

When Biomarkers Are Not Enough: FDA Evaluation of Effectiveness of Neuropsychiatric Devices for Disorders of Consciousness

Megan S. Wright, Keturah James, Adam Pan,
and Joseph J. Fins*

21 STAN. TECH. L. REV. 276 (2018)

ABSTRACT

Studies demonstrating the efficacy of interventions for neuropsychiatric disorders often rely on the assessment of a trained observer or patient reports, evidence that can be more subjective than biomarkers. The problem with relying upon subjective evidence of efficacy is that objective evidence is advantaged in the medical product approval process, notwithstanding the increase in submission of patient-reported outcomes in applications to the U.S. Food and Drug Administration (FDA). The FDA's preference for biomarkers may thus adversely impact the approval process for interventions targeting neuropsychiatric conditions, for which there are no biomarkers or only emerging biomarkers (with limited sensitivity or specificity) associated with the disorders.

In this Article, we first review how the FDA evaluates different types of evidence of an intervention's efficacy, focusing on Class III medical devices. We use the case of disorders of consciousness to analyze relevant regulations and guidance for the development of novel neurotechnologies when subjective data is used in an approval process. We then offer suggestions for reform, arguing

* Megan S. Wright, J.D., Ph.D. is a Postdoctoral Associate of Medical Ethics at Weill Cornell Medical College and a Research Fellow at the Solomon Center for Health Law and Policy at Yale Law School. Keturah James and Adam Pan, Ph.D. are J.D. candidates at Yale Law School. Joseph Fins, M.D., M.A.C.P. is the E. William Davis, Jr., M.D. Professor of Medical Ethics, Professor of Medicine, and Chief of the Division of Medical Ethics at Weill Cornell Medical College, co-director of the Consortium for the Advanced Study of Brain Injury (CASBI) at Weill Cornell Medical College and Rockefeller University, and the Solomon Center Distinguished Scholar in Medicine, Bioethics, and the Law at Yale Law School. The Authors wish to acknowledge the support of the Jerold B. Katz Foundation, the Solomon Center for Health Law and Policy at Yale Law School, and Weill Cornell Medical College, and thank Professor Abbe Gluck for helping to make this collaboration possible and for her support of CASBI@YLS.

for a more expansive approach to what constitutes evidence of efficacy en route to more objective and vetted biomarkers. Specifically, we argue that patient-provided information is valuable and should be included in applications to the FDA, yielding a more integrated and holistic approach, or what some scholars refer to as a “pluralistic way of knowing.” We also argue for greater transparency from the FDA about how patient-provided information is evaluated, valued, and weight as part of the application for approval of new drugs and devices. While our Article focuses on disorders of consciousness and neurotechnology, our arguments have broader implications for the evaluation of emerging drugs and devices designed to ameliorate other neuropsychiatric conditions where the subjective experience of patients has particular relevance.

TABLE OF CONTENTS

INTRODUCTION	278
I. HIERARCHY OF EVIDENCE USED TO ASSESS EFFICACY OF INTERVENTIONS: THE CASE OF DISORDERS OF CONSCIOUSNESS...	282
II. FDA MEDICAL DEVICE APPROVAL PROCESS: EVIDENCE OF EFFECTIVENESS	287
A. <i>Medical Device Approval Process</i>	288
B. <i>Demonstrating Effectiveness? FDA Preference for Biomarkers in Drug Studies</i>	291
C. <i>The Role of PRO and PPI in Medical Device Applications</i>	294
III. POLICY RECOMMENDATIONS	298
A. <i>Require Submission of PPI and PRO</i>	300
1. <i>Regulatory and Scientific Considerations</i>	301
2. <i>Ethical Considerations</i>	303
3. <i>Case of Disorders of Consciousness</i>	304
4. <i>Additional Considerations</i>	305
B. <i>Permit Proxy Reports in PPI and PRO</i>	305
C. <i>Transparency in How PPI and PRO Assessed During Regulatory Review</i>	308
IV. CONCLUSION	309

INTRODUCTION

New drugs and medical devices must demonstrate effectiveness prior to approval by the U.S. Food and Drug Administration.¹ There has been increasing focus on the use of biological markers in evaluations of an intervention’s efficacy, and recent legislation even created a Biomarker Qualification Program to help expedite drug development.² Researchers affiliated with the National Institutes of Health explain that “[t]he term ‘biomarker,’ a portmanteau of ‘biological marker,’ refers to a broad subcategory of medical signs—that is, objective indications of medical state observed from outside the patient—which can be measured accurately and reproducibly.”³

1. See generally 21 C.F.R. § 312 (2017) (describing requirements for new drugs) and 21 C.F.R. §§ 800-898 (2017) (describing requirements for medical devices).

2. See 21st Century Cures Act § 3011, 21 U.S.C. § 357 (2016); Susan McCune, *Biomarker Qualification Program Educational Module Series-Module 2: How Biomarkers Can Improve the Drug Development Process*, FDA (May 12, 2017), <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/ucm539845.htm> [<https://perma.cc/RQ98-HQN3>] (“Biomarkers are playing an increasingly important role in drug development. They can help improve clinical trial efficacy and reduce uncertainty in regulatory decision making.”).

3. Kyle Strimbu & Jorge A. Tavel, *What are Biomarkers?*, 5 CURR. OPIN. HIV

Not all diseases and disorders have clear biomarkers, however, and for other conditions, there are better validated and more reliable *non*-biomarker instruments that can be used to measure effects of medical interventions. A common type of non-biomarker assessment may be one in which a clinician or investigator administers validated, reliable scales, which are usually produced through a questionnaire designed to measure the concept of interest.⁴ Alternatively, an instrument known as a patient reported outcome (PRO) can assess the patient's perceptions of the effect of the intervention post-drug intake or post-device use.⁵ Biomarkers are more objective than patient reports and clinical assessments because biomarkers can be quantitatively measured by someone other than the patient. Moreover, biomarkers are mediated primarily through technology rather than through a clinician using their judgment, thus mitigating an opportunity for error.

More subjective instruments are by necessity used in the assessment of emerging therapies for neurological and psychiatric conditions, given that "the living brain is protected from the prying eyes and instruments of investigators by its bony skull, and . . . human experimental neurobiology is mostly limited to indirect, non-invasive methods of investigation."⁶ Assessing the efficacy of an intervention to treat somatic disorders thus may be easier than assessing the efficacy of an intervention for neuropsychiatric disorders given the greater likelihood of being able to use biomarkers in the former category of disorders. And within the broad category of neuropsychiatric disorders, it may be easier to assess efficacy of interventions targeted toward physical symptoms (e.g., tremors) as compared to nonphysical symptoms (e.g., obsessive thoughts). For example, one can observe tremors as a physical sign at the bedside but to assess obsessive thoughts one would need to administer the Yale-Brown Obsessive Compulsive Scale (YBOC Scale), a neuropsychological instrument administered and rated by clinicians. The YBOC asks patients about, for example, the amount of time spent each day thinking obsessive thoughts or their perceived degree of control over a compulsive behavior.⁷

AIDS 463, 463 (2010).

4. One example of such a scale is the Coma Recovery Scale-Revised (CRS-R). See Joseph T. Giacino et al., *The JFK Coma Recovery Scale-Revised: Measurement Characteristics and Diagnostic Utility*, 85 ARCHIVES PHYS. MED. & REHABILITATION 2020, 2020 (2004) (describing reliability and validity of scale).

5. One example of a commonly used patient reported outcome measure is the Visual Analog Scale (VAS). See, e.g., Richard Shikhar et al., *Validity and Reliability of Patient Reported Outcomes Used in Psoriasis: Results from Two Randomized Clinical Trials*, 1 HEALTH & QUALITY OF LIFE OUTCOMES 53 (2003) (describing validity and reliability of VAS in studies of psoriasis).

6. Steven E. Hyman, *A Glimmer of Light for Neuropsychiatric Disorders*, 455 NATURE 890, 890 (2008).

7. Wayne K. Goodman et al., *The Yale-Brown Obsessive Compulsive Scale*, 46 ARCHIVES GENERAL PSYCHIATRY 1006, 1007 (1989).

For interventions targeting neuropsychiatric disorders, or disorders that are understood by examining the “interaction of neurobiological and psychological-social factors,”⁸ there are thus two primary sources of evidence of efficacy, both of which are somewhat subjective: the patient’s self-reports of the intervention’s effect, or the assessment of an observer. The problem with these more subjective kinds of evidence of efficacy is that objective evidence (i.e., chemical or electrophysiological biomarkers) is advantaged in the drug and device approval process,⁹ despite the increase in submission of PROs in applications to the FDA.¹⁰ This preferential treatment for certain kinds of evidence may thus decrease the likelihood of drug or device approval for neuropsychiatric conditions, which already lack efficacious treatments.¹¹

In this Article, we first review how the FDA evaluates different types of evidence¹² of an intervention’s efficacy, focusing on Class III medical devices. We focus on medical devices because (1) regulatory guidance about evidence of efficacy is not as well developed as guidance for drug development, (2) potentially groundbreaking neurotechnologies are in development for neuropsychiatric conditions that lack biomarkers as

8. Perminder Sachdev, *International Neuropsychiatric Association*, 1 NEUROPSYCHIATRIC DISEASE & TREATMENT 191, 191 (2005) (“Neuropsychiatry is a field of scientific medicine that concerns itself with the complex relationship between human behavior and brain function . . .”); see also Paul Krack et al., *Five-Year Follow-up of Bilateral Stimulation of the Subthalamic Nucleus in Advanced Parkinson’s Disease*, 349 NEW ENGLAND J. MED. 1925, 1932 (2003) (describing how DBS for Parkinson’s, typically thought of a neurological disorder, results in depression, typically thought of as a psychiatric disorder).

9. See discussion *infra* Part II.

10. CENTER FOR DEVICES AND RADIOLOGICAL HEALTH (CDRH), VALUE AND USE OF PATIENT-REPORTED OUTCOMES (PROS) IN ASSESSING EFFECTS OF MEDICAL DEVICES: CDRH STRATEGIC PRIORITIES 2016-2017 8 [hereinafter CDRH], <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHVisionandMission/UCM588576.pdf> [<https://perma.cc/FQ7Q-AUXJ>].

11. I-han Chou & Tanguy Chouard, *Neuropsychiatric Disease*, 455 NATURE 889, 889 (2008) (describing treatment options as “far from optimal in terms of efficacy and specificity” and introducing NATURE issue no. 7215 in which all articles make this point); Vaishnav Krishnan & Eric J. Nestler, *The Molecular Neurobiology of Depression*, 455 NATURE 894, 901 (2008) (“enormous gaps in the knowledge of depression and its treatment persist”).

