invalid (907:5, 1021:9 (Gerraughty)), and casts doubt on those batches that passed through testing without incident. (576:14 (Mulligan))

7. Averaging¹⁵

48. Although averaging test data can be a rational and valid approach, (1578:4 (Rhodes)), as a general rule, firms should avoid this practice, because averages hide the variability among individual test results.
49. This phenomenon is particularly troubling if testing generates both out-of-specification and passing individual results which when averaged are within specification. Here, relying on the average figure without examining and explaining the individual out-of-specification results is highly misleading and unacceptable. (812:20 (Mulligan); 1255:11 (Gerraughty)); see also (1481:20, 1482:8 (Bolton) (averaging blend assay results invites problems at finished product stage where values of 89, 89, and 92, given range of 90 to 110, should be met with more testing even though average equals 90)).
50. Although averaging camouflages variability, an average may provide more information about the batch's true assay value than any single test result.¹⁶ (1672:4 (Atwater)).

Thus, in the case of microbiological assay testing, the USP prefers an average, (1932:2 (Cooper))¹ when reaching an ultimate judgment about the quality of the product. (1929:22 (cooper)). It is good practice to include out-of-specification results in the average, unless an outlier test indicates that an out-of-specification result is an anomaly. (1579:23 (Rhodes)).

¹³ Based on the examples of sampling error provided at the hearing, see e.g. (1534:17 (Rhodes) (quality assurance personnel taking sample from first fifty tablets from tableting machine rather than retrieving tablets from the beginning, middle, and end of the run, as required)), firms should be able to identify such errors readily by reviewing the steps used to generate the sample or questioning the employee responsible for sampling.

¹⁴ Although Barr ceased remixing after the 1991 FDA investigation and the FDA found no evidence of remixing during their 1992 investigation (388:19 (Mulligan)), Barr's prior remix practice may affect the reliability of its retrospective validation studies.

¹⁵ Unless indicated otherwise, this discussion applies to assay testing only. Content uniformity results never should be averaged to obtain a passing value with the limited exception noted in ¶ 51 infra.

¹⁶ Assuming a homogeneous sample, statistically an average is more precise than the individual test results that make up the average. (1332:16 (Bolton)).

51. Finally, the average of individual content uniformity tests at the finished product stage can act as a proxy for the assay value. (1231:33 (Gerraghty)). Though this estimate cannot substitute for final product assay testing, it can provide some information about a batch.

8. Releasing a Batch for Distribution

52. Section 211.165(a), in relevant part, provides:

For each batch of drug product, there shall be appropriate laboratory determination of satisfactory conformance to final specifications for the drug product, including the identity and strength of each active ingredient, prior to release.

21 C.F.R. 211.165(a). And section 211.165(f) specifies:

Drug products failing to meet established standards or specifications and any other relevant quality control criteria shall be rejected. Reprocessing may be performed. Prior to acceptance and use, reprocessed material must meet appropriate standards, specifications, and any other relevant criteria.

53. The USP provides the "established standards" to which the CGMP regulations refer and upon which firms rely to release their drug products for distribution to the public. (486:23 (Mulligan)). These specifications are absolute. (1484:18 (Bolton)) (noting that firms cannot stretch USP -- if standard is 93 then cannot release batch on result of 91)).
54. Under section 211.165(f), the government argues, a single out-of-specification result that cannot be invalidated defeats a batch. (312:12 (Mulligan); 1271:22 (Gerraghty) (because variability of assay and instrument built into USP procedure, batch fails if have one failing result)). Barr reads the statute more liberally and refuses to equate a failure to meet specifications with obtaining one out-of-specification result from many quality performance measurements. (1465:14 (Bolton)).
55. Similarly, because the USP specifies the standards for release testing, the government contends that section 211.165(f) removes any uncertainty that might require the exercise of scientific judgment. (486:23 (Mulligan)). The absolute approach the government recommends, under which a batch fails if one tablet out of one million tablets tested produces an out-of-specification result, (1273:24 (Gerraghty)), is extreme. Instead, in the Court's view, section 211.165, through the "appropriate laboratory determination" language and the allowance of reprocessing, suggests that scientific judgment can play a role when firms decide whether to release a batch to the public.
56. In light of the case-by-case variability scientific judgment introduces, the Court cannot articulate specific procedures for release decision-making. Instead, in this context, firms should follow the general retesting guidelines. See ¶¶ 37-42. However, firms cannot justify release when only fifty or sixty-six percent of the finished product tests for a particular batch produce passing results. For example, with an assay limit of 90 to 110, test results of 89 and 90, or 89, 91, and 91, or two 89s and two 92s all should be followed by more testing. (1398:1 (Bolton)). The amount of additional tests required would be a matter of scientific judgment, as informed by other relevant data. (816:9 (Mulligan); 1398:21 (Bolton)). The goal is to distinguish between an anomaly and a reason to reject the batch. (1390:1 (Bolton)).
57. While experts disagree about the relative importance of finished product and blend testing, see (1324:9-20, 24, 1325:12 (Bolton) (blend testing serves to enhance confidence in finished product tests, but finished product assay and content uniformity testing more important); (907:17: (Gerraghty) (in-process testing more important because finished product testing very minimal)), it is clear that the release evaluation depends, in part, on the background of the batch and product. Secondary parameters,

such as physical properties, blend evaluations, time of mix, weight, thickness, and friability. affect the actual finished product results as well as their reliability. (1317:7 (Bolton)) The lesson for firms and the FDA is that context and history inform many final conclusions. (1486:1 (Bolton) (marginal blend assay results require stricter scrutiny at finished product stage); 1273:10 Gerraughty) (considering past content uniformity problems); 1368:23 (Bolton) (when reaching conclusions, important to look at all data)); 1419:25 (Bolton) (history of product useful in determining whether error is cause for concern)).

9. Blend Testing

58. An important aspect of drug manufacturing, blend testing gives firms an opportunity to discover and remedy in-process problems before batches reach the final stages of production. Because finished product testing is limited, blend testing is necessary to increase the likelihood of detecting inferior batches. (1181:1, 1221:2 (Gerraughty)); *see also* (1036:18 (Bolton) (cannot waive blend content uniformity testing and rely solely on finished product results))

(a) Sample Size

59. An element of blend testing which influences the ultimate test results is sample size. No regulation, guideline, or publication requires any specific blend sample size. (1119:15 (Gerraughty)). The CGMP regulations merely provide that "[i]n-process materials shall be tested for identity, strength, quality, and purity as appropriate." 21 C.F.R. § 211.110(c) (emphasis added).
60. In accordance with the preamble to the CGMP regulations, the Court must construe the "as appropriate" phrase to permit "reasonable, albeit variable interpretations." (765:16 (Mulligan)). The testing procedures a firm chooses, however, must be logical and effectual. (767:13, 776:14 (Mulligan)). Paraphrasing Inspector Mulligan, the Court must ask whether "what they are saying makes sense." (769:15 (Mulligan)).
61. Implicit in the "as appropriate" language, the government argues, is a sample size requirement equal to one to three times the product's run weight. (432:17 (Mulligan)). Driving this theory is the concern that a larger sample will dilute or even negate any nonuniformity in the blend. (672:1 (Mulligan)). Barr argues that a variety of sample sizes comply with CGMP and that the final choice is a matter of scientific judgment. Barr Findings ¶ 122.
62. Although sample size is a question of scientific judgment, the sample chosen must advance the purpose of the test. Thus, what is "appropriate" may vary with the type of test performed.
63. Content uniformity testing, designed to detect the adequacy of the mix by measuring variations in the potency of the blend, (1150:15 (Gerraughty)), should be conducted with a sample that resembles the dosage size. Any other practice likely would blur differences in portions of the blend, (1069:18 (Gerraughty)), and defeat the object of the test.
64. The Court appreciates the difficulty companies experience taking minute samples from large-volume blends. Indeed, testimony revealed that the smallest thief available can retrieve a 250-milligram sample, (1285:1 (Gerraughty)), so in some cases firms cannot obtain a single-run-weight sample. As such, the Court will follow Dr Gerraughty's testimony and hold that the appropriate sample for content uniformity testing in both validation and ordinary production batches, (922:15 (Gerraughty)), is three times the active ingredient dosage size. (921:4, 922:8 (Gerraughty)).

65. In addition to assuring a more accurate measure of uniformity, this rule accommodates the need for retesting. In order to conduct an initial test and two retests, a standard testing practice in the industry, (671:16 (Mulligan)), analysts need a three-run-weight sample. (671:21 (Mulligan)). Under Inspector Mulligan's one-run-weight rule, in order to retest the same sample, firms must take additional samples from the same spot in the blend. (681:12 (Mulligan)) Such a requirement would be onerous.
66. Implicit in the sample size determination for content uniformity testing is a prohibition on compositing multiple individual samples taken from different areas of the blend. Again, in order to detect uniformity problems, firms must avoid this practice which would conceal variations in the blend. (1970:6 (Cooper)).
67. In contrast, blend assay or potency testing, designed to measure the strength of the blend, can accommodate larger samples. Although averaging the differences in the mixture, here a larger sample also provides a better indication of the overall percentage of active ingredient in the blend. (1150:15 (Gerraghty)). In fact, a blend assay test conducted with a larger sample will be more representative of the final assay. (1328:19 (Bolton)). similarly, as variation detection is not the object of assay testing, the Court does not object to compositing.¹⁷

(b) Site of Sampling

68. Also important in content uniformity testing are the number of samples taken and how representative they are of the mix. (1329:7 (Bolton)). The government again cites the "as appropriate" language of section 211.110(c) to support its view that blend content uniformity testing should be conducted with samples taken from the mixer and not the drum. Barr maintains that firms are free to sample from either the mixer or the drum and, further, that in some cases, testing from the drum is preferable. Barr Findings ¶ 123. Once again, no regulation, guideline, or publication expresses a preference for blend of testing in the mixer or the drum. (1120:1 (Gerraghty))⁴
69. Expert testimony revealed that firms test from both the drum and the mixer and that either practice is acceptable under CGMP. (1012:19 (Gerraghty))¹ 1336:14 (Bolton)). Thus, the Court is not prepared to prescribe a particular location for blend testing.
70. Rather, the factor of more concern to the experts, and thus the Court, is the representativeness of the sampling technique. (1013:9 (Bolton)). To detect poor uniformity, firms must take samples from "places that might be problems," (1335:2 (Bolton)), such as weak and hot spots in the blend. (921:4, 1223:13 (Bolton)). Indeed, Barr conceded that Inspector Mulligan's concern about in-process blend testing in weak spots of the mixer is valid at the process validation stage. (2212:3 (Pierpaoli)). Thus, whether sampling from the mixer or the drum, firms must demonstrate through validation that their technique is representative of all portions and concentrations of the blend.¹⁸
71. Experts did suggest, however, that sampling from the mixer is preferable, (1012:19 (Gerraghty)), especially if firm has experienced content uniformity problems. (436:16 (Mulligan)). While this sampling method is not CGMP-required, the Court's view, it has advantages. Sampling from the mixer facilitates both remixing and targeting areas of likely nonuniformity

E. Criticisms of Barr

¹⁷ This sampling strategy is consistent with the USP parameters for finished product assay and content uniformity testing. Applying the same limits at the blend stage, the narrow testing parameters compliment a large assay sample, while wider parameters allow for the small sample used in content uniformity testing. (1331:4 (Bolton)).

¹⁸ Compare Barr's ability and willingness to identify the specific areas of its equipment that are difficult to clean in their cleaning validation studies. Barr Findings ¶ 152.

72. Having defined the terms underlying the CGMP compliance determination, the Court now can turn to the criticisms the FDA has levied against Barr's practices. Specifically, the FDA contends that: (1) Barr's manufacturing processes are invalid or not validated; (2) Barr's failure investigations are inadequate; and that (3) Barr releases drug products that fail their specifications. Also of concern to the FDA are Barr's cleaning validation and record-keeping deficiencies as well as Barr's failure to control steps in the manufacturing process, to validate its testing methods and to file ANDA supplements and field alerts. In light of these shortcomings, the government asks the Court for restrictions on Barr's current product line. The Court will discuss each criticism in turn.

1. Manufacturing Process Validation

(a) General Requirements

73. To comply with CGMP, drug manufacturers must develop "written procedures for production and process control designed to assure that the drug products have the identity, strength, quality and purity they purport or are represented to possess" 21 C.F.R. § 211.100.

74. These written procedures, in turn, must be verified. More specifically, the CGMP provides:

To assure batch uniformity and integrity of drug products, written procedures shall be established and followed that describe the in-process controls, and tests, or examinations to be conducted on appropriate samples of in-process materials of each batch. Such control procedures shall be established to monitor the output and to validate the performance of those manufacturing processes that may be responsible for causing variability in the characteristics of in-process material and the drug product. Such control procedures shall include, but are not limited to, the following, where appropriate:

- (1) Tablet or capsule weight variation;
- (2) Disintegration time;
- (3) Adequacy of mixing to assure uniformity and homogeneity;
- (4) Dissolution time and rate;
- (5) Clarity, completeness, or pH of solutions.

21 C.F.R. § 211.110(a)

75. This verification, known as process validation, is a quality control measure through which drug manufacturers become confident that their processes consistently will produce products which meet their predetermined specifications. (873:8 (Gerraughty)).

76. Process validation requires that "[i]n-process materials be tested for identity, strength, quality and purity as appropriate, and approved or rejected by the quality control unit, during the production process, e.g., at commencement or completion of significant phases or after storage for long periods." *Id.* § 211.110(c).

77. In its Guideline on General Principles of Process Validation ("FDA Guideline"), published in 1987 and re-noticed unchanged in 1990, the FDA provides further guidance to firms in the area of process validation. *See* Exhibit 143.

78. Types of in-process testing commonly performed include assay, content uniformity and dissolution tests. Assay testing measures the potency of the active ingredient in the drug product. (21:13 (Mulligan)). Content uniformity testing measures the variability of potency between drug dosages. *Id.* Dissolution testing determines how the active ingredient will dissolve and work in the human body. (25:7-15 (Mulligan)).

79. Such testing occurs at two stages in the production process, the final blend stage and the finished product stage. (26:9 (Mulligan)). Assay and content uniformity tests are performed at the blend stage, while all three tests are conducted at the finished product stage.

80. The data resulting from this barrage of testing forms the basis for the CGMP-required validation study. Firms may choose a prospective, concurrent or retrospective validation, each an accepted method under CGMP, alone or in combination.
81. Prospective validation requires a manufacturer to make at least three consecutive batches under closely-controlled conditions and to perform the studies and intensive testing necessary to determine that each step in the process is under control and will yield a product that consistently meets its specifications. This method is used for new products that have not been manufactured and marketed.
82. Concurrent validation also requires the manufacture of at least three consecutive batches under closely-controlled conditions followed by intensive testing. This method is used for existing products.
83. Retrospective validation allows a manufacturer to rely on the past production of a series of problem-free batches to show that the process is under control and yields a product that consistently meets its specifications. (223-24 (Mulligan)). Although no single batch is subject to intensified testing, making a product for a long period of time generates a history that contains much useful information about the product. (1314:11 (Bolton))

(b) Specific Process Problems

84. The government argues that Barr's manufacturing processes are invalid due to a high percentage of batch failures and, further, that Barr's attempts to validate these processes retrospectively are unavailing since Barr has excluded failing data and failing batches from their studies and relied on an insufficient number of batches.

(1) Batch Failures

85. The government first argues that the failure rates associated with Barr products demonstrate the need to revise the underlying manufacturing processes. To the extent that batches included in retrospective studies exhibit a failure rate of ten percent or more, the Court agrees. (1384:10 (Bolton))
86. Citing assay, content uniformity, dissolution, and stability test failures, among others, the government has calculated batch failure rates ranging from 2.0 to 44.4% for thirty-one of Barr's products. Because the Court has rejected the government's definition of "failure," the Court will not rely solely on the failure percentages calculated on this basis to discredit Barr's manufacturing processes.
87. Yet, the frequency with which Barr obtains out-of-specification results does give the Court pause. As such, it is with a sharp eye that the Court examines Barr's validation studies.

(2) Retrospective Validation

88. Pointing to its retrospective validation studies, Barr claims that it manufactures each of its sixty products in its current product-line under a validated process and, as a result, that it is in compliance with CGMP. The government challenges this conclusion, pointing to numerous deficiencies in Barr's testing practice.

(a) Omitting Failing Results

89. The government first argues that Barr consistently omits "failing" test results from batches included in its retrospective validation studies.

90. Only those test results determined through an appropriate failure investigation to be caused by analyst or operator error can be excluded. Those results that are not explained, but merely called into question by successful retesting, must be included.
90. Barr's former practice of rejecting initial out-of-specification results based upon acceptable retest results does not meet this standard. Currently, however, Barr considers all initial results. see ¶ 109 *infra*.