12. We are focusing on types of data rather than study design. Many other commentators have expressed concern that since the passage of the 21st Century Cures Act, the FDA is relaxing its requirements of randomized controlled trials, especially in the medical device context for which such trials are already difficult to conduct. See, e.g., Jerry Avorn & Aaron S. Kesselheim, *The 21st Century Cures Act—Will It Take Us Back In Time?*, 372 NEW ENG. J. MED. 2473, 2474 (2015); Aaron S. Kesselheim & Jerry Avorn, *New “21st Century Cures” Legislation: Speed and Ease vs Science*, 317 J. AM. MED. ASS’N 581, 581 (2017); Vinay K. Rathi et al., *Characteristics of Clinical Studies Conducted Over the Total Product Life Cycle of High-Risk Therapeutic Medical Devices Receiving FDA Premarket Approval in 2010 and 2011*, 314 J. AM. MED. ASS’N 604, 610 (2015).

part of the National Institutes of Health (NIH) BRAIN Initiative,¹³ (3) the FDA “recognizes the value of medical device innovation to address unmet clinical needs and improve patient care . . . [for] neurological diseases or conditions,”¹⁴ and (4) leaders in the field of neuropsychiatric disorders are currently calling for the creation of scientifically valid diagnoses and the establishment of objective measures of the disorders.¹⁵ We use the case of disorders of consciousness (DOC) after severe brain injury, which are neuropsychiatric disorders¹⁶ that include coma, vegetative state (VS), minimally conscious state (MCS), and the acute confusional state¹⁷ and that lack sensitive and specific biomarkers, to analyze potential pitfalls of the relevant regulations and guidance for the development of novel neurotechnologies. This class of disorders also does not yet have established therapies for treatment, which provides a good opportunity to highlight what types of evidence could be used in developing effective therapeutic interventions. We

13. See CORNELIA BARGMANN & WILLIAM NEWSOME, NAT’L INSTS. OF HEALTH, BRAIN 2025: A SCIENTIFIC VISION 11, 13 (2014); Kathy L. Hudson & Francis S. Collins, *The 21st Century Cures Act—A View from the NIH*, 376 NEW ENG. J. MED. 111, 113 (2017); Nicholas D. Schiff, *Central Thalamic Stimulation for Traumatic Brain Injury*, NIH REPORTER 5UH3NS095554-02, https://projectreporter.nih.gov/project_info_description.cfm?aid=9148199 [<https://perma.cc/CT8P-MPDY>] (NIH BRAIN Initiative funded study).

14. FDA, CLINICAL CONSIDERATIONS FOR INVESTIGATIONAL DEVICE EXEMPTIONS (IDES) FOR NEUROLOGICAL DEVICES TARGETING DISEASE PROGRESSION AND CLINICAL OUTCOMES 4 (2016) [hereinafter NEUROLOGICAL DEVICES], <https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm489111.pdf> [<https://perma.cc/BE5W-NJC2>].

15. Thomas Insel, *Transforming Diagnosis*, NATIONAL INSTITUTE OF MENTAL HEALTH (Apr. 29, 2013), <https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2013/transforming-diagnosis.shtml> [<https://perma.cc/H96R-6AFF>].

16. Sachdev, *supra* note 8, at 191. For a description of the link between brain injury and psychological conditions, see Daan P.J. Verberne et al., *Psychological Interventions for Treating Neuropsychiatric Consequences of Acquired Brain Injury: A Systematic Review*, 13 NEUROPSYCHOLOGICAL REHABILITATION 1, 2 (2018).

17. The vegetative state (VS) is characterized as “wakeful unconsciousness” in which the patient’s eyes are open but there is no awareness of self, others, or the environment. Functions are autonomic and localized to the brain stem without higher integrated cortical function. Under current nosological categorization as this Article goes to press, the VS can be persistent (lasting over a month) or permanent (three months after a non-traumatic brain injury and twelve months after a traumatic brain injury). Joseph T. Giacino et al., *Disorders of Consciousness After Acquired Brain Injury: The State of the Science*, 10 NATURE REV. NEUROLOGY 99, 100 (2014). The MCS “is a condition of severely altered consciousness characterized by minimal but definite behavioral evidence of self or environmental awareness.” *Id.* at 100. A person with a severe brain injury may be covertly conscious, unable to evidence their cognitive functions through their behaviors but evidencing consciousness through imagining volitional actions on investigative neuroimaging studies. *Id.* at 103. Furthermore, when patients emerge from the MCS and have more reliable functional communication, they may be confused and also have disabilities, and this is known as the acute confusional state. *Id.* at 101.

then offer suggestions for reform, arguing for a more expansive approach to what constitutes evidence of efficacy. Specifically, we argue that when objective evidence of efficacy is possible to obtain, then it should be required for approval, but, *en route* to more objective and vetted biomarkers, other types of evidence, particularly patient-provided information, should be considered valuable as well. These patient reports should also be required as part of the application process, yielding a more integrated and holistic approach, or what some scholars refer to as a pluralistic way of knowing.¹⁸ This approach is consistent with the ethos of the 21st Century Cures Act that provides for greater patient input in the regulatory process.¹⁹ We also argue for greater transparency from the FDA about how patient-provided information is evaluated and titrated as part of the application for approval of new drugs and devices. While this Article will focus on DOC and neurotechnology, our arguments will have broader implications for the evaluation of emerging drugs and devices designed to ameliorate other neuropsychiatric conditions.

I. HIERARCHY OF EVIDENCE USED TO ASSESS EFFICACY OF INTERVENTIONS: THE CASE OF DISORDERS OF CONSCIOUSNESS

It is possible to classify evidence used to assess the effect of a treatment as existing on a continuum from the most objective to least objective. The most objective evidence consists of biomarkers. While there are multiple definitions and functions of biomarkers,²⁰ for the purposes of this Article, we will be using the joint FDA-NIH definition. A biomarker is

A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers. A biomarker is not an assessment of how an individual feels, functions, or survives.²¹

So, for example, blood pressure is a biomarker that can be used to determine whether a blood pressure medication works for hypertension.²² And in the case of ovarian cancer, serum cancer antigen CA-125

18. MIRIAM SOLOMON, MAKING MEDICAL KNOWLEDGE (2015).

19. See, e.g., 21st Century Cures Act § 3001-03, 21 U.S.C. § 360bbb-8c (2016).

20. See, e.g., Strimbu & Tavel, *supra* note 3, at 463; FDA-NIH BIOMARKER WORKING GROUP, BEST (BIOMARKERS, ENDPOINTS, AND OTHER TOOLS) RESOURCE 3-35 (2018), <https://www.ncbi.nlm.nih.gov/pubmed/27010052> [<https://perma.cc/DQE9-2Y9D>].

21. *Id.* at 48.

22. *Id.* at 13.

is a known tumor marker,²³ and an intervention assessed for its efficacy in treating ovarian cancer is evaluated based on how it affects CA-125 levels in the blood. The levels of CA-125 are thus biomarkers of disease burden and tumor load. In the neuropsychiatric context, neuroimaging biomarkers, such as MRI images of structural changes in the brain, are in the process of being validated as biomarkers for diseases such as Alzheimer's,²⁴ but for the most part, biomarkers for neuropsychiatric disorders are lacking.²⁵

Biomarkers are considered objective because they are quantifiable²⁶ and do not depend on patient reports or clinician judgment, but instead rely on technology. There can still be elements of subjectivity in using biomarkers, however. For example, statistical and clinical analysis of output from an electroencephalogram (EEG), which is a potential biomarker for neuropsychiatric disorders,²⁷ can be interpretive.²⁸

Scales created from a clinical assessment or questionnaire administered by a clinician or investigator, while often useful and informative, are less objective than biomarkers because they are mediated by a person rather than technology or may not be observable by an outsider. However, if they are validated and reliable,²⁹ they become more objective. For example, the Karnofsky Performance Status Scale was developed to assess the functional status of patients with cancer, and it is used to assess a patient's activities of daily living following the administration of different therapies.³⁰ The scale is both reliable and valid.³¹ But,

23. Nathalie Scholler & Nicole Urban, *CA125 in Ovarian Cancer*, 1 *BIOMARK MED.* 513 (2007). This tumor marker has poor sensitivity and specificity, however. S. Pignata, et al., *Follow-up with CA125 After Primary Therapy of Advanced Ovarian Cancer: In Favor of Continuing to Prescribe CA125 During Follow-up*, 22 *ANNALS ONCOLOGY* viii40, viii40 (2011).

24. Bradford C. Dickerson & Reisa A. Sperling, *Neuroimaging Biomarkers for Clinical Trials of Disease - Modifying Therapies in Alzheimer's Disease*, 2 *NEURORX* 348 (2005).

25. KEWAL K. JAIN, *THE HANDBOOK OF BIOMARKERS* 463, 719 (2d ed. 2017); Chou & Chouard, *supra* note 11; Hyman, *supra* note 6; Insel, *supra* note 15.

26. Strimbu & Tavel, *supra* note 3, at 2.

27. Andras Horvath et al., *EEG and ERP Biomarkers of Alzheimer's Disease: A Critical Review*, 23 *FRONTIER BIOSCIENCES* 183, 184-186, 198, 200-202 (2018).

28. Andrew M. Goldfine et al., *Reanalysis of "Bedside Detection of Awareness in the Vegetative State: A Cohort Study,"* 381 *LANCET* 289 (2013).

29. A scale is valid if it measures what it purports to measure, and it is reliable if it results in the same measure regardless of who uses it or when it is used. Martyn Hammersley, *Some Notes on the Terms 'Validity' and 'Reliability,'* 13 *BRITISH EDUCATIONAL RESEARCH J.* 73 (1987). For discussions of validity and reliability of measurement instruments in a medical context, see J. Martin Bland & Douglas G. Altman, *Validating Scales and Indexes*, 324 *BMJ* 606 (2002). See also discussion *infra* Part II.C.

30. See Vincent Mor et al., *The Karnofsky Performance Status Scale: An Examination of Its Reliability and Validity in a Research Setting*, 53 *CANCER* 2002, 2003-2006 (1984) (discussing scale and its validity and reliability).

31. *Id.*

as one research team described, it is “a subjective value assigned by clinicians within a matter of seconds.”³²

Patient reports of their experiences with illness and a treatment or investigational intervention, collected through a standard structured questionnaire,³³ also can constitute evidence of efficacy in this intermediary category of evidence. PROs are defined as “any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else.”³⁴ Because chronic pain is a condition without well-established biomarkers, it is assessed through a variety of instruments that rely on patient reports of their pain,³⁵ including the Visual Analogue Scale (VAS).³⁶

The least objective types of evidence are patient reports or clinician observations that are only qualitative in nature. Examples include unsolicited patient comments about their experiences in a clinical trial, or an investigator’s observations that are not collected in a systematic manner using a vetted instrument. While such evidence is most subjective, it is still important to consider. Indeed, the unjustified marginalization of patient reports is captured by an often-cited comment by Dr. Eugene A. Stead, Jr., the late long-time chairman of medicine at Duke School of Medicine. Observing a patient on rounds, he wryly commented, “I have nothing to add. Patient says she is feeling better and her chart seems to agree with this.”³⁷ In his comment, Dr. Stead warns his students neither to discount the patient’s voice, which is a qualitative self-report not collected from a validated, reliable questionnaire, nor

32. Dominik Peus et al., *Appraisal of the Karnofsky Performance Status and Proposal of a Simple Algorithmic System for Its Evaluation*, 13 BMC MED. INFORMATICS & DECISION MAKING 1, 2 (2013).