(b) Omitting Failing Batches

92. The government next argues that Barr omits failing batches from its validation studies, thereby obscuring results obtained during the specified time-frame.
93. As a general rule, a retrospective study must contain all batches produced within the time frame chosen for analysis. To avoid a comparison of apples and oranges, however, these batches must be: (1) made in accordance with the process that is being validated; and (2) representative of this process. (884:23 (Gerraghty); 2124:18 (Baldassarre)). Batches that do not meet these criteria should not be included in the retrospective report.
94. Thus, exclusion is warranted in situations in which firms have made the same product under superseded or unapproved processes.¹⁹ This occurs, for example, if a firm alters a procedure, thinking the change does not require prior approval, and later learns from the FDA that approval was needed. (2110:9 (Baldassarre)). Although made under the correct process, those batches that are aborted or rejected due solely to a nonprocess-related error should be excluded as nonrepresentative. (1450:15 (Bolton)).
95. Because the default is inclusion, (2111:15 (Baldassarre) (firm should include all batches unless it has very good explanation)), firms bear the burden of explaining and documenting any gap in the chronological series of batches included in a retrospective study. (2111:15 (Baldassarre); 1456:19 to 1457:4 (Bolton)). Firms also must indicate the process used for each batch. (1003:9 (Gerraghty)).
96. Currently, Barr selects those batches made under the same manufacturing process for its retrospective validations, (2108:12 (Baldassarre)). provided that these batches did not experience failures attributed to operator error. (2124:18 (Baldassarre)). Barr does include batches that exhibit process-related failures. (2125:10 (Baldassarre)). Barr explains operator errors and anomalous situations in footnotes, (2127:18 (Baldassarre)). but leaves un-noted any gaps resulting from different manufacturing processes. (2110:22 (Baldassarre)).
97. While Barr's selection of batches for retrospective studies is acceptable, the government's request for an explanation is well-taken. (2127:11 (Ms. Yeoh)). Barr must remedy its failure to note and explain gaps in its retrospective validations. Specifically, Barr must document the reason for exclusion, through a failure investigation if the problem is an operator error, or a written explanation of the process and its modifications if different processes are used.

(c) Insufficient Number of Batches

98. Finally, the government contends that Barr's retrospective studies are unreliable, because Barr used fewer than the required minimum number of batches.
99. The CGMP regulations and the FDA Guidelines do not discuss individual validation methods and thus provide no guidance on the issue of minimum batches for retrospective studies. (1118:3, 12 (Gerraghty)). In the context of both prospective and retrospective validation, the FDA Guidelines express a preference for specific test results, rather than simple pass/fail determinations, since statistical analysis of specific tests generates an expected variance in data. FDA Guidelines at 24.

¹⁹ Batches made with an unapproved process can be included in a prospective or retrospective validation study if the process later is approved (1653:19 (Rhodes)). Before approval, however, the study is provisional. (1654:2 (Rhodes)), and the batch cannot be released. (1655:3 (Rhodes)).

100. The Government's expert advocated a ten-batch-minimum rule, a figure the industry allegedly accepts. (1003:20 (Gerraughty)). Barr contends that with the help of statistical analysis it can validate a process retrospectively with very few batches. Barr Findings ¶ 133.
101. While any number of batches can support a validation study¹ (1439:19 (Bolton)), because validation is a matter of degree, the degree of assurance about the validity of the process necessarily increases with the number of batches. (1440:11 (Bolton)). More batches produce more reliable statistical conclusions about the process characteristics. (1346:2 (Bolton)).
102. The simple more-is-better principle does not translate easily into a specific minimum number of batches. Although the number of batches required is a matter of scientific judgment, (1345:16 (Bolton)), the hearing testimony provides some parameters in which these decisions must be made. First, the number of batches used in retrospective studies must be greater than those used in prospective validation. Less intensive testing on the retrospective batches and the inability of firms to retest batches examined retrospectively require this rule. (879:4 (Gerraughty)). Second, studies conducted with only five batches, a number Barr's expert was unwilling "to hang (his) hat on," (1345:16, 1346:16 (Bolton)), are unacceptable.²⁰ Instead, as a general rule firms should strive to include twenty to thirty batches in their retrospective studies. (1346:5 (Bolton)).
103. In the absence of any direction from the CGMP regulations or the FDA Guidelines, however, the Court is unwilling to hold that one precise number of batches is needed to complete an adequate retrospective validation study. Any such figure would be arbitrary and might encourage some firms to modify their current practice downward. In addition, an absolute rule would make it impossible to validate retrospectively those drug products with low production rates. The Court is confident that firms will act in accordance with this principle. Any decision to ignore the general rule must be supported by statistical or other evidence that indicates a reliable validation.²¹

²⁰ By concluding that studies based on five batches are unacceptable, the Court does not imply that the inverse of this statement also is true. Studies based on six or more batches may not be acceptable.

²¹ Because a 10% batch failure rate is unacceptable, see § 85 supra, if one batch fails, more than ten batches would be required for a satisfactory retrospective validation study

2. Failure Investigations

104. The government next attacks Barr's failure-investigation practice. Citing FDA criticism and investigations dating back to 1985, it alleges that until recently Barr rarely performed for-cause failure investigations. Government Findings ¶¶ 63-64. Even under the Court's more generous interpretation of the section 211.192-prescribed failure investigation, the evidence clearly indicates that Barr's past failure investigations fall far short of the CGMP mandate. (1513:4 (Rhodes) (concern about failure investigation documentation and product mix-ups); 1536:8 (Rhodes) (Barr not documenting out-of-specification results and decision to resample); 1991:19 (Cooper) (inadequate documentation prevents judgment about procedures and investigations proper)). At best, Barr carried out some investigations, but did not document its efforts (1953:25 Cooper)).
105. The testimony of Barr experts and employees, however, suggests that Barr recently has begun to conduct for-cause inquiries in a timely and adequate fashion (1046:12 (Gerrughty); 1515:1, 1551:6, 1152:24, 1553:2,20, 1555:6, 1557:18, 1573:17, 1574:20 (Rhodes)); see also (1519:19) (Rhodes) (Barr has established central file for failure investigation reports)). A step-by-step critique of Barr's current failure investigation practice follows. See ¶¶ 107-115 infra.

3. Release of “Failing” Batches

106. The government, in a three-part argument also contends that Barr has and continues to release batches of product which fail their predetermined specifications. Specifically, it claims that Barr: (a) discards failing results based on an assumption of laboratory error; (b) without exception discards failing results based on passing retest results, a procedure that is statistically and scientifically invalid; and (c) discards initial results after retesting different parts of the batch. Barr denies each allegation.

(a) and (b) Rejecting Failing Results As Laboratory Errors After Retesting

107. The government first focuses on Barr's alleged procedure of rejecting initial out-of-specification results as laboratory error based solely on passing retest results. Although separate criticisms, the complaints aimed at Barr's failure investigations²² and retest procedure are closely related and best considered together. The Court will conduct a step-by-step analysis of Barr's procedures for chemical and microbiological assays in turn.
108. Currently, Barr meets out-of-specification results from chemical tests with a laboratory investigation. Barr Findings ¶¶ 54-55. If this limited inquiry successfully identifies a laboratory error, Barr rejects the initial results. *Id.* ¶ 55. Consistent with the Court's interpretation of section 211.192, this procedure complies with CGMP.
109. If the laboratory investigation is inconclusive, as is sometimes the case, two additional analysts retest the same sample. If these retests are within specification, Barr averages the three test results and releases the batch if the average also is within specification. *Id.*; Ex. 14 at 5; (405:22, 685:13, 1040:18 (Mulligan); 2157:23 (Baldassarre)). Although retesting is preceded by an investigation as CGMP requires, this procedure nevertheless is in direct conflict with expert testimony. As the Court noted previously the amount of retesting required, a matter of scientific judgment, will vary on a case-by-case basis depending on the history of the product and the batch. See ¶¶39-42 supra. Thus, Barr cannot rely on its two-out-of-three retest/release procedure.
110. Next, “[i]f the results appear to be inconsistent the laboratory performs the appropriate outlier test”. Whether this phrase contemplates a comparison between the retest results and the initial result, or between the retest results themselves, the use of an outlier test in the chemical testing context is impermissible. See ¶¶ 35-36 supra.
111. Finally, if the average of the results is out-of-specification, Barr proceeds to resample in order to determine whether there was a sampling error and begins a formal investigation. Barr Findings ¶ 56.

²² Adequately discussed above, the Court will merely reiterate that Barr's former failure investigation practice was insufficient, since for-cause failure inquiries must follow each out-of-specification result. ¶¶ 104-05 supra.

Again, this reflexive approach is highly suspect. (1244:21 (Mulligan)). Because resampling is appropriate only in limited circumstances, the Court cannot condone this resample rule. (1245:5 (Mulligan) not practice of industry)). No action can be taken with respect the batch until the results of the formal investigation are known.

112. The problems with Barr's retesting procedure are not erased by the protocol's popularity at other firms. See Barr Findings ¶ 57. As the Court indicated earlier, industry practice while instructive also is subject to the CGMP-compliance test. See ¶ 12 supra. With the exception of the testimony of Dr. Rhodes, (1542:1-11; 1544:22; 1666:15; 1669:7 (Rhodes)), the record contains resounding condemnation of Barr's retesting protocol. (1041:6 (Gerraughty) (two of three chance that passing result is correct insufficient); 1044:14 (Gerraughty) (this practice not standard in industry or in compliance with CGMP); 1476:1 (Bolton) (cannot discard out-of-specification result based on two acceptable retests of same sample); 1476:9 (not a good practice in every case as need to consider all data);²³ 1477:15 (Bolton) (specifically rejecting TM-196C); 1244:13 (Gerraughty) (that no discussion of results themselves problematic in that procedure would allow firm to pass batch based on results of 89, 91, 91, which is inappropriate; 1246:16, 22 (Mulligan) (if all results are within same precision of instrument, "real gamble" to assume that true value is 91)).
113. Barr's microbiology protocol provides for additional testing using aliquots from each of the three original containers if the initial results are out-of-specification. Barr Findings ¶ 58. If all individual retests are within specification, and the average of the initial and retest results is within specification, the batch passes. Id. ¶ 59; (1946:9 (Cooper)).
114. If, however, any retest results are out of specification, Barr proceeds to resampling. (1949:1-12 (Cooper)). If the resampling results and the average value of the initial tests, retests and resamples is within specification, the batch passes. Id.
115. Given the variability inherent in microbiology testing, Barr's retesting procedure is appropriate,²⁴ (1955:8 (Cooper)), provided that Barr continues to use the outlier test to discredit out-of-specification results, (2192:7 (Pierpaoli)), and does not return to its former practice of simply disregarding initial out-of-specification results. (2191:21 (Pierpaoli) (Barr previously threw away initial result without evaluating it); 1951:1 (Cooper) (Barr previously would not include average of initial test results)). However, as the Court has noted, firms cannot remedy test and retest problems through reflexive resampling. See ¶ 44-45 supra.

(c) Rejecting Failing Results After Resampling

116. Here, the government focuses its concern on Barr's practice of conducting retests using samples from a different 20-tablet grinds taken from the same laboratory sample. (2191:19 (Pierpaoli)). Instead, the government contends, Barr should retest the sample used in the initial test. Expert testimony, however, determined that both of these procedures properly are termed a "retest."

²³ In his treatise, Dr. Bolton also rejects similarly inflexible retesting rules, specifically: (1) conducting three assay tests and then selecting the two results that are closer together; and (2) conducting two assay tests, and if they disagree, conducting one more to determine which of the first two test results should be discarded. (1487:14 (Bolton)).

²⁴ With respect to microbiological assay, the USP speaks in terms of a "combined result" or average (1932:4 (Cooper)), and provides that: "two or more independent assays are required for a reliable estimate of the potency of a given assay preparation or unknown," (Ex 219, page 1493; (1931:3 (Cooper)), and further that if the second assay result differs significantly from first the test, firms should conduct one or more additional assays. (1931:11 (Cooper)).

117. Of more concern to the Court is Barr's practice of resampling¹ discussed supra. Although Barr changed its operating procedure to alleviate FDA criticisms, leaving out the paragraph that allowed resampling when a reconstitution error or a sample integrity problem was suspected, (2079:15 (Baldassarre)), a Barr employee testified that resampling, a necessary element of a failure investigation, (2152:19 (Baldassarre)), would continue. To the extent that Barr contemplates a greater role for resampling than does the Court, see ¶¶ 43-45 supra, its practice must be revised.

4. Failure to Control Manufacturing Process Steps

118. The government provides a laundry list of Barr manufacturing processes that are deficient. Specifically, it challenges Barr's failure to validate physical characteristics of inactive ingredients, to conduct optimum mixing time studies, to draw representative blend samples, to control particle size, particle size distribution and the physical characteristics of tablets and capsules as well as Barr's sample size practices. Government Findings ¶¶ 91-99.

(a) Barr Blend Testing

119. Currently Barr conducts blend testing from a greater number of drums,²⁵ (2199:14 (Pierpaoli)), retrieving samples from the top, middle, and bottom of the drums at positions corresponding to 12:00, 4:00 and 8:00 o'clock. Barr combines the resulting nine samples and then splits this composite into test and reserve samples (2200:11 (Pierpaoli)). With respect to their testing parameters,²⁶ Barr has tightened the blend range specification, (2202:5 (Pierpaoli))¹ and thereby narrowed the variability in the beginning of the process to facilitate corrective action. (2203:1 (Pierpaoli)).

(1) Site of sampling

120. Barr's decision to retrieve blend samples from the drum rather than the mixer grows out of its concern that the movement of the blend from the mixer to the drum might invalidate tests based on mixer samples, because moving blends and powders can change their physical characteristics." (2205:4 (Pierpaoli)). As a result, Barr contends, it is better to defer testing to the end of the manufacturing process, (2206:9 (Pierpaoli)), because a firm cannot expect earlier stage testing to control a later stage. (2210:16 (Pierpaoli)).

121. This argument proves too much. Under Barr's theory, only tests of the product in final form would be reliable. Barr itself concedes that it would be inappropriate to eliminate in-process blend testing, since, in their view, any gross nonuniformity should be detected. (2208:1 (Pierpaoli)). Further, Barr's concern that mixtures may declassify upon handling, (2226:17 (Pierpaoli)), can be dealt with through particle size distribution studies. Particle size distribution affects both mixing and content uniformity. (892:21-893:3 (Gerraughty)).

122. Through content uniformity testing, firms must attempt to discover even subtle uniformity problems at the blend stage. Deferring detection of these inconsistencies to the finished product stage, (2203:13 (Pierpaoli)), is inappropriate. To facilitate this goal, firms must sample from representative portions of the batch as outlined above, whether from the drum or mixer. See ¶¶ 68-71 supra. If Barr continues to draw samples from the drums, it must demonstrate the representativeness of this testing technique.

²⁵ Before Dr. Pierpaoli's arrival, Barr tested only three drums. Barr now tests a minimum of four drums. (2199:14-25 (Pierpaoli)).

²⁶ The USP does not specify acceptable ranges for blend testing.

(2) Sampling Procedure

123. Barr's practice of compositing samples also frustrates attempts to detect subtle blend differences. This combination practice yields an assay value that is representative of the contents of the entire drum, (2200:23 (Pierpaoli)), and therefore does not provide information about intra-drum variability. (2203:13; 2223:1, 12 (Pierpaoli) (not measuring within drum nonuniformity)). Because this procedure does not address possible weak and hot spots in the blend, (2210:16 (Pierpaoli)), Barr must abandon this compositing practice. Cf. 21 C.F.R. § 211.84(c)(4) (prohibiting compositing of samples used for raw material testing). The Court notes that this requirement is not onerous since Barr currently retrieves nine samples, no additional sampling is required. This requirement simply maximizes the information Barr's blend testing can yield.

(3) Sample Size

124. Currently, Barr takes 20-gram blend samples. (1330:9). Although this sample size is acceptable for assay testing, as the Court noted above, see ¶¶ 59-67 *supra*, the proper blend sample size for content uniformity testing is three times the run weight of the finished product. Thus, for the latter tests Barr must retrieve much smaller samples.

(b) *Mixing Time*

125. Section 211.110(a)(3) lists "adequacy of mixing to assure uniformity and homogeneity" as a control procedure that "shall be established to monitor the output and to validate the performance of those manufacturing processes that may be responsible for causing variability." 21 C.F.R. 211.110(a), (a) (3). The government argues that this provision requires Barr to validate its mixing time. (919:4 (Gerraughty)).
126. The language of the regulation seems to require only that an adequate mixing time be specified. Generally, mix times are specified in the firm's ANDAs. (1155:13 (Gerraughty)). Expert testimony revealed that when firms conduct retrospective validation, they may rely on either passing finished product content uniformity tests, (1325:22 (Bolton)), or mixing times used in prior successful batches, (1154:13; 1157:13 (Gerraughty)), as a form of proxy to satisfy the mixing time requirement needed for proper validation. However, a time of mix study should be included in a prospective validation program, see Nash Declaration, ¶ 35(b), and also follow any problems that surface in retrospective validation batches. (1157:17 (Gerraughty)).

(c) *Particle Size Distribution*

127. Neither the USP nor the CFR contains particle size specifications for active ingredients. (2097:3 (Baldassarre)). currently there is no. particle size specification for excipients or inactive ingredients. (1630:6-18 (Rhodes) (witness is USP excipient subcommittee member)).
128. Particle size and particle size distribution checks occur at the raw materials stage, and also during the manufacturing process if problems arise. (917:12 (Gerraughty)); 21 C.F.R. § 211.84.
129. Generally, firms can rely on the certification of their wholesaler with respect to particle size, provided they conduct spot checks for verification. (914:16 (Gerraughty)). Nevertheless, it is a widely accepted practice in the industry to include particle size distribution in a validation study. (893:16-21 (Gerraughty)).
130. Barr's raw materials, both active and inactive, have been the subject of ANDA application and approval. (411:25 (Mulligan)). Raw material testing at Barr is acceptable for quality control and validation purposes. (1317:19, 1318:21 (Bolton)).

5. **Failure to Validate Testing Methods**

131. An analytical method is a laboratory procedure which measures an attribute of a product or raw material by generating a numerical value for comparison with a prescribed specification. (1747:18 (Atwater)).
132. Analytical methods validation is the process by which firms prove that their chosen method is reliable and capable of performing the laboratory analysis it is designed to accomplish. (1759:12 (Atwater)). The aspects of a method that should be confirmed include accuracy, precision, linearity, specificity and ruggedness. (1763:1 (Atwater)).
133. Section 211.165(e) of the CGMP regulations governs method validation, providing:

The accuracy, sensitivity, specificity, and reproducibility of test methods employed by the firm shall be established and documented. Such validation and documentation may be accomplished in accordance with § 211.194(a)(2).