33. Prasanna R. Deshpande et al., *Patient-Reported Outcomes: A New Era in Clinical Research*, 2 PERSP. CLINICAL RES. 137, 138 (2011).

34. FDA, PATIENT PREFERENCE INFORMATION — VOLUNTARY SUBMISSION, REVIEW IN PREMARKET APPROVAL APPLICATIONS, HUMANITARIAN DEVICE EXEMPTION APPLICATIONS, AND *DE NOVO* REQUESTS, AND INCLUSION IN DECISION SUMMARIES AND DEVICE LABELING 8 (2016) [hereinafter PPI] (citing FDA, GUIDANCE FOR INDUSTRY: PATIENT-REPORTED OUTCOME MEASURES: USE IN MEDICAL PRODUCT DEVELOPMENT TO SUPPORT LABELING CLAIMS (2009), <http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf> [https://perma.cc/X95V-T9FK]), <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446680.pdf> [https://perma.cc/WH5A-TDX3]).

35. Mathias Haefeli & Achim Elfering, *Pain Assessment*, 15 EUR. SPINE J. S17, S18-S19 (2006).

36. Deshpande et al., *supra* note 33, at 141. Other examples of PROs are the Health Assessment Questionnaire (HAQ) and Disability Index (DI) score for physical function. PPI, *supra* note 34; *id.*

37. JOSEPH J. FINS, *A PALLIATIVE ETHIC OF CARE: CLINICAL WISDOM AT LIFE’S END* 164 (2006).

overinvest the medical record, which may contain more objective measures such as lab results, as the arbiter of the patient's experience.

In neuropsychiatric disorders, feedback from the patient, both as quantified PROs and in more qualitative patient reports, becomes more critical as it can be difficult to measure drug or device efficacy using an objective, biomarker metric.³⁸ This is because the symptoms in these disorders are subjective by their nature. Neither signs of depressed mood nor evidence of an intervention's effect on these symptoms may always be evident on brain imaging or blood tests, for example. Although promising research with deep brain stimulation (DBS) in treatment-resistant depression has demonstrated changes in metabolic activity on positron emission tomography (a type of neuroimaging) associated with a clinical anti-depressant effect,³⁹ such correlations using neuroimaging as a biomarker of an intervention's efficacy are on the investigative vanguard and as yet, not the evidentiary norm. Therefore, assessing the effectiveness of the tested intervention using less objective modes of assessment, such as the administration of valid and reliable psychological scales, remains especially important given the nature of the disorder and symptoms being targeted. So, as noted, the YBOC Scale is used to assess symptoms of obsessive compulsive disorder and has been found to be reliable.⁴⁰ It is used in clinical trials to evaluate the effect of drugs being evaluated to treat obsessive compulsive disorder⁴¹ given that there are not yet established biomarkers for this disorder.⁴²

The importance of more subjective evidence is clear in efficacy studies of medical devices in which biomarkers neither exist nor are an appropriate indicator of disease activity. With devices "intended to slow, stop, or reverse the effects of neurological disease,"⁴³ for instance, the challenges to obtaining the most objective evidence are two-fold: (1) there may not be any known biomarkers that correlate with disease progression, and (2) even if there were, collection of biometric data tends to be invasive and alternative approaches may be more well-accepted in the clinical community.⁴⁴ Especially in cases where treatment

38. JAIN, *supra* note 25, at 719.

39. Helen S. Mayberg et al., *Deep Brain Stimulation for Treatment-Resistant Depression*, 45 NEURON 651, 655, 657 (2005).

40. Wayne K. Goodman et al., *The Yale-Brown Obsessive Compulsive Scale*, 46 ARCH. GEN. PSYCHIATRY 1006, 1009-1011 (1989) (discussing reliability of YBOC).

41. See, e.g., David F. Tolin et al., *Defining Response in Clinical Trials for Obsessive-Compulsive Disorder: A Signal Detection Analysis of the Yale-Brown Obsessive Compulsive Scale*, 66 J. CLINICAL PSYCHIATRY 1549, 1549 (2005).

42. Ilana Frydman et al., *Can Neuroimaging Provide Reliable Biomarkers for Obsessive-Compulsive Disorder?: A Narrative Review*, 18 CURRENT PSYCHIATRY REP. no. 90, Aug. 22, 2016, at 1, 2.

43. NEUROLOGICAL DEVICES, *supra* note 14, at 4.

44. *Id.* See also Hyman, *supra* note 6, at 890 (describing how the brain is difficult to penetrate to study).

with a device is meant to restore function, as in implantation of DBS devices, the best way of assessing improvement is currently with clinical observation and administration of psychometric scales.⁴⁵ And even as assessment evolves towards the utilization of biomarkers, these more objective assessments will not obviate the need for more subjective measures against which the biomarker will need to be assessed and calibrated to ensure the biomarkers are valid.⁴⁶ As the FDA notes in guidance about neurological devices, “[w]hen biomarkers are chosen as a metric, there should be well established evidence and agreement in the clinical community that the chosen biomarker reflects a characteristic that is important to the underlying disease process and that it is associated with a clinically meaningful outcome measure.”⁴⁷

There is also the need to rely on multiple types of evidence of efficacy for interventions designed to target DOC resulting from brain injury. Without a reliable physiological biomarker that directly corresponds to consciousness, less objective assessments are the primary means of determining the severity of the disorder and ascertaining the efficacy of a particular therapeutic intervention.⁴⁸ For patients in the MCS, for instance, initial evaluations of consciousness and subsequent improvements in cognitive functioning are often measured with a standardized and vetted behavioral assessment such as the JFK Coma Recovery Scale Revised (CRS-R),⁴⁹ which is administered by a trained clinician or investigator.⁵⁰ In an early proof-of-principle study assessing whether thalamic DBS could assist in recovery for patients in the MCS, neuroimaging could not be used safely.⁵¹ The CRS-R was used to show how one patient became interactive and had an increase in the functional domains assessed by the CRS-R.

45. Refer to, for example, use of the UPDRS for symptoms of Parkinson’s Disease. See Christopher G. Goetz et al., *Movement Disorder Society-Sponsored Revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS): Process, Format, and Clinical Testing Plan*, 22 MOVEMENT DISORDERS 41, 42 (2007).

46. JAIN, *supra* note 25, at 713 (explaining validating and quantifying biomarkers “is the process of assessing . . . measurement performance characteristics and . . . providing evidence to link a biomarker with biology and clinical endpoints”).

47. NEUROLOGICAL DEVICES, *supra* note 14, at 6-7.

48. Neuroimaging may also be used in combination with the CRS-R when diagnosing and evaluating interventions for DOC even if they are not yet validated and reliable biomarkers. Giacino et al., *supra* note 17, at 109-110.

49. Giacino et al., *supra* note 4, at 2020.

50. An alternative to the CRS-R also relies on clinical observation rather than biomarkers. The Wessex Head Injury Matrix is a tool, which relies on observation, that can be used by any trained member of the medical treatment team in a period of observation. Lynne Turner-Stokes et al., *Serial Measurement of Wessex Head Injury Matrix in the Diagnosis of Patients in Vegetative and Minimally Conscious States: A Cohort Analysis*, 5 BMJ OPEN (2015).

51. See generally Nicholas D. Schiff et al., *Behavioral Improvements with Thalamic Stimulation After Severe Traumatic Brain Injury*, 448 NATURE 600 (2007).

Moreover, even if possible to use neuroimaging to assess the effect of a drug or device for DOC and other neuropsychiatric disorders,⁵² there can be subjectivity or bias when reading/interpreting the neuroimaging results for neuroimaging biomarkers. Additionally, the understanding of these disorders is still in its infancy; therefore, formal and systematic clinical assessments as well as clinical and family member observation remain crucial when assessing impacts of interventions. Studies have shown that families are often aware of the level of consciousness a person with a DOC has,⁵³ and since they may spend more time with a patient than the healthcare team or investigators, families can be instrumental in communicating evidence of consciousness and describing the best ways and times to communicate with the patient. However, while the CRS-R is currently considered the most trustworthy behavioral diagnosis tool for patients with DOC⁵⁴—and is therefore the clear choice to assess the efficacy of medical devices designed to treat these disorders—its reliance on behavioral measures presents opportunities for interpretive error, notwithstanding good interobserver reliability.

For patients with DOC, it is vital they have access to safe and effective neurotechnology that can assist them in regaining function. In this next section, we will explore how the U.S. Food and Drug Administration evaluates evidence of efficacy and how this matters for medical devices targeted for neuropsychiatric conditions generally and DOC specifically.

II. FDA MEDICAL DEVICE APPROVAL PROCESS: EVIDENCE OF EFFECTIVENESS

Pursuant to the Federal Food, Drug, and Cosmetic Act, FDA approval is required before new drugs and medical devices can be legally marketed in the United States. Though safety and effectiveness are the

52. Currently, such imaging is used in an investigational manner and is not a conclusive biomarker. See, generally, e.g., Christine Brefel-Courbon et al., *Clinical and Imaging Evidence of Zolpidem Effect in Hypoxic Encephalopathy*, 62 ANNALS NEUROLOGY 102 (2007) (using clinical exams as the primary measure of intervention effects, including neuroimaging results).

53. There is a high positive correlation between family assessments of consciousness and results from the CRS-R. Ralf J. Jox et al., *Diagnosis and Decision Making for Patients with Disorders of Consciousness: A Survey Among Family Members*, 96 ARCHIVES PHYS. MED. & REHABILITATION 323 (2015).

54. See Caroline Schnakers et al., *Diagnostic Accuracy of the Vegetative and Minimally Conscious State: Clinical Consensus Versus Standardized Neurobehavioral Assessment*, 9 BMC NEUROLOGY 35 (2009); Caroline Schnakers et al., *A French Validation Study of the Coma Recovery Scale-Revised (CRS-R)*, 22 BRAIN INJURY 786 (2008); Ronald T. Seel et al., *Assessment Scales for Disorders of Consciousness: Evidence-Based Recommendations for Clinical Practice and Research*, 91 ARCHIVES PHYS. MED. & REHABILITATION 1795 (2010).

central determinants of FDA approval for both drugs and certain devices, the specific approval processes differ for each. The FDA prefers biomarker evidence of efficacy, but there is a role for more subjective evidence in the review process.