Section 211.194(a)(2), in turn, provides:

- (a) Laboratory records shall include complete data derived from all tests necessary to assure compliance with established specifications and standards, including examinations and assays, as follows:
- (2) A statement of each method used in the testing of the sample. The statement shall indicate the location of data that establish that the methods used in the testing of the sample meet proper standards of accuracy and reliability as applied to the product tested. (If the method employed is in the current revision of the United States Pharmacopoeia, . . . or in other recognized standard references, or is detailed in an approved new drug application and the referenced method is not modified, a statement indicating the method and reference will suffice.) The suitability of all testing methods used shall be verified under actual conditions of use.
134. In addition, the USP contains guidelines for analytical methods validation, including specific recommendations as to how validation should be completed (1760:2 (Atwater)). Thus, a firm can establish that its methods are validated in one of three ways:(1) if the method was approved as part of its ANDA; (2) if the method is the same as that used in the current revision of the USP; or (3) if a firm validates its method through a validation study. Government Findings ¶ 101; see also (1762:12 (Atwater)). If the method falls into one of these three categories, firms need only show that the method works under the conditions of actual use.
135. However, if firms either adopt methods the USP does not recognize or modify USP procedures, they must validate these methods. (726:10 (Mulligan)). Systems suitability data alone is insufficient for validation. (1860:21 (Atwater)).
136. Barr has demonstrated that each of its analytical methods is either prescribed in the USP or the CFR or the subject of ANDA approval. See Barr Findings, Appendix I. The government, nevertheless, disputes the reliability of Barr's methods, noting that under the revised USP, which requires method precision data, Barr's ANDA applications are incomplete.
137. Because Barr's analytical methods conform to the USP, the CFR or Barr ANDAs, they are entitled to a presumption of validity. The government's reservations, however, cannot be disregarded lightly. Although Barr attempted to address these concerns through validation studies submitted during trial, the government has been unable to evaluate these studies since they do not contain raw data. Hill Declaration, ¶ 18.

138. Barr must supply the government with the raw data used to prepare the validation studies. After the government has had an opportunity to review this information, the government may voice further objections. In light of Barr's former willingness to update its methods in conformance with USP revisions, *see e.g.*, Barr Findings, Appendix I, at 16, 65, 73, 76, 98, the Court is confident that the parties will be able to resolve any subsequent disagreements without judicial intervention.

6. Cleaning Validation Deficiencies

139. CGMP requires that drug makers develop and verify procedures for cleaning their manufacturing equipment.²⁷ Specifically, section 211.67 provides:

- (a) Equipment and utensils shall be cleaned, maintained, and sanitized at appropriate intervals to prevent malfunctions or contamination that would alter the safety, identity, strength, quality, or purity of the drug products beyond the official or other established requirements.
- (b) Written procedures shall be established and followed for cleaning and maintenance of equipment, including utensils, used in the manufacture, processing, packing, or holding of a drug product. These procedures shall include, but are not necessarily limited to, the following:
 - (1) Assignment of responsibility for cleaning and maintaining equipment;
 - (2) Maintenance and cleaning schedules, including, where appropriate, sanitizing schedules;
 - (3) A description in sufficient detail of the methods, equipment, and materials used in cleaning and maintenance operations, and the methods of disassembling and reassembling equipment as necessary to assure proper cleaning and maintenance;
 - (4) Removal or obliteration of previous batch identification;
 - (5) Protection of clean equipment from contamination immediately before use;
 - (6) Inspection of equipment for cleanliness immediately before use.
- (c) Records shall be kept of maintenance, cleaning, sanitizing, and inspection as specified in §§211.180 (general requirements for records and reports) and 211.182 [equipment cleaning and use log].

21 CFR § 211.67.

140. Repeated FDA criticism of Barr's failure to validate its cleaning procedures leads the Government to challenge Barr's compliance with this section of the regulations. Barr points to its cleaning validation studies produced at trial which the government in turn argues are incomplete.

141. Once again, the passage of time greatly has benefitted Barr. Although it is unlikely that Barr's former cleaning procedures complied with CGMP, (Rhodes 1606:8, "they met minimal standards"), Barr's current efforts, at least since June 1992, are much improved. (1606:9; 1604:24 (Rhodes) (operating procedures for cleaning and maintenance are "fully appropriate"); 1605:2, 13 (Rhodes) (approving cleaning and maintenance logs)). Barr has revised its cleaning methods and developed written cleaning validation procedures. *See* Torres Declaration ¶¶ 2-8.

²⁷ Barr disputes the need to verify cleaning processes. Barr Findings ¶ 150. In order for the cleaning rules to be effective, however, the specific methods chosen must be shown to be effective. Firms cannot wait for contamination or product mix-ups to reveal inadequate procedures.

142. In addition, to facilitate its cleaning efforts, Barr has remodeled its physical plant, has created self-contained manufacturing rooms and installed waterproof lighting. (1607:17 (Rhodes)); Torres Dec. ¶¶ 5, 9. These and other physical improvements greatly have reduced the likelihood of mix-ups. (1612:1-9 (Rhodes)).
143. The government insists that these efforts fall short of the mark. In particular, it argues, the cleaning validation studies are insufficient since Barr has not: (1) identified the cleaning agents used; (2) tested for cleaning agent residues; or (3) demonstrated the reproducibility of the cleaning process. Government Findings ¶ 109. In addition, the government complains, Barr did not produce a cleaning validation study for a milling machine. *Id.*
144. Because section 211.67 applies to "equipment" generally, Barr cannot find regulatory support for its decision to validate only "major pieces of equipment." Barr Findings ¶ 150. As such, Barr must produce a study for its milling machine.
145. With respect to the identification of cleaning agents, section 211.67 leaves little room for argument, requiring a "description in sufficient detail" of the methods and materials used for cleaning. Barr therefore must identify the agents used.
146. The remaining criticisms of the government are without merit. Presumably, the identification requirement reduces the need to test for cleaning agent residues, unless Barr chooses to clean with a substance known to have such difficulties. Further, one run-through of the cleaning procedure, in the absence of problems, is not insufficient for validation.

7. Record-Keeping Deficiencies

147. The detail in which the CGMP record-keeping provisions are written highlights the importance of records in the drug industry. Indeed, satisfactory record-keeping practices enhance the ability of Barr to manufacture drugs as well as the FDA's attempts to monitor these efforts.
148. The government voices numerous strenuous objections to Barr's records, characterizing them with adjectives such as unreliable and inconsistent. Government's Findings ¶ 116. Of particular concern are the government's allegations that Barr has misplaced records, recorded test data on scrap paper, and included only averages of test results in batch records.²⁸ *See* Government's Findings ¶ 111, 113, 114.
149. Hearing testimony revealed a transformed Barr. Indeed, as of October 1991 Barr's laboratory records were found comprehensive and complete. (162:11-25 (Rhodes)) (conclusion based on review of Barr's laboratory records and operating procedures).
150. However, the use of averages should be avoided except in the limited circumstances outlined by the Court. *See* ¶¶ 48-51 *supra*. Similarly, the Court expects that Barr will record individual test results, even in those situations in which averages also are useful.
151. Test data must be recorded in laboratory notebooks. The use of scrap paper, even if its contents also are transcribed in laboratory notebooks, see Baldassarre Dec ¶¶ 48-49, is highly irregular and inappropriate. (1734 (Rhodes)). Currently, Barr's written testing procedures expressly forbids analysts from using loose or scrap paper to record test results. *See* Decker Dec. ¶ 17.

²⁸ The Government also cites Barr's destruction of Management Quality Review ("MQR") Minutes and its failure to reconcile raw material inventories as examples of weak record-keeping. Government's Findings ¶ 112, 116. While not admirable, the practice of destroying these records is not proscribed in the CGMP record-keeping regulations. Barr maintains that its MQR Minutes are internal, self-evaluative records. Since the government has presented no evidence to the contrary, (619:20, 621:4, 622:8 (Mulligan)), their preservation is within the firm's discretion.

With respect to raw material reconciliation, the CGMP record-keeping rules are similarly silent. Barr maintains inventory records for each drug component, container and closure, and reconciles each lot of the drug components. *See* Decker Dec. ¶ 31.

Finally, the government objects to Barr's failure to describe the sample used for retesting. Having found that additional testing from the initial test sample and additional testing from the laboratory sample both are characterized properly as retests, see ¶ 37 *supra*, Barr's failure to distinguish between these two samples is not a violation of § 211.194(a)(1) which requires a "description of the sample received for testing." The "retest" designation satisfies this requirement.

8. Failure to File ANDA Supplements and Field Alerts

152. The government contends that Barr has adopted manufacturing procedures that vary from those specified in their ANDAs~without filing a supplement or obtaining FDA approval. In addition, it alleges, Barr's field alert practice is unsatisfactory. Government Findings ¶¶ 120, 122.