A. *Medical Device Approval Process*

The path to approval for drugs requires submitting an investigational new drug (IND) application, agency approval of the IND and clinical trial protocol, and approval from the local institutional review board (IRB) at the institutions where the proposed studies will take place.⁵⁵ FDA approval for new drugs requires “substantial evidence of clinical effectiveness,” which consists of “adequate and well-controlled investigations” that have clinically meaningful endpoints, where “[t]he methods of assessment . . . are well-defined and reliable.”⁵⁶

Medical devices, on the other hand, follow a different route to FDA approval and constitute a distinct niche in the regulatory process.⁵⁷ Each new device is assigned a classification that, among other things, reflects the degree of perceived risk to patients.⁵⁸ Class I includes low risk devices, such as bandages, while Class II includes moderate risk devices, such as a powered wheelchair, that are subject to more extensive regulations.⁵⁹ Class III devices are those that support or sustain human life, are of “substantial importance in preventing impairment of human health,” or pose an unreasonable risk to patient safety.⁶⁰ A cardiac pacemaker or deep brain stimulator is an example of a Class III medical device.⁶¹

To market a Class III device, a manufacturer must obtain a pre-market approval (PMA) that provides reasonable assurance of device safety and effectiveness. There is reasonable assurance of effectiveness “when it can be determined, based on valid scientific evidence, that . . .

55. 21 C.F.R. § 312.23(a) (2017).

56. 21 C.F.R. § 314.126(a), (b)(6) (2017).

57. Joseph J. Fins, *Devices, Drugs & Difference: Deep Brain Stimulation and the Advent of Personalized Medicine*, in SPRINGER HANDBOOK ON NEUROETHICS 607 (N. Levy & J. Clausen eds., 2015) (comparing drug and device approval processes, size of industries, and clinical trials).

58. 21 U.S.C. § 360c(a) (2016).

59. *Learn if a Medical Device Has Been Cleared by FDA for Marketing*, FDA (Dec. 29, 2017), <https://www.fda.gov/MedicalDevices/ResourcesforYou/Consumers/ucm142523.htm> [<https://perma.cc/7M42-YZJR>].

60. 21 U.S.C. § 360c(a)(1)(C)(ii) (2016).

61. *Learn if a Medical Device Has Been Cleared by FDA for Marketing*, FDA (Dec. 29, 2017), <https://www.fda.gov/MedicalDevices/ResourcesforYou/Consumers/ucm142523.htm> [<https://perma.cc/7M42-YZJR>].

the use of the device for its intended use . . . will provide clinically significant results.”⁶² The FDA requires that “valid scientific evidence” should be primarily from “well-controlled investigations,” but may also be from

partially controlled studies, studies and objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience with a marketed device, from which it can fairly and responsibly be concluded by qualified experts that there is reasonable assurance of the safety and effectiveness of a device under its conditions of use.⁶³

While the FDA regulations governing Class III devices acknowledge that “[t]he evidence required may vary according to the characteristics of the device, its conditions of use, the existence and adequacy of warnings and other restrictions, and the extent of experience with its use,” they specifically preclude “[i]solated case reports, random experience, reports lacking sufficient details to permit scientific evaluation, and unsubstantiated opinions.”⁶⁴

The “reasonable assurance of safety and effectiveness” requirement creates a potential conundrum for device manufacturers seeking pre-market approval: obtaining a PMA for the use of a Class III device requires data derived from clinical trials or other valid scientific evidence to demonstrate safety and effectiveness, but a Class III device first must be used without approval in order to obtain that required evidence. To resolve this problem, the FDA created the Investigational Device Exemption (IDE), which allows a device to be used in a clinical study limited to the collection of safety and efficacy data required to support a PMA application.⁶⁵ Investigators can submit an IDE for first-in-human studies, early feasibility studies, traditional feasibility studies, and pivotal trials (i.e., trials that serve as a basis of FDA approval), although the type of information included in the IDE will differ depending on the clinical study type.⁶⁶ From the data acquired from an IDE study, reviewed in conjunction with documentation pertaining to the device’s manufacture, design history, technical, and biological safety, the FDA determines whether the manufacturer has provided sufficient evidence

62. 21 C.F.R. § 860.7(c)(1), (e)(1) (2017).

63. *Id.* at § 860.7(c)(2), (e)(2) (2017).

64. *Id.* at § 860.7(c)(2).

65. *Id.* at § 812.2 (2016).

66. FDA, INVESTIGATIONAL DEVICE EXEMPTIONS (IDE) FOR EARLY FEASIBILITY MEDICAL DEVICE CLINICAL STUDIES, INCLUDING FIRST IN HUMAN (FIH) STUDIES 4 (2013), <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm279103> [<https://perma.cc/SG3J-ENJJ>].

of safety and effectiveness within the PMA parameters. An IDE was employed to assess the effect of DBS in the MCS in a clinical trial.⁶⁷

In some cases, a Class III device may be legally marketed without undergoing the PMA process. Class III medical devices are exempt from premarket approval if the manufacturer can demonstrate that the device is “substantially equivalent” in terms of intended use, technological characteristics, and performance testing to another legally marketed device that is not subject to PMA.⁶⁸ The determination of “substantial equivalence” is based on § 510(k) of the Food, Drug and Cosmetic Act, which also requires device manufacturers to notify the FDA of their intent to market a medical device.⁶⁹ A device that reaches market following the 510(k) “premarket notification” procedure is not considered “approved” by the FDA, but rather “cleared” to be commercially distributed in the United States.⁷⁰ The 510(k) clearance process “has been criticized as introducing additional risks to consumers, because the assumption that the device is ‘equivalent’ to another already on the market may be unsound.”⁷¹ Some neurotechnologies have been cleared through this process, including transcutaneous electrical nerve stimulators.⁷²

Rarely, a Class III device may bypass both the PMA and 510(k) processes. Humanitarian Use Devices (HUDs), that is, devices with intended use in diseases or conditions that affect fewer than 8000 individuals in the United States per year,⁷³ may reach market through a non-standard premarket approval procedure known as the Humanitarian Device Exemption (HDE).⁷⁴ The HDE authorizes marketing of a

67. Schiff et al., *supra* note 51; see also Joseph J. Fins & Nicholas D. Schiff, *Conflicts of Interest in Deep Brain Stimulation Research and Ethics of Transparency*, 21 J. CLINICAL ETHICS 125 (2010) (describing the experience of the study).

68. 21 C.F.R. § 807.92(a)(3) (2017).

69. *Id.* § 807.81 (2016).

70. *Premarket Notification 510(k)*, FDA (Mar. 3, 2018), [https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/How-to-Market-Your-Device/Premarket-Submissions/Premarket-Notification-510\(k\)-ucm2005718.htm](https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/How-to-Market-Your-Device/Premarket-Submissions/Premarket-Notification-510(k)-ucm2005718.htm) [<https://perma.cc/465V-BYJ7>].

71. Gail A. Van Norman, *Drugs, Devices, and the FDA: Part 2; An Overview of Approval Processes; FDA Approval of Medical Devices*, 1 JACC: BASIC TO TRANSLATIONAL SCI. 277, 280 (2016). The process has also been critiqued for not incentivizing innovation in device development. *Id.*

72. Leigh Anderson et al., *FDA Regulation of Neurological and Physical Medicine Devices: Access to Safe and Effective Neurotechnologies for All Americans*, 92 NEURON 943, 943 (2016). The Empi Active transcutaneous electrical nerve stimulator was cleared based on the predicate SELECT device. See Empi Active Transcutaneous Nerve Stimulator Summary, FDA Releasable 510(k) Premarket Notification Database (June 5, 2009) https://www.accessdata.fda.gov/cdrh_docs/pdf9/K090922.pdf [<https://perma.cc/3AU6-K7VU>].

73. 21 C.F.R. § 814.3(n) (2017).

74. *Id.* §§ 814.100-26.

HUD under the condition of individual and continuous IRB approval.⁷⁵ While regulations call for a thorough review of such requests by the FDA's Office of Orphan Product Development,⁷⁶ the approval threshold is lower for efficacy data requirements. Under the HDE procedure, the manufacturer only has to demonstrate that the device does not pose an unreasonable risk of harm, and that the probable benefits of using the device outweigh the risk of harm.⁷⁷ An HDE device is therefore exempt from the requirement of providing reasonable assurance of effectiveness, a constraint to which Class III devices approved under PMA or 510(k) are subjected.⁷⁸ The HDE was used to market a DBS device for OCD, and some scholars have argued that the HDE mechanism is sometimes misused to bypass the need for scientifically rigorous research.⁷⁹

B. Demonstrating Effectiveness? FDA Preference for Biomarkers in Drug Studies

Apart from general descriptions of what constitutes valid scientific evidence, FDA regulations do not provide directions for *how* a clinically meaningful effect should be determined. Aside from the indication that valid scientific evidence of device efficacy might be generated through "studies and objective trials without matched controls," the rules neither explain whether the FDA considers objective methods of assessment to be more valid than more subjective methods nor delineate how this data is weighed in the approval process. That information instead comes primarily from supplemental sources like official FDA guidance documents.

Guidance documents, while not legally binding, represent the FDA's "current thinking" on a particular topic.⁸⁰ For drugs, guidance documents strongly imply that the FDA prefers objective measures to more subjective evidence. The guidance document for submission of the efficacy section of the Common Technical Document for pharmaceutical registration, for instance, mentions several objective metrics that

75. *Id.* § 814.124.

76. *Id.* § 814.102.

77. *Id.* §§ 814.104(b)(2)-(3).

78. *Id.* § 814.100(b)(2).

79. Joseph J. Fins et al., *Misuse of the FDA's Humanitarian Device Exemption in Deep Brain Stimulation for Obsessive-Compulsive Disorder*, 30 HEALTH AFF. 302, 303-304, 306 (2011).

80. This is stated explicitly in every guidance document. *See, e.g.,* NEUROLOGICAL DEVICES, *supra* note 14, at 4. Some legal scholars have written critically about the role of administrative guidance, which can be viewed as helpful information or, if participants who do not follow the guidance will not obtain approval for a drug or device, as mandatory rules. *See* Richard A. Epstein, *The Role of Guidance in Modern Administrative Procedure: The Case for De Novo Review*, 8 J. LEGAL ANALYSIS 47, 70 (2016).

might be considered in analyzing clinical data from trials, including assay sensitivity, plasma concentration monitoring, and dose-blood level relationships—metrics available for assessing candidate drugs but not for assessing potential devices.⁸¹ Discussion of more subjective measurements such as scales or indices, on the other hand, is summarily addressed in the directive that “[v]alidation of any scales used should be discussed.”⁸² Another guidance document, describing the recommended content of the integrated summary of effectiveness section in a new drug application, characterized subjective measures—such as PROs—as posing “particular analytic challenges.”⁸³ A third guidance document, outlining the documentation usually submitted to the agency to support an effectiveness claim, asserts that “[c]learly appropriate endpoints that can be objectively assessed and are not dependent on investigator judgment (e.g., overall mortality, blood pressure, or microbial eradication) . . . are more readily interpreted than [are] more subjective endpoints such as cause-specific mortality or relief of symptoms.”⁸⁴

Much of the FDA’s guidance on drug applications favors biomarker evidence of efficacy.⁸⁵ Because biomarkers are widely considered to be the most objective, quantifiable medical values that science can measure reproducibly,⁸⁶ they are often chosen as primary endpoints, or main results, in assessments of clinical efficacy.⁸⁷ When used as substitutes for actual clinical endpoints (i.e., how a patient “feels, functions, or survives”⁸⁸), the biomarkers are termed “surrogate endpoints.”⁸⁹ Leading

81. FDA, GUIDANCE FOR INDUSTRY: M4E: THE CTD – EFFICACY 11-12, 25 (2001) [hereinafter M4E], <https://www.fda.gov/downloads/Drugs/Guidances/ucm073290.pdf> [perma.cc/HM64-LQWJ]. The CTD is the internationally agreed-upon format for the preparation of applications for the registration of pharmaceuticals for human use in participating countries, which includes the United States.