(a) ANDA Supplements

153. Drug manufacturers may alter ANDA-approved procedures. If the proposed change is significant, firms must file an ANDA supplement and obtain FDA approval. (1649:17 (Rhodes)). Minor changes, however, can be adopted without formal procedure, as long as notice is given to the FDA at time of the change or in an annual report. See 21 C.F.R. § 314.70.

154. The Court is unwilling to intervene in an area in which the categorization of changes as significant or minor often is the subject of industry dispute. (2111:17 (Baldassarre)). Barr has an incentive to adopt a conservative approach here. Otherwise, if Barr regularly includes changes that the FDA may find significant in an annual report, it runs the risk of being unable to use the batches made under altered, but unapproved, processes.

(b) Field Alerts

155. If the firm discovers a failure in a previously distributed product, it must file an ANDA field alert, within three working days after it experiences a failure, such as degradation or contamination problems or failure to meet specifications (106:2 (Mulligan)). The purpose of the field alert is to notify the FDA about potential problems with marketed products. (106:11 (Mulligan)).

156. Having rejected the government's definition of failure, the Court necessarily is reluctant to conclude that Barr's field alert practice was deficient. In any event, Barr has agreed to comply with the three-working day notification rule. See Solas Declaration, Ex. 1 at 1, observation 4.

II. CONCLUSIONS OF LAW

A. Jurisdiction

1. The Court has jurisdiction over the parties and the subject matter of this action pursuant to 21 U.S.C. § 332(a) and 28 U.S.C. §§ 1331 and 1345.

B. Standard for Injunctive Relief

2. Before a preliminary injunction can issue, the movant must demonstrate: (1) that it is reasonably likely that he will succeed on the merits of the litigation; and (2) that he will suffer irreparable injury if injunctive relief is denied. See Instant Air Freight Co. v. C.F. Air Freight, Inc., 882 F.2d 797, 800 (3d Cir. 1989). In addition, the Court should consider the possibility of harm to the parties and other interested persons from the grant or denial of injunctive relief as well as the interests of the public. Id.
3. A different standard applies in cases "where the public interest [is] directly affected," Id. at 803, especially if "the public interest in question has been formalized in a statute." Id. For example, enforcement of a federal statutory provision "by the agency charged with that duty may alter the burden of proof of a particular element necessary to obtain injunctive relief." United States v. Odessa Union Warehouse Corp., 833 F.2d 172, 175 (9th Cir. 1987); see also Instant Air Freight, 882 F.2d at 800 ("authorizing preliminary injunctive relief upon a showing of probable cause to believe that the statute is being violated may be considered a substitute for a finding of irreparable harm for purposes of a preliminary injunction issued under Rule 65" (citing Government of Virgin Islands v. Virgin Islands Paving, Inc., 714 F.2d 283, 284 (3d Cir. 1983))); United States Postal Service v. Beamish, 46 F.2d 804, 806 (3d Cir. 1972) (granting preliminary injunction upon "a showing of probable cause to believe" statute violated).²⁹

C. The Federal Food, Drug and Cosmetic Act

4. In the instant case, Congress codified the public interest in safe and effective drugs with the Act and expressly granted the Court injunctive powers to protect this interest. Section 332(a) of the Act provides:

The district courts of the United States . . . shall have jurisdiction₁ for cause shown, and subject to the provisions of section 381 (relating to notice to opposite party) of Title 28, to restrain violations of section 331 of this title₁ except paragraphs (h)-(j) of said section. 21 U.S.C. § 332(a).
5. Section 331(a) prohibits: "The introduction or delivery for introduction into interstate commerce of any food, drug, device, or cosmetic that is adulterated or misbranded." 21 U.S.C. § 331(a).
6. A drug is deemed adulterated if:

[T]he methods used in, or the facilities or controls used for, its manufacture, processing, packing, or holding do not conform to or are not operated or administered in conformance with current good manufacturing practice to assure that such drug meets the requirements of this chapter as to safety and has the identity and strength, and meets the quality and purity characteristics, which it purports or is represented to possess. 21 U.S.C. § 351(a)(2)(B).
7. Similarly, the "failure to comply with any regulation set forth in this part (21 C.F.R. 210) and in parts 211 through 226 of this chapter in the manufacture, processing, packing or holding of a drug shall

²⁹ However, in requiring courts issuing injunctions under the Clean Water Act to apply traditional equitable standards, the Third Circuit suggested a movement away from this dual standard. See Natural Resources Defense Council, Inc. v. Texaco Refining & Marketing, Inc., 906 F.2d 934, 937 (3d Cir. 1990) (following Weinberger v. Romero-Barcelo, 456 U.S. 305, 313, 102 S. Ct. 1798, 1803 (1982)); International Union v. Amerace Corp., 740 F. Supp. 1072, 1086 (D.N.J. 1990). Faced with direct threats to public safety, district courts in the Third Circuit, nevertheless, have continued to make a distinction between statutory and general equity injunctions. See United States v. Toys "R" Us, Inc., 754 F. Supp. 1050, 1053 (D.N.J. 1991) (issuing injunction under Consumer Product Safety Act); United States v. Richlyn Laboratories, Inc., slip op. at 10-11 (E.D. Pa. October 1992) (No. 92-5464) (granting injunction under Federal Food, Drug and Cosmetic Act)

render such drug to be adulterated under section 531(a)(2)(B) of the act and such drug, as well as the person who is responsible for the failure to comply, shall be subject to regulatory action." 21 C.F.R. § 210.1(b). Thus, the relevant inquiry is whether individual drugs are "adulterated" under the law, not whether such drugs are "pharmaceutically perfect" See United States v. Lit Drug Co., 333 F. Supp. 990, 998 (D.N.J. 1971).

8. Preliminary injunctions based on violations of section 331 "do not require a showing of immediate and irreparable injury." United States v. Spectro Foods Corp., 544 F.2d 1175, 1181 (3d Cir. 1976); United States v. Premo Pharmaceuticals Labs. Inc., 511 F. Supp. 958, 977 (D.N.J. 1981) (courts have power to enjoin present and future violations of section 331 "solely on the basis that such violations have been established"); United States v. Nutrition Service Inc., 227 F. Supp. 375, 389 (W.D. Pa. 1964) (citing United States v. Ingersoll-Rand Co., 320 F.2d 509, 524 (3d Cir. 1963) (enjoining violation of Clayton Act without requiring government to show precise way in which violation might injure public interest)), aff'd, 347 F.2d 233 (3d Cir. 1965); see also Odessa, 833 F.2d at 174-75 (no showing of irreparable harm necessary where injunction authorized by section 332(a)); United States v. Diapulse Corp. of America, 457 F.2d 25, 28 (2d Cir. 1972) (showing of precise way in which statutory violation harms public not required to support injunctive relief under FC&D Act). However, if a court does not base its order on a particular violation of the Act from which injury may be presumed, "there must be an independent showing of irreparable harm to warrant its issuance." Spectro Foods, 544 F.2d at 1181.
9. Courts entering injunctions under the Act have required the government to show: (1) violations of the Act on the part of the defendant; and (2) a cognizable danger of recurrent violations. See e.g., United States v. Vital Health Products Ltd., 786 F. Supp. 761, 770 (E.D. Wis. 1992); United States v. Baxter Healthcare Corp., 712 F. Supp. 1352, 1355 (N.D. Ill. 1989), aff'd, 901 F.2d 1401 (7th Cir. 1990); Richlyn, slip op. at 10-11 (No. 92-5464).
10. Factors to consider when determining whether there is a reasonable chance of future infractions include:
 - (1) the degree of scienter involved on the part of the defendant;
 - (2) the isolated or recurrent nature of the infraction;
 - (3) the defendant's recognition of the wrongful nature of his conduct;
 - (4) the sincerity of defendant's assurances against future violations;
 - (5) the nature of defendant's occupation. It is deemed important to consider as well the defendant's voluntary cessation of challenged practice, the genuineness of the defendant's efforts to conform to the law, the defendant's progress toward improvement and the defendant's compliance with any recommendations made by the government.

Toys "R" Us 754 F. Supp. at 1058-59 (citations omitted); see Richlyn, slip op. at 11-12 (No. 92-5464) (same).

11. Even if the Court were to find that Barr violated the Act in the past and that future violations are likely, an injunction automatically will not issue. Because the language of section 332(a) is not mandatory, the Court retains discretion to grant or deny equitable relief. As one court stated:

The grant of authority to issue an injunction is not the same as a mandate that an injunction issue whenever the government proves a violation and cognizable danger of recurrence. The court must employ its sound discretion in imposing all equitable remedies, including those which Congress has authorized.