82. *Id.* at 12.

83. FDA, INTEGRATED SUMMARY OF EFFECTIVENESS: GUIDANCE FOR INDUSTRY 8 (2001), <https://www.fda.gov/downloads/drugs/guidances/ucm079803.pdf> [https://perma.cc/WSV3-Y3U8].

84. FDA, GUIDANCE FOR INDUSTRY: PROVIDING CLINICAL EVIDENCE OF EFFECTIVENESS FOR HUMAN DRUG AND BIOLOGICAL PRODUCTS 19 (1998), <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072008.pdf> [https://perma.cc/7PJQ-6TX4].

85. *See, e.g.*, M4E, *supra* note 81.

86. Identifying Potential Biomarkers for Qualifications and Describing Contexts of Use to Address Areas Important to Drug Development: Request for Comments, 80 Fed. Reg. 8089 (Feb. 13, 2015).

87. Strimbu & Tavel, *supra* note 3, at 2-3.

88. *Id.* at 2.

89. *Id.*; INST. OF MED. OF THE NAT’L ACADS., EVALUATION OF BIOMARKERS AND SURROGATE ENDPOINTS IN CHRONIC DISEASE 25-26 (Christine M. Micheel & John R. Ball eds., 2010).

experts have critiqued the use of surrogate rather than clinical endpoints in clinical trials, however, because while an intervention may affect a biomarker, a patient may experience no clinical benefit.⁹⁰ The FDA may, for example, approve cancer drugs because they reduce the size of tumors (biomarkers and surrogate endpoints), but the drugs then have no effect on patient mortality.⁹¹

Indeed, the importance of biomarkers has been emphasized through recent government initiatives such as the recent passage of the 21st Century Cures Act. In response, the FDA created a “Biomarker Qualification Program” for drug development,⁹² to facilitate biomarker tool development.⁹³ Recent studies also indicate the relative importance of biomarker evidence in FDA applications and approvals for new drugs. One study that characterized study endpoints for recent FDA-approved interventions found that pivotal trials using biomarkers as a primary endpoint comprised about 45% of all approved drug indications.⁹⁴ By contrast, clinical outcomes accounted for about a third,⁹⁵ and clinical scales for about 18%.⁹⁶

Evaluation and approval on the basis of biomarker evidence likely works differently for interventions targeting neuropsychiatric conditions. Unlike a clear correlation between a new statin and its effect on cholesterol level, which is a putative biomarker for cardiovascular health, for example, in many neurological conditions, connections between interventions and effects are less clear. This becomes even more complicated when implantable devices affect complex neurological circuits in various directions. With no clear biomarkers to gauge effects,

90. See, e.g., Avorn & Kesselheim, *supra* note 12, at 2474; Kesselheim & Avorn, *supra* note 12, at 581-82; Diana M. Zuckerman et al., *21st Century Cures Act and Similar Policy Efforts: At What Cost*, 351 *BMJ* no. h6122, Nov. 23 2015 at 1, 2.

91. Vinay Prasad et al., *The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-Analyses*, 175 *JAMA INTERNAL MED.* 1389, 1389-90 (2015); JAIN, *supra* note 25, at 722, 728.

92. FDA, *CDER Biomarker Qualification Program* (April 2, 2018), <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/default.htm> [<https://perma.cc/MAS8-J2M5>]. Even prior to the 21st Century Cures Act, however, an FDA research priority was developing biomarkers. JAIN, *supra* note 25, at 723.

93. Identifying Potential Biomarkers for Qualifications and Describing Contexts of Use to Address Areas Important to Drug Development: Request for Comments, 80 Fed. Reg. 8089 (Feb. 13, 2015). The responses to the request for comments can be found at <https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/UCM459168.pdf> [<https://perma.cc/UF7A-YQDY>].

94. Nicholas S. Downing et al., *Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005-2012*, 311 *J. AM. MED. ASS'N* 368, 371 (2014).

95. *Id.* at 371.

96. *Id.*

investigators are left to evaluate observable behaviors and patient reports. For example, one study relying on neuroimaging demonstrated the variable presentation of VS in five patients. This would indicate that neuroimaging would not be a reliable biomarker since the scans of these patients looked very different *even though behaviorally the patients all were in the VS*.⁹⁷

C. *The Role of PRO and PPI in Medical Device Applications*

The FDA's preference for biomarkers in pharmaceutical applications may thus give applicants pause when submitting more subjective evidence. It is unknown whether this preference extends to evidence of medical device effectiveness, however, because there is less FDA guidance for medical devices than for pharmaceuticals. The guidance that has been issued manages to simultaneously convey a preference for objective standards and carve out a space for more subjective metrics. For instance, one guidance document created to advise about design considerations for device studies noted that "[d]evice performance and clinical outcomes should be objectively measured with minimal bias," and "[r]elying on the subjective clinical assessments to determine an endpoint in a clinical outcome study is typically inadequate when more objective assessment methods exist."⁹⁸ In cases where more subjective assessments are appropriate, the guidance notes that "it is important to select a scoring assessment that is validated for the subject population and condition being treated, and consistent with the desired intended use."⁹⁹ However, it concludes by adding that "these more subjective measures are often used in conjunction with more objective assessments as part of a composite endpoint,"¹⁰⁰ suggesting that more subjective measures on their own may be insufficient and it is preferable to use them with objective assessments. Even when objective assessments do not exist, and a subjective assessment must be used, the guidance document suggests making evaluation of evidence as objective as possi-

97. Nicholas D. Schiff et al., *Residual Cerebral Activity and Behavioral Fragments Can Remain in the Persistently Vegetative Brain*, 125 *BRAIN: J. NEUROLOGY* 1210, 1228-29 (2002).

98. FDA, DESIGN CONSIDERATIONS FOR PIVOTAL CLINICAL INVESTIGATIONS FOR MEDICAL DEVICES: GUIDANCE FOR INDUSTRY, CLINICAL INVESTIGATORS, INSTITUTIONAL REVIEW BOARDS AND FOOD AND DRUG ADMINISTRATION STAFF 24-25 (2013), <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM373766.pdf> [<https://perma.cc/2YZ7-2V6Y>].

99. *Id.* at 25.

100. *Id.*

ble, perhaps by using “an independent adjudication committee” to evaluate an endpoint (e.g., for interpreting a radiograph, which is an image produced through radiation).¹⁰¹

PROs are one of the main forms of less objective, non-biomarker evidence that also may be submitted as part of a drug or device approval application (the other being standardized clinical assessments such as the CRS-R). The FDA encourages submission of PRO data as part of the device application, and also permits its use as a primary or secondary endpoint.¹⁰² Indeed, studies of medical devices may collect both biomarker data and PROs—for example, the effects of cardiovascular devices targeting heart failure can be measured by Kansas City Cardiomyopathy Questionnaire (a PRO that assesses quality of life, symptoms, and function) and biomarkers.¹⁰³

Devices can be approved on the basis of PROs,¹⁰⁴ but as stated in one guidance document, before the FDA will accept a more subjective assessment as an adequate metric of device effectiveness, the assessment must be shown to have:

1. content validity (i.e., it must measure the aspects of the disease that are relevant to the study and to patients);
2. construct validity (i.e., it must show that the “documented relationships between results gathered using the instrument and results gathered using other measures are consistent with pre-existing hypotheses concerning those relationships”);¹⁰⁵
3. reproducibility (i.e., it must demonstrate test-retest reliability, which yields consistent, reproducible estimates of true treatment effect); and
4. responsiveness (i.e., it must have the ability to detect change in patient condition).¹⁰⁶

101. *Id.*

102. CDRH, *supra* note 10, at 8.

103. *Id.* app. at 2-3.

104. See, e.g., Cartiva, Inc., *A Study to Evaluate the Effectiveness and Safety of an Implant in the Treatment of Osteoarthritis of the Great Toe*, CLINICALTRIALS.GOV (Nov. 15, 2016) <https://www.clinicaltrials.gov/ct2/show/NCT00969969?term=NCT00969969&rank=1> [<https://perma.cc/VVK3-SJ5Q>] (listing a study approved on the basis of VAS and FAAM, two different PROs); FDA, CDRH PRO COMPENDIUM, (2017) <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHVisionandMission/UCM588577.xlsx> [<https://perma.cc/X56V-LADT>] (listing many other studies approved on basis of PROs).

105. FDA, GUIDANCE FOR INDUSTRY: PATIENT-REPORTED OUTCOME MEASURES: USE IN MEDICAL PRODUCT DEVELOPMENT TO SUPPORT LABELING CLAIMS 19 (2009) [hereinafter PRO], <https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf> [<https://perma.cc/Y7DJ-AEP9>].

106. *Id.* at 20. See also Laurie B. Burke et al., *The Use of Patient-Reported Outcome Measures in the Evaluation of Medical Products for Regulatory Approval*, 84 CLINICAL

Several widely used scales have been shown to be internally and externally valid, reliable, and responsive.¹⁰⁷

Although the FDA has not specified how evidence establishing each type of validity is weighed, guidance suggests that evidence pertaining to content validity is the most important. Because the FDA “evaluate[s] instrument adequacy to measure the concept represented by the labeling claim,” establishing content validity is a threshold requirement.¹⁰⁸ Moreover, “[e]vidence of other types of validity (e.g., construct validity) or reliability (e.g., consistent scores) will not overcome problems with content validity.”¹⁰⁹ The guidance document discusses, in great detail, eight types of documentation that sponsors are encouraged to submit in order to support the instrument’s content validity.¹¹⁰ For instance, in describing the submission of information that establishes the consistency of item response options with their purpose and intended use, the guidance document includes examples of item response options in PROs.¹¹¹ In describing documentation that addresses how the instrument minimizes administrator and respondent burden, it provides examples of factors that cause “physical, emotional, or cognitive strain on patients [that] decrease the quality and completeness of PRO data.”¹¹²

By contrast, descriptions of documentation pertaining to construct validity, reproducibility, and responsiveness are relatively brief. Discussion of these measures is consolidated into one section addressing “additional measurement properties,” and very few examples of appropriate documentation are given.¹¹³ While the FDA is silent on how they assess these criteria, the FDA’s emphasis on establishing content validity before evaluating other measurement properties indicates that, while these other properties are certainly necessary to the successful development of a PRO instrument, the bulk of FDA review is dedicated to evaluating content validity. Indeed, a recent study found that while almost half of new drug applications in the years 2006-2010 included PRO evidence, just a quarter of these applications were granted PRO

PHARMACOLOGY & THERAPEUTICS 281, 283 (2008).

107. See, e.g., Brett J. Deacon & Jonathan S. Abramowitz, *The Yale-Brown Obsessive Compulsive Scale: Factor Analysis, Construct Validity, and Suggestions for Refinement*, 19 J. ANXIETY DISORDERS 572, 583 (2005); Giacino et al., *supra* note 4, at 2020; Claudia Ramaker et al., *Systematic Evaluation of Rating Scales for Impairment and Disability in Parkinson’s Disease*, 17 MOVEMENT DISORDERS 867, 870, 873 (2002).

108. PRO, *supra* note 105, at 12.

109. *Id.* (emphasis omitted)

110. *Id.* at 12-18.

111. *Id.* at 15.

112. *Id.* at 17.

113. *Id.* at 18-20.

label claims, mostly because the FDA flagged issues with content validity of the PRO instrument.¹¹⁴ This study demonstrates that the FDA demands statistical rigor even from more subjective types of evidence.

Another form of subjective evidence that may be submitted to the FDA as part of a device application is Patient Preference Information (PPI), which the FDA defines in a guidance document as “qualitative or quantitative assessments of the relative desirability or acceptability to patients of specified alternatives or choices among outcomes or other attributes that differ among alternative health interventions.”¹¹⁵ Put more simply, the FDA values patient views in assessing the risks and benefits of medical devices.¹¹⁶ Whereas PROs measure a patient’s perception of their health and an intervention’s effect, PPI studies measure patient preference for type of therapeutic intervention or attributes of such intervention.¹¹⁷ The FDA may consider PPI, along with clinical and nonclinical evidence, in its benefit-risk determinations for PMAs, HDE applications, and de novo¹¹⁸ reviews of medical devices,¹¹⁹ but notes that PPI may *not* be useful for all interventions, and is likely to be more relevant when patient treatment decisions are “preference sensitive.”¹²⁰

114. Carla DeMuro et al., *Reasons for Rejection of Patient-Reported Outcome Label Claims: A Compilation Based on a Review of Patient-Reported Outcome Use Among New Molecular Entities and Biologic License Applications: 2006-2010*, 15 VALUE HEALTH J. 443, 445-46 (2012). Other issues for rejection were related to study design, such as bias due to open-label studies. *Id.* We were unable to locate similar information for results of PRO submission in the medical device context.

115. PPI, *supra* note 34, at 6.

116. *Id.* at 1. The FDA distinguishes “patient input,” “patient perspectives,” and “patient preference information.” “Input” is the broadest category, and includes all patient feedback, from qualitative comments to quantitative PROs. “Perspectives” refer to “information relating to patients’ experiences with a disease or condition and its management.” The guidance document is about PPI, which is a subset of patient perspectives, solely about the risk-benefit profile of an intervention. *Id.* at 3. The FDA may take into account perspectives others than patients, including family members (“care-partners”) and clinicians, if relevant. *Id.* at 6.

117. *Id.* at 8.

118. “De novo” devices are those that are initially classified as Class III devices because there is no predicate device, but are reclassified after sponsor petition because they have lower risk than a Class III device. Van Norman, *supra* note 71, at 278.

119. PPI, *supra* note 34, at 3-4.

120. *Id.* at 3. Patient decisions regarding treatment options are “preference sensitive” when:

- 1) multiple treatment options exist and there is no option that is clearly superior for all patients; 2) when the evidence supporting one option over others is considerably uncertain or variable; and/or 3) patients’ view about the most important benefits and acceptable risks of a technology vary considerably within a population, or differ from those of healthcare professionals.

Id.

For the FDA to consider PPI in reviewing an application, PPI needs to be considered valid scientific evidence. In evaluating whether a PPI study constitutes valid scientific evidence, the guidance document lists eleven recommended study qualities related to three domains: patients, good study design, and good study conduct.¹²¹ The FDA uses qualitative PPI to help identify “which outcomes, endpoints or other attributes are valued most by patients and which factors affect patients’ perspectives on risk and benefit.”¹²² In contrast, quantitative PPI provides estimates of “how much different outcomes, endpoints or other attributes are valued by patients, and the tradeoffs that patients state or demonstrate they are willing to make among them.”¹²³

The guidance document lists encouraging submission of PPI to aid in FDA decision making as the first objective of the guidance document, but also emphasizes that PPI submission does not change any review standards for device applications and PPI submissions remain voluntary for device manufacturers.¹²⁴ It is unclear how many submissions include PPI data, but during a webinar on this particular guidance document about PPIs, a participant posed the question of “how many patient preference studies have contributed to either a clearance or approval to date” and the Center for Devices and Radiological Health (CDRH) Assistant Director for Strategic Programs who oversaw the creation of the guidance did not give a quantitative answer but merely repeated that the use of patient perspectives is an increasing trend. The influence of PPI on device clearance or approval to date is thus unknown. The Assistant Director further noted that the use of quantitative patient preference studies “are relatively more rare,” which is perhaps a hint that if more quantitative patient preference studies are conducted they may contribute to future approvals.¹²⁵

III. POLICY RECOMMENDATIONS

Given that the FDA prefers the most objective evidence of efficacy, investigators are incentivized to prioritize objective endpoints, even if biomarkers do not result in clinical benefit. Objective evidence should be collected not just to satisfy regulators, however. In the context of medical device studies, evidence of efficacy should be collected with the

121. A full list of criteria the FDA considers in evaluating the quality of PPI is in the guidance document. *Id.* at 11-15.

122. *Id.*

123. *Id.* at 6. PPI can assess safety, efficacy, and other considerations such as how a device is designed for use. *Id.*

124. *Id.* at 2.

125. Irene Aihie et al., *Webinar: Final Guidance on Patient Preference Information*, FDA 17 (Sept. 27, 2016), <https://www.fda.gov/downloads/Training/CDRHLearn/UCM523904.pdf> [<https://perma.cc/GSN4-8XNP>].

purpose of increasing the likelihood that only truly clinically effective Class III devices make it to the market. Such devices pose significant risks to patients, and in the case of implantable devices, there may be difficulty removing an implanted device should it be determined not to be effective.¹²⁶

It is not clear whether or how the FDA preference for biomarker evidence in drug applications extends to review of medical device applications. Although the FDA acknowledges the need for and the practice of using less objective metrics in device studies, this acknowledgment does not preclude the possibility that studies that rely exclusively on patient reports and clinical assessments are disadvantaged in the approval process with greater frequency or encounter other barriers in the path to approval. More subjective evidence, while often proven to be scientifically sound before use as clinical trial endpoints, still relies on administrator judgment as well as patient honesty, transparency, and accuracy. Concerns about deficiency of any one of these factors may lead to reviewer skepticism about the validity of these measures, regardless of their methodological acceptance from the clinical or scientific community.

We would recommend that the FDA not discount studies, especially of neurotechnologies, that do not include biomarker evidence of medical device effectiveness. As noted previously, reliable biomarkers do not exist for many neuropsychiatric conditions including DOC. It is not always possible to assess efficacy with neuroimaging or EEG biomarkers for DOC, which means that valid, reliable behavioral assessments such as the vetted CRS-R remain valuable, if not essential, given our current state of knowledge and ability to engage in assessment. Further, patient or surrogate input, even if not part of the PRO questionnaire, may also be helpful in evaluating an intervention's effect. Finally, experience with some drugs approved on the basis of surrogate endpoints has demonstrated that relying solely on biomarkers is sometimes problematic for establishing clinical benefit.¹²⁷ This would likely be true for medical devices relying on biomarkers as well.

We do not mean to suggest, however, that biomarkers have no value. They do, and modern neuroscience should strive to develop them. They may be appropriate surrogates for clinical outcomes in some cases. Furthermore, biomarkers may elucidate a pathology that is not yet known to clinicians or investigators.¹²⁸ Additionally, in the case of DOC, there

126. Given the burden of implantation of neurotechnology, it is imperative to ensure effectiveness beyond a placebo effect. See also Megan S. Wright, *A Case for Randomized, Double-Blinded, Sham-Controlled Class III Medical Device Trials*, 34 *YALE L. & POL'Y REV.* 199, 207 (2016) (arguing that it is often unethical to market Class III medical devices that have not been studied using placebo-controlled trial designs).

127. See, e.g., Prasad et al., *supra* note 91, at 1389-92.

128. For example, EKGs may show evidence of pathology even in the absence of

are limitations to clinician assessments and patient reports, because of the possibility of covert consciousness (i.e., an inability to behaviorally signal consciousness) making neuroimaging a useful corrective source of information. We would argue, however, that if biomarkers are submitted as evidence of efficacy, they need to be validated, reflect clinical plausibility, and be analyzed in a statistically sound manner in order to make it more likely that they will predict clinical benefit.¹²⁹

Because no type of evidence in isolation is fully adequate, we need to use all available data to assess the efficacy of an intervention,¹³⁰ especially in the context of neuropsychiatric conditions where scientific knowledge is still in relative infancy. Thus, both emerging biomarkers and validated, reliable clinician assessments and patient reports are important types of evidence to demonstrate efficacy for neurotechnology for DOC.

Given that it seems that the regulatory trend is toward preferring biomarkers, however, the remaining focus of the Article will be on the role of more subjective types of evidence in FDA evaluation of medical devices. Specifically, we would recommend that the FDA require submission of results from PROs and PPI (when possible) to evaluate drugs and devices, especially novel neurotechnology. Adopting such a policy is consistent with the documented benefits of pluralistic ways of knowing.¹³¹

A. *Require Submission of PPI and PRO*

Currently, submission of PPI and PRO as part of a medical device application to the FDA is encouraged but *voluntary*.¹³² But as we have discussed, multiple ways of assessing effectiveness of an intervention can be scientifically superior to relying solely on biomarkers. This is especially true when there are no valid biomarkers associated with a condition, and patient provided information (both quantitative as in the case of PROs and also qualitative) may be useful in assessing both safety and efficacy of an intervention. We thus recommend that investigators be required to submit PPI and PRO when feasible to do so.¹³³

chest pain, a clinical symptom. Insel, *supra* note 15.

129. See Goldfine et al., *supra* note 28, at 290-91 (discussing problems with EEG evidence of state of consciousness for patients with DOC).

130. See SOLOMON, *supra* note 18.

131. *Id.*

132. PPI, *supra* note 34, at 6.

133. It may not always be feasible to collect such information. If a patient is unconscious or unable to communicate, collecting PRO and PPI will not be possible (unless a proxy assists).

1. Regulatory and Scientific Considerations

Questions arise with respect to how the FDA should evaluate patient-provided information in the drug and device approval process, and it is likely that the evaluation of patient-provided evidence will differ depending upon whether and what type of more objective evidence is available to assess efficacy. For some conditions without well-established biomarkers, such as chronic pain,¹³⁴ patient-provided information in the form of valid and reliable scales may constitute sufficient evidence for evaluating an intervention in all phases of research,¹³⁵ particularly when the intervention is a drug, which can be discontinued should later studies find that it is not effective (beyond a placebo effect). For other conditions without existing biomarkers, less objective evidence may be sufficient for assessing signs of efficacy in nonpivotal feasibility studies. However, once a Class III medical device or, more specifically, implantable neurotechnology such as DBS, is in pivotal trials, regulators should generally require additional objective evidence of efficacy, including results from well validated, reliable clinical assessments that are not mediated through patient interpretation of the intervention's effect (e.g., CRS-R). These results should be considered prior to approval being granted given the relative difficulty of removing a non-efficacious implanted medical device. The calculus may change, however, if there are no other available treatments for the particular condition or no viable objective means of assessment.