Baxter Healthcare, 712 F. Supp. at 1355. Thus, the Court also must determine whether injunctive relief is needed to remedy the threatened violations

D. Evaluating Barr Laboratories

12. The Court concludes that the government has demonstrated that it is likely to succeed on the merits³⁰ With regard to past violations, there can be no dispute that Barr has violated the Act by failing to follow manufacturing practices that comply with CGMP as required under section 351(a) (2) (B) and, therefore, has introduced adulterated drugs into commerce in violation of section 331(a). Until recently, for example, Barr did not conduct failure investigations, released batches on the basis of selective data, and refused to validate its cleaning processes, thereby ignoring specific provisions of the CFR. See 21 C.F.R. 211.165(a), 211.192, 211.67, respectively. Barr's own attempts to remake itself through a vigorous overhaul prevents any other conclusion.
13. Turning to future violations, the government has demonstrated that a threat of recurrence exists, as a consideration of the appropriate factors illustrates. See ¶ 9 supra. First, defendants' violations properly are characterized as "recurrent" and not "isolated." Problems at Barr persisted despite repeated criticism from the FDA over at least a four-year period.
14. Barr's reluctance to ameliorate its methods in the face of these warnings is troublesome. This behavior requires the Court to attach a greater degree of scienter to Barr's actions. Further, Barr's refusal to comply with the recommendations of the government casts doubt on Barr's recognition of the wrongful nature of its conduct as well as the genuineness of its efforts to conform to the law.
15. While defendant has made many improvements, these efforts are long overdue. Because the threat of this litigation served as the catalyst for Barr's renovation, the Court cannot with confidence conclude that Barr's current efforts are the product of a new philosophy rather than a reflection of a desire to deflect this suit. See Odessa, 833 F.2d at 176 (noting that "[c]ourts must beware of attempts to forestall injunctions through remedial efforts and promises of reform that seem timed to anticipate legal action").
16. Due to established past violations and the risk of future violations, the Court now must consider whether an injunction is required under the particular facts of this case.
17. The Court concludes that injunctive relief is necessary to safeguard the public interest. Many of the practices the Court condemns today are used in the day-to-day operations of Barr and memorialized in standard operating procedures. Examples include Barr's blend sampling strategy, retesting procedure,³¹ outlier technique and reliance on averaging. Only through an injunction can the Court be confident that these forbidden methods, defended with vigor by Barr's employees, will be abandoned and the products made under their auspices shielded from the public.
18. Nor does any threatened harm to Barr prohibit injunctive relief. See Finnerty Declaration, ¶ 7. The interests of Barr, specifically its shareholders and employees, do not outweigh the health and safety of the public. See Diapulse, 457 F.2d at 29 (injunction which puts defendant out of business not impermissible).

E. Fashioning a Remedy

(1) General CGMP violations

19. The government cites Barr for general CGMP-compliance problems. Although the Court recognizes that Barr has had much difficulty satisfying the often reasonable demands of the FDA, injunctive relief

³⁰ With this finding, the government is entitled to a presumption of irreparable injury. United States v. NutriCology, Food, Drug, Cosm. L. Rep. (CCH) ¶ 38,298, 38,933 (9th Cir. Dec. 29, 1992) (No. 91-16419).

³¹ The retesting protocol Pierpaoli seeks to implement, see Ex. 274, TM-387, which provides for the adoption of retest results as the batch value if the initial out-of-specification result is statistically discrepant and thus discountable, (2215:6 (Pierpaoli)), also is unacceptable in its current form.

must be used sparingly, to prevent future harm, and not to punish past violations. See SEC v. Bonastia, 614 Fe2d 908, 912 (3d Cir. 1980).

20. In light of Barr's recent makeover, both personal³² and physical, the Court is unwilling to order a temporary shut-down. While Barr's transformation from an ugly duckling to a swan is neither natural nor complete, the Court cannot ignore Barr's remedial efforts, as reflected in the satisfaction of many of the concerns of its experts. See e.g., Nash Dec., ¶¶ 24, 36.

(2) Product Validation

21. Of more concern, however, are the specific products. To the extent that Barr relied upon investigations which do not satisfy section 211.192, as construed by the Court, to release batches or to complete retrospective and prospective validation studies, these actions and studies are invalid. Reliance on faulty methods cannot be cured by subsequent compliance.
22. Similarly, to the extent that Barr's validation studies and product release decisions rest on their current retesting practice, they are invalid. Only if these criticisms are addressed Barr will have an acceptable protocol. (1043:19 (Gerraughty)).³³
23. Barr cannot rely on the claim that the FDA previously approved their procedures in ANDAs, as a defense to these criticisms. An ANDA cannot shield a process that generates failures. In order to comply with CGMP, firms must correct any process that demonstrates its own inadequacies in practice. (1038:19 (Gerraughty)). ANDA approval often is obtained on the basis of a smaller batch, and, not surprisingly, the move to full production may reveal problems. (1039:2 (Gerraughty)).
24. Similarly, the control charts created by Dr. Bolton and included in Barr's retrospective studies are insufficient adequately to support their products.³⁴ These control charts were created based on average test result figures, without any review of the underlying data. (1460:6-11 (Bolton)) In addition, Bolton himself conceded that "statistics in itself is absolutely not a validation process," (1364:10 (Bolton)), but rather that statistics merely aids in the evaluation of validation data. (1364:13 (Bolton)).
25. Based on these findings, the Court will order Barr to validate its products. This order will reach only those products of particular concern to the government.³⁵

³² The Court, however, must point out that it has little confidence in Barr's decision to place the responsibility of the final failure investigation sign-off on the shoulders of an employee who has no science background or training. (2290:14 (Solas))

³³ Barr's sampling strategy and improper reliance on averaging, outlier tests and resampling provide additional support for the rejection of Barr's validation studies.

³⁴ Dr. Bolton created two types of charts for Barr's validation studies. The assay chart contains the average assay value for each batch plotted in sequence, (1339:10 (Bolton)), and the so-called range chart for content uniformity plots the difference between the highest and lowest values for each batch (1345:1 (Bolton)). From this data, Bolton is able to generate control limits, (1340:21 (Bolton)), which represent the limits within which the process is expected to perform (1341:5 (Bolton)). The conclusions drawn from these charts are independent of the size of batch produced (1358:19 (Bolton)).

³⁵ While the government based its validation requests on the failure percentages it calculated for these batches, Government Findings ¶ 43; Ex. 300; Mulligan Decl. ¶¶ 1-5; Daurio Decl. ¶ 14 (updated percentages reflecting , batch failures disclosed following trial); 1/65 (Mulligan)); Exh. 300; Mulligan Decl. ¶¶ 1-5, 19, as well as Barr's failure to include at least ten batches in their retrospective studies, 7/880, 955 81003 (Gerraughty)); Exh. 30; Mulligan Decl. ¶ 6, the Court's focus on these products is consistent with its pledge to give extra scrutiny to the batches which experienced repeated out-of-specification results.

26. Barr must conduct concurrent or prospective validation studies for each of the following products:

1. Acetohexamide tablets, 250 mg.;
2. Acetohexamide tablets, 500 mg.;
3. Cephalexin for Oral Suspension 125 mg./5 ml.;
4. Cephalexin for Oral suspension 250 mg./5 ml.;
5. Cephadrine capsules, 250 mg.;
6. Cephadrine for Oral Suspension, 250 mg./5 ml.;
7. Chlordiazepoxide HCL capsules, 5 mg.;
8. Chlordiazepoxide HCL capsules, 25 mg.;
9. Diphenhydramine HCL capsules, 25 mg.,
10. Dipyridamole tablets, 25 mg.;
11. Erythromycin Delayed Release capsules, 250 mg.;
12. Erythromycin Ethylsuccinate~tablets, 400 mg.;
13. Errthromycin Ethylsuccinate for Oral Suspension, 200 mg./5 ml
14. Erythromycin Ethylsuccinate and Sulfisoxazole Acetyl for Oral Suspension, 200 mg./600 mg./5 ml. ("ESP");
15. Erythromycin Stearate tablets, 250 mg.;
16. Erythromycin Stearate tablets, 500 mg.;
17. Hydroxyzine Pamoate capsules, 25 mg.;
18. Hydroxyzine Pamoate capsules, 100 mg.;
19. Meclofenamate Sodium capsules, 50 mg.;
20. Meclofenamate Sodium capsules, 100 mg.;
21. Meperidine HCL tablets, 50 mg.;
22. Meperidine HCL tablets, 100 mg.;
23. Propylthiouracil tablets, 50 mg.; and
24. Spironolactone with HCTZ tablets, 25 mg./25 mg.

Barr must cease all distribution of these products until such studies are completed.

27 Barr also must conduct concurrent or prospective validation studies for each of the following products:

1. Aspirin with Codeine tablets. 325 mg./15 mg.;
2. Aspirin with Codeine tablets, 325/60 mg.
3. Cephalexin tablets, 500 mg.,
4. Cephadrine capsules, 500 mg.;
5. Ibuprofen tablets, 400 mg.;
6. Isoniazid tablets, 100 mg.,
7. Isoniazid tablets, 300 mg.;
8. Methotrexate tablets, 2.5 mg.;
9. Cephadrine for Oral Suspension, 125 mg./5 ml.;
10. Diazepam tablets, 2 mg.;
11. Indomethacin capsules, 25 mg.;
12. Indomethacin capsules, 50 mg.;
13. Propranolol with HCTZ tablets, 40 mg./25 mg.;
14. Propranolol with HCTZ tablets, 80 mg./25mg.; and
15. Sulfipyrazone tablets, 100 mg.