Non-biological, patient-provided evidence is also useful when there are emerging biomarkers, as this evidence can enhance the reliability of biomarkers in all stages of research.¹³⁶ For example, when attempting to identify biomarkers for depression using neuromodulation,¹³⁷ the intervention can also be compared with existing instruments used to identify severity of depression such as the Beck Depression Inventory¹³⁸ (a PRO) as well as other types of patient reports. In the case of severe brain injury and DOC, interventions such as DBS can be assessed by examining emerging neuroimaging biomarkers and comparing this objective evidence against the CRS-R, patient and family reports, and clinical outcomes such as death. The comparison between biomarker evidence and more subjective evidence may provide greater evidence of

134. There are attempts to find brain-based biomarkers for pain using fMRI technology. See, e.g., Marianne C. Reddan & Tor D. Wager, *Modeling Pain Using fMRI: From Regions to Biomarkers*, 34 NEUROSCIENCE BULL. 208 (2018).

135. It is also important to note that PROs and even qualitative findings can help inform the construction of more systematic scales.

136. Patient reports can also assist in hypothesis generation for future studies.

137. Allison C. Waters & Helen S. Mayberg, *Brain-Based Biomarkers for the Treatment of Depression: Evolution of an Idea*, 23 J. INT. NEUROPSYCHOL. SOC. 870 (2017).

138. See Aaron T. Beck, et al. *Psychometric Properties of the Beck Depression Inventory: Twenty-five Years of Evaluation*, 8 CLINICAL PSYCHOL. REV. 77 (1988) (describing reliability and validity of this scale).

the biomarker's utility, or help refine the biomarker. And, as stated earlier, biomarkers may provide important information even if they do not correlate perfectly with the clinical outcome of interest. For example, neuroimaging may show other evidence of pathology, which may be clinically relevant, or help reveal an underlying mechanism or brain circuit not previously known to researchers, which is scientifically relevant and could lead to improved treatments in the future. Conversely, more subjective data could lead to hypothesis generation that could be scientifically evaluated.¹³⁹

There is also a role for more subjective evidence even if there are well-established, valid biomarkers for a condition. One could imagine a study in which biomarkers do not show evidence of efficacy but clinical assessments or patient-provided information indicates effectiveness. In such a study, if biomarkers were the primary endpoint, what could be promising intervention may never be approved. Alternatively, the biomarkers could show evidence of efficacy, but there is no effect shown through use of clinical assessments or PROs or other patient-provided information. In this latter case, there is a physiologic effect, but no patient-centered clinical benefit.¹⁴⁰ The FDA acknowledges this possibility in guidance related to neurological devices.¹⁴¹ An example of this occurred when a patient said to be in the VS for fifteen years had a vagus nerve stimulator implanted, which by report moved him from the VS into the MCS with neuroimaging changes (i.e., emerging biomarker), but offered no additional improvement in consciousness over a period of nine months.¹⁴² While experts were intrigued by this case, they agreed this report did not provide sufficient evidence to use this intervention on persons with DOC as a therapy given the effect did not result in a clinical benefit (i.e., further state changes) and that other unobserved phenomena could have caused the movement from the VS to MCS. Nonetheless, the case provided insight into DOC and advanced the state of the science.¹⁴³

139. See, e.g., Joseph J. Fins, *The Reagan Diaries Reconsidered*, 48 J. ALZHEIMER'S DISEASE 59 (2015).

140. For a discussion of concerns about approving drugs solely on the basis of biomarkers, see Kesselheim & Avorn, *supra* note 12, at 581-82; Zuckerman et al., *supra* note 90. This distinction between effect and patient-centered benefit also has been discussed extensively in the medical futility literature. See, e.g., Lawrence J. Schneiderman et al., *Medical Futility: Its Meaning and Ethical Implications*, 112 ANN. INTERNAL MED. 949, 950 (1990). Another example of an effect without benefit is when a person is hypotensive and their blood pressure is raised (an effect), but they do not wake up (no benefit). Lawrence J. Schneiderman, *The Futility Debate: Effective versus Beneficial Intervention*, 42 J. AM. GERIATRICS SOC'Y 883, 883 (1994).

141. NEUROLOGICAL DEVICES, *supra* note 14, at 6 (explaining that "biological markers may not be accompanied by clinically meaningful observable changes").

142. Martina Corrazol et al., *Restoring Consciousness with Vagus Nerve Stimulation*, 27 CURRENT BIOLOGY R994 (2017) (based on neuroimaging and EEG evidence).

143. Michael Price, *Experimental Nerve-Stimulation Therapy Partially Revives Man in Long-term Vegetative State—But Experts Urge Caution*, SCIENCE (Sept. 25, 2017),

If there is such a conflict between biomarkers and supplementary, more subjective evidence with respect to indication of efficacy, this may be a sign that the understanding of the condition or the intervention is incomplete, that the length of the trial was insufficient to demonstrate effect, or that what matters to participants in the trial and influences their reports of effectiveness is not what is being assessed with the objective evidence. This preliminary information can be critical to additional hypothesis generation and the advancement of science longitudinally.¹⁴⁴

While the drug or device may progress past early stage research because the results on at least one measure of efficacy are scientifically suggestive and can point to evidence of mechanism, the FDA may choose not to approve the intervention in later stages of research if the conflict is not reconciled or explained.¹⁴⁵ Regardless, however, just as researchers are required to identify the endpoints of their study in advance, the FDA should be transparent before a trial begins about how possible conflicts between multiple endpoints in a single trial will be adjudicated.

2. Ethical Considerations

There are also ethical considerations that require that investigators collect PPI. If researchers and sponsors aspire to respect the personhood of research participants, investigators should ask research participants about their experiences in the trial, and how participants assess risks and benefits of novel technologies.¹⁴⁶

Respondent answers may reveal that some clinical trials are not focusing on endpoints that are important to patients.¹⁴⁷ For example, persons with severe brain injury and subsequent DOC may prefer that a

<http://www.sciencemag.org/news/2017/09/experimental-nerve-stimulation-therapy-partially-revives-man-long-term-vegetative-state> [https://perma.cc/SYT2-PPFK]. It is important to note, however, that in this case, even a state change from vegetative to minimally conscious may be viewed highly positively by families of the person with the DOC.

144. See Joseph J. Fins, *A Review of Making Medical Knowledge*, NOTRE DAME PHIL. REV. (2016), <https://ndpr.nd.edu/news/making-medical-knowledge/> [https://perma.cc/Y9E4-L95P] (reviewing MIRIAM SOLOMON, MAKING MEDICAL KNOWLEDGE (2014)).

145. See JAIN, *supra* note 25, at 722 (providing an example of how imaging biomarkers are used as primary outcomes in early drug research for Multiple Sclerosis but not in pivotal trials because of the lack of consistent correlation with clinical assessments).

146. See NEUROLOGICAL DEVICES, *supra* note 14, at 9 (advising that patients' perspectives be taken into consideration with respect to risk-benefit assessment); Anderson et al., *supra* note 72, at 946.

147. See PPI, *supra* note 34, at 5 (describing how patients can provide information on outcomes important to them).

device aid in restoring motor function, but a device is designed to restore cognitive function (or vice versa).¹⁴⁸ As part of the social contract—that patients participate in clinical research if not for direct benefit, then at least the prospect of benefit for patients with their same condition—it seems necessary that patient preferences should be taken seriously by investigators and regulators. This stance is consistent with the 21st Century Cures Act, which explicitly calls for greater patient engagement.¹⁴⁹ As some have noted, “patients seek treatment for their diseases, not for the numerical measures that frequently but not perfectly correlate with their illnesses.”¹⁵⁰ Having recognized the importance of the patient’s perspective, we would not, however, suggest that patients inappropriately direct scientific and research agendas. This could lead to health disparities if only some groups are assertive about their preferences.¹⁵¹ This would distort research agendas as patients are not subject-matter experts.

There are thus both scientific and ethical reasons to collect PPI and PROs.¹⁵² Therefore, we would advocate for making submission of such information (both qualitative and quantitative) a mandatory part of the application to the FDA if feasible to collect such data.

3. *Case of Disorders of Consciousness*

The benefits of collecting such data can be seen clearly in the case of interventions targeted toward neuropsychiatric disorders generally and DOC specifically. With respect to submitting and evaluating PPI, it may be the case that given the severity of the brain injury and the significant disability it results in, patients or their surrogates may be more willing to tolerate burdensome, risky interventions such as DBS that may offer uncertain or only modest benefit. This is similar to the case in which patients with advanced end stage cancer may be willing to tolerate toxicity for additional days/months of life. But we will only know what matters to patients if we ask and submission of PPI is mandatory. With respect to PRO and other, more qualitative patient reports, these types of subjective evidence may demonstrate an effect that biomarkers cannot, especially in early stages of research. For example, neuroimaging may not reveal any changes to brain activity with DBS, but the patient’s

148. This was the case for one participant in the study of DBS for patients with severe brain injury. Schiff, *supra* note 13.

149. See, e.g., 21st Century Cures Act § 3001-03, 21 U.S.C. § 360bbb-8c (2016).

150. Strimbu & Tavel, *supra* note 3, at 2.

151. See, e.g., Norman Sharpless, Commentary at the Solomon Center for Health Law and Policy Conference: The Politics, Policy and Law of Cancer Conference (Feb. 9, 2018) (describing this potential outcome in the context of cancer).

152. Other scholars have also noted that PROs can be useful, especially when it comes to understanding symptoms, so long as they are collected in randomized controlled trials. Kesselheim & Avorn, *supra* note 12, at 582.

family may notice that the patient moves some part of their body differently and can communicate this to researchers. Investigators should be called upon to incorporate patient provided information in some way in their study design, especially for feasibility or proof-of-principle studies, including the perspectives of family members and patient surrogates, which may also be instructive in this evolutionary epistemic process.¹⁵³

4. *Additional Considerations*

Some may argue that collecting more data slows down the approval process, and also increases costs, which may disincentivize research and development. However, if collecting such data is useful for demonstrating efficacy or justifying risks for a subpopulation of patients, which could contribute to the success of the application to the FDA, these benefits may outweigh administrative and economic costs associated with increased data collection.¹⁵⁴ This is especially true for medical conditions without biomarkers or reliable biomarkers. Furthermore, there is evidence that more applications are including PROs,¹⁵⁵ which demonstrates that this requirement would not constitute a significant disincentive. Finally, there may be a benefit to sponsors and investigators if they prioritize patient input—patient advocacy can help play a role in clinical trial recruitment, which is notoriously difficult for some types of trials,¹⁵⁶ like those for implantable neurotechnology. Additionally, such information will also provide better understanding of adherence for therapeutic interventions that come to market.