Barr nevertheless retains the right to manufacture and distribute these products, provided that existing batches are retested in accordance with the procedures outlined herein and newly manufactured batches are both manufactured and tested in accordance with the same.

(3) Product Recall

28. The Court may recall any drug product found to be manufactured in violation of the Act that has been released to the public for distribution. See Lit Drug, 333 F. Supp. at 992; see also United States v. K-N Enterprises, Inc., 461 F.Supp. 988, 991 (N.D. Ill 1978) (recalling firm's misbranded drugs). Although not authorized expressly in the Act, this remedy is consistent with the broad equitable relief powers district courts enjoy. See Spectro Foods, 544 F.2d at 1182; but see United States v. Superpharm Corp., 530 F. Supp. 408, 410 (E.D.N.Y. 1981) (refusing to order recall given FDA seizure remedy). Any other rule would allow Barr to use its channels of distribution to shield adulterated products from the Court's remedial arm, thereby frustrating the purpose of the Act. See United States v. An Article Cf Drug . . . Bacto-Unidisk, 394 U.S. 784, 798, 89 S. Ct. 1410, 1418 (1969) (courts "must give effect to congressional intent in view of the well-accepted principle that remedial legislation such as the (Act) is to be given a liberal construction consistent with the Act's overriding purpose to protect the public health").
29. Citing a variety of failures, the government asks the Court to recall fifteen batches of ten different drug products.³⁶ (Exh. 2, ¶¶ 4, 38 at 379, 390; Exhs. 11, 21-23, 29-32, 37-50, 304; 2/304-05, 310 (Mulligan); 989-996, 1045-47, 1061-62 (Gerraughty)). In a batch-by-batch defense, Barr attempts to refute these charges.
30. In many instances, Barr relied on passing resampling results to release batches that experienced both initial and retest out-of-specification results. In this circumstance, Barr maintains, resampling determines "whether the sample preparation or the random variance of the microbiological assay rather than the product caused the ambiguous results." See e.g. Barr Findings, Appendix II, tab 5.
31. Having previously concluded that a successful resample alone cannot invalidate out-of-specification results, see ¶ 45 supra, the Court will order the recall of batches released on the basis of such data.³⁷ These batches include: Erythromycin Delayed Release capsules, 250 mg., batch 1F584EV; Erythromycin Stearate tablets, 240 mg., batch OIO13FQ; Erythromycin Ethylsuccinate and Sulf-isoxazole Acetyl for Oral Suspension, 200 mg./600 mg., batches 1HH445DQ and 1I445EF; Meperidine HCL tablets, 50 mg., batches 0B38IA0, 0B381AT, 0B38IAW, and 0B38IAV.
32. Content uniformity problems form the basis of other recall requests. The USP contains detailed requirements for finished product content uniformity testing. The first stage involves testing ten tablets. If each tablet falls within 85 to 115 and the standard deviation is 6.0 or less, the batch passes. If one of the ten tablets falls between either 75 and 85 or 115 to 125 or the standard deviation is above 6.0, firms then proceed to the second stage and test 20 more tablets. (1177:18-25; 1263:12 to 1274:4 (Mulligan)), (1209:22, 1211:14 (Gerraughty)).³⁸ Unless a firm with certainty establishes grounds to reject the tablet falling outside the 75 to 125 range, the batch should not be released. (2258:8 (Pierpaoli)).
33. First, the government cites Aspirin with Codeine tablets, 325/60mg, batch 1C28ODY, for content uniformity problems. One out of the first ten tablets tested was out-of-specification for both aspirin and codeine ingredients, yielding values of 122.2 and 64.8, respectively. (1263:21 (Mulligan)). Barr conducted fifty additional tests, all of which were within specification and released the batch. See Barr Findings, Appendix II, tab 5.

³⁶ The government also asks the Court to recall Dicyclomine HCL tablets, 20 mg., batch OE126AA and Dicyclomine HCl tablets, 20mg, batch 1H126CX. Since Barr temporarily suspended these products prior to this litigation, they are not the subject of the instant action.

³⁷ Another problem in this circumstance is Barr's failure to investigate the out-of-specification initial and retest results. Even experts that approved resampling after unsatisfactory retests required a full investigation. See e.g., 1536:14-21 (Rhodes).

³⁸ Although not required by CGMP, the conservative, responsible procedure is to use the 90 to 110 scale, normally used for assays, for content uniformity testing also. (1217:4 (Gerraughty)).

34. Thus, at the first stage of testing, batch IC28ODY yielded a result of 64.8 for codeine, a figure below the outside lower limit of 75, and a standard deviation of 11.1. (1264:10, 15 (Mulligan)). Given the prior difficulties this product experienced with content uniformity tests, (1273:10 (Mulligan)), and the absence of any reason to believe that the initial ten-tablet sample which generated one failing result was not representative of the batch, (2264:10-12 (Pierpaoli)), Barr should not have proceeded to the subsequent testing stage. This batch must be recalled.
35. Erythromycin Ethylsuccinate tablets, 400 mg., batch 0B259FA, had content uniformity and assay difficulties. With respect to content uniformity Barr cited an imprecise analytical method and retested ninety tablets. The assay values were low with one out-of-specification result. Subsequently, the FDA revised the test method, and Barr retested using the updated procedure. Receiving passing content uniformity and assay results, Barr released the batch.
36. The assay retesting is both justified and well-documented. The failure investigation report includes a memorandum explaining the retest under the updated FDA procedure. The Court has little confidence in the content uniformity testing, however, since Barr relied upon average content uniformity results to release the batch. Such averages hide the very variability the content uniformity test is designed to detect. This batch must be recalled.
37. Hydrocodone Bitartrate with APAP tablets, 5 mg./500 mg., batch 1D325CL, had finished product assay troubles. Barr retested in duplicate, obtained passing values and released the batch. In light of hearing testimony, the Court is not satisfied with a sixty-six percent chance that the passing results represent the true assay value of the batch. This batch must be recalled.
38. The government cites Erythromycin Stearate, 500 mg., batch 2A219HJ, for content uniformity problems. Here, during the initial content uniformity testing of the core, three tablets of the thirty tested were below specification and the relative standard deviation was 9.9. Barr retested two sets of thirty tablets, received passing results and released the batch. Based on a statistical evaluation, Barr concluded that the initial tests were the result of analyst error.
39. Although this batch had passing assay values at both the core and finished product stages, and Barr attempted to establish grounds for rejecting initial out-of-specification results, the Court will recall this batch. In the blend uniformity testing context, passing and nonpassing results are not inherently inconsistent. See note 5 supra. Thus, where, as here, the sample population does not enjoy an assumption of homogeneity, the use of an outlier test is inappropriate, see Hammerstrom Declaration, ¶ 12, especially when applied to average values. Id. ¶ 15.
40. Hydroxyzine Pamoate Capsules, 100 mg., batch 1K324E1, exhibited both finished product assay and content uniformity out-of-specification results. Barr attributed the assay problems to "deviant capsule fill weights" obtained by analytical error, although there was no evidence of analyst error in the fill weight testing. (2276:6-9 (Pierpaoli)). Based on statistical evaluation, the fill weights were determined to be discrepant. The three retests which followed were within specification see Ex. 38 at B-1572, as was the average of all eight fill weight tests.
41. With regard to the out-of-specification content uniformity results, Barr discovered that the analysts conducting the first two series of tests failed to follow the testing protocol and discounted these results accordingly. (2280:14 (Pierpaoli)). Two further sets of content uniformity tests fell within the required specifications. See Ex. 38 at B-1574.
42. Despite the various difficulties with this batch, the Court will not order its recall. The Court agrees that these separate and unrelated problems are "happenstance". (2280:21 (Pierpaoli)). Barr responded to these mishaps with thorough investigations, adequately supported its conclusions, and based its final determination on the entire aggregate of data and other information. (2281:10 (Pierpaoli)).
43. The Court will not order a recall of Propylthiouracil tablets, 50 mg., batch 2A089HG. Here, Barr cited analyst error as the cause of the initial out-of-specification results. Consistent with a proper failure investigation, Barr documented the suspected errors, conducted retests with the correct standards and obtained passing retest results.
44. The Court will not order a recall of Spironolactone, batch 9E265EA, in light of the government's concession during the hearing that the cited dissolution failure was not required by the USP until 1990. Barr made this batch in 1989. (336:4 (Mulligan)).

CONCLUSION

For the reasons stated above, the Court will grant in part and deny in part the government's motion for a preliminary injunction.³⁹ Finally, the Court will retain jurisdiction of this action in order to ensure that defendants comply with the Act and the CGMP regulations as construed by the Court.

An appropriate order is attached.

Dated: February 4, 1993

ALFRED M. WOLIN, U.S.D.J.

³⁹ The injunctive relief granted today binds defendant Barr and its officers, agents, employees and attorneys.