B. *Permit Proxy Reports in PPI and PRO*

Because we are arguing for mandatory submission of PPI and PRO, it is necessary to address the question of whether proxy reports are permitted for research participants' experiences with illness or disability and in the clinical trial; preferences; assessment of risks and benefits; and other vital information. Proxy reports are when a representative of

153. JOSEPH J. FINS, *RIGHTS COME TO MIND: BRAIN INJURY, ETHICS, AND THE STRUGGLE FOR CONSCIOUSNESS* (2015); *see also* Joseph J. Fins & Jennifer Hersh, *Solitary Advocates: The Severely Brain Injured and Their Surrogates*, in *TRANSFORMING HEALTH CARE FROM BELOW: PATIENTS AS ACTORS IN U.S. HEALTH POLICY* 21 (B. Hoffman et al. eds., 2011).

154. Patient reports can also advance science and translational work, an important benefit even if not the mandate of the FDA, an agency charged with regulating drugs and devices to protect the public health.

155. Indeed, the vast majority of medical device applications currently contain PROs. CDRH, *supra* note 10, at 8.

156. INSTITUTE OF MEDICINE, *PUBLIC ENGAGEMENT AND CLINICAL TRIALS: NEW MODELS AND DISRUPTIVE TECHNOLOGIES* 17-30 (2012) (describing recruitment challenges).

the patient speaks on their behalf. The FDA guidance on the form that PPI and PRO should take is that individuals have to be able to provide information themselves, and the FDA advises against proxy reports in the case of cognitively impaired research participants.¹⁵⁷

This guidance is inconsistent with existing disability rights law, policy, and ethics, all of which promote equality and require reasonable accommodations so that persons with disabilities can fully participate in all aspects of society,¹⁵⁸ which in our view includes the research enterprise. This guidance discriminates against persons who are unable to reliably communicate and need a proxy to assist with relaying information, or who prefer communicating via a proxy. Precluding this assistance in research and treating such persons differently because they have disabilities is discriminatory. We thus suggest that the FDA revise its guidance to explicitly permit submission of proxy reports, as long as such proxy reports are justified by investigators as a reasonable accommodation for persons with disabilities.

Examining the case of patients and research participants with DOC makes clear why proxy reports should be permitted. With respect to PPI, proxies can relay information to regulators about what risks and benefits matter to the person with a DOC (both their preferences prior to their brain injury and their current experiential interests). Proxies may also be helpful in ensuring compliance and adherence. And, given the nature of the disorder, not permitting such reports means that important qualitative information relevant to assessing the risk-benefit ratio is lost. Furthermore, if proxies are not permitted to assist in the collection of PRO, then investigators may not collect PROs for a population that cannot reliably communicate even though the proxy report may be sufficient to generate a quality PRO.¹⁵⁹ Changing the guidance to permit proxy reports would benefit not only persons with DOC who participate in clinical trials, but also other groups of people with neuropsychiatric disorders such as dementia or other conditions that lead to lack of or fluctuating capacity to provide feedback.

157. PRO, *supra* note 105, at 21. This is a change from previous practice in which proxy reports were encouraged to be submitted along with patient reports. FDA, PATIENT REPORTED OUTCOME MEASURES: USE IN MEDICAL PRODUCT DEVELOPMENT TO SUPPORT LABELING CLAIMS 22-23 (2006), reprinted in FDA, *Patient Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*, 4 HEALTH QUALITY LIFE OUTCOMES 79, 92 (2006).

158. Americans with Disabilities Act, 42 U.S.C. §§ 12101-213 (1990). See also Joseph J. Fins & Megan S. Wright, *Rights Language and Disorders of Consciousness: A Call for Advocacy*, 32 BRAIN INJ. 670 (2018); Megan S. Wright & Joseph J. Fins, *Rehabilitation, Education, and the Integration of Individuals with Severe Brain Injury into Civil Society: Towards an Expanded Rights Agenda in Response to New Insights from Translational Neuroethics and Neuroscience*, 16 YALE J. HEALTH POL'Y L. & ETHICS 233, 261-267 (2016).

159. See also Fins & Hersh, *supra* note 153; FINS *supra* note 153.

It is important to note that permitting proxy reports for PPI and PRO may introduce bias in the data. Indeed, many studies have shown discordance between proxies and patients in surrogate healthcare decision making, for example.¹⁶⁰ And in the documented case of a woman with a DOC, her proxy thought she was frustrated but when the patient was asked and given eye tracking software to assist with generating a response, it turned out that she was actually happy.¹⁶¹ Pilot studies thus should be conducted with patient and proxy pairs to determine the quality of proxy-provided patient information. If there is good reliability between patient and proxy in the PRO, then this may counter any scientific objection to the use of proxies. Indeed, one study demonstrated no average differences in proxy-patient agreement on some items of a PRO assessing quality of life post-stroke.¹⁶² Because there is room for error in proxy interpretation of an incapacitated person's preferences and experiences, we caution that proxy generated PROs should not be the sole basis for FDA approval of a drug or device, but for reasons described above, argue that proxy generated PROs be permitted.

If it turns out that for this population of patients, it is not possible to collect PROs, there is still a scientific role for proxies to report qualitative data on patient's experience with brain injury and subsequent DOC, and how the patient seems to be reacting to the tested intervention, and these data should be analyzed and reported so that this information can also be evaluated and compared to other types of assessments. Indeed, it is part of the protocol for patient evaluation at the Consortium for the Advanced Study of Brain Injury to have their proxies asked what the patient with a DOC is like when they are alert, so that the investigators have a baseline.¹⁶³

160. See, e.g., Stephen C. Hines et al., *Improving Advance Care Planning by Accommodating Family Preferences*, 4 J. PALLIATIVE MED. 481 (2001) (finding discordance between individual preferences and surrogate decisions).

161. FINS, *supra* note 153, at 260.

162. Allan J. Kozlowski et al., *Agreement Between Responses from Community-Dwelling Persons with Stroke and Their Proxies on the NIH Neurological Quality of Life (Neuro-QoL) Short Forms*, 96 ARCHIVES PHYSICAL MED. & REHABILITATION 1986, 1990-91 (2015).

163. *Neuroethics and Disorders of Consciousness: A Qualitative and Quantitative Assessment of Clinical Practice and Healthcare System Epidemiology/Barriers to Care*, WEILL CORNELL MEDICINE CONSORTIUM FOR THE ADVANCED STUDY OF BRAIN INJURY (CASBI) (May 30, 2017), <https://casbi.weill.cornell.edu/neuroethics-and-disorders-consciousness-qualitative-and-quantitative-assessment-clinical-practice> [<https://perma.cc/Q8UA-ENH7>] (Screening Form, Protocol #0606008614, for this study includes such questions for proxies).

C. *Transparency in How PPI and PRO Assessed During Regulatory Review*

Although federal law requires the FDA to “solicit the views of patients” and “consider the perspectives of patients during regulatory discussions,”¹⁶⁴ there is currently a lack of transparency about how some forms of patient perspectives such as PPI and PRO are used, evaluated, and weighed by the FDA throughout the regulatory process: from the clinical trials, to the drug, device, or biologic approval process (including approval only for specific uses or specific patient populations), and concluding with any post-market surveillance. For example, how is a study with a PRO as a primary endpoint evaluated? What about when a PRO is used as a co-primary endpoint? Or part of a composite endpoint with more objective measures? Or as a secondary endpoint?

As others have argued (e.g., CDRH, patient advocacy groups, researchers),¹⁶⁵ we too advocate for transparency in this realm. While sponsors and investigators have pre-submission meetings with the FDA to discuss data collection and study design,¹⁶⁶ which may answer some of the investigators’ questions, by the time a research team is prepared to have this meeting, they have already invested significant resources. Transparent policies in the form of guidance documents, rather than case-by-case confidential review of projects, will reduce uncertainty and may also increase all stakeholders’ trust in the regulatory process.¹⁶⁷ Additionally, the FDA believes transparency will also decrease regulatory review time,¹⁶⁸ and leading policy experts believe that transparency will speed innovation.¹⁶⁹

Transparency about how PROs and PPI are evaluated for a particular intervention is important for all types of diseases, but transparency may be especially important for neuropsychiatric conditions. This is be-

164. 21 U.S.C. § 360bbb-8c(a)(1) (2016).

165. See, e.g., CDRH, *supra* note 10; FDA, CENTER FOR DEVICES & RADIOLOGIC HEALTH: 2016-2017 STRATEGIC PRIORITIES 8, <https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHVisionandMission/UCM481588.pdf> [<https://perma.cc/P68G-YVFL>] (describing transparency as a goal); Anderson et al., *supra* note 72; Joshua M. Sharfstein et al., *Blueprint for Transparency at the U.S. Food and Drug Administration: Recommendations to Advance the Development of Safe and Effective Medical Products*, 45 S2 J. L. MED. & ETHICS 7 (2017).

166. Van Norman, *supra* note 71, at 278-79; FDA, GUIDANCE FOR INDUSTRY AND FOOD AND DRUG ADMINISTRATION STAFF: REQUESTS FOR FEEDBACK ON MEDICAL DEVICE SUBMISSIONS (2017), <https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM311176.pdf> [<https://perma.cc/4LVT-SE48>].

167. Sharfstein et al., *supra* note 165.

168. Anderson et al., *supra* note 72, at 944-45.

169. Sharfstein et al., *supra* note 165.

cause the locus of the disease is often the mind, and therefore, more subjective, non-biological evidence of efficacy, often collected through administration of psychometric scales that results in a PRO, may be the primary endpoint of a clinical trial. It is thus especially important to be transparent about how PROs and PPI will be evaluated for interventions targeting neuropsychiatric conditions.

IV. CONCLUSION

The FDA privileges the most objective evidence (i.e., biomarkers) of efficacy when determining whether a new drug or device should be approved. But other types of evidence—clinical assessments and patient reports—are also valuable. Indeed, relying on multiple types of evidence is especially important for interventions targeted to treat neuropsychiatric conditions given the lack of established biomarkers or the burden of obtaining biomarker evidence. We have thus argued for the collection and review of many types of evidence of an intervention's effectiveness in order to ensure that the development, approval, and dissemination of neurotechnology to patients is not hampered by the FDA's preference for biomarker-based evidence.

We argue that there is scientific value in assessing multiple types of evidence. There are also ethical reasons, such as respect for personhood, to collect patient feedback. Submission of more subjective types of evidence of efficacy, particularly in the form of PROs, should therefore be a required part of an FDA application when PROs are possible to collect. But if such evidence is submitted to the FDA, it is necessary to know in advance how it will be evaluated, and so we join others in calling for greater transparency about how the FDA will weigh such evidence. While some may be skeptical of the value of PROs, our recommendations are meant to promote sound science and regulatory policy. We support the approval and marketing of only safe and effective interventions, and we believe our recommendations will help accomplish this.

Finally, while we have primarily focused on neurotechnological devices in this Article, our arguments are relevant to all medical interventions for neuropsychiatric conditions, which currently lack validated biomarkers and for which clinical judgment and patient provided information can provide important insights into efficacy.