

## **B. Resources Within State Insurance Departments**

Many of these resources, such as a state insurance department consumer complaint resolution unit, are discussed in detail in the body of this handbook. Other key resources include:

### **Market Conduct and Financial Examinations**

Market conduct examinations focus on such areas as operations/management, complaint handling, marketing and sales, producer licensing, policyholder service, underwriting and rating, and claims. The financial condition examination system focuses on financial and corporate matters. Market conduct compliance issues can have a significant effect on legal and compliance risks, which in turn can create material solvency issues. Coordination with the financial examination function is an important area for market conduct examiners to understand. Guidance on financial condition examinations is provided in the *Financial Condition Examiner's Handbook* and is available through the Insurance Products and Services Division of the NAIC.

### **Financial Analysis**

Financial reporting and analysis information is shared with the NAIC, which assembles a wide range of data compilations on a multistate basis. An insurance department's financial analysis and examination staff can provide valuable assistance in interpreting this information. [Additionally, market regulators are encouraged to coordinate with a company's domestic financial regulator to obtain information related to the company's group capital calculations, liquidity stress test results, corporate governance, and Own Risk and Solvency Assessment \(ORSA\).](#)

### **Rates and Forms Information**

Tools such as the System for Electronic Rate and Form Filing (SERFF) and the insurance department posting of state filing review requirements provide a wide range of new data in formats that are more readily comparable across state and regional lines. As of April 2021, 53 jurisdictions including the District of Columbia, Puerto Rico, Guam and the Virgin Islands – plus more than 6,500 insurance companies, third-party filers, rating organizations and other companies—are using SERFF to efficiently and effectively speed insurance products to the market. The SERFF system provides an indicator of marketplace trends, such as overall increases in premiums or changes in coverages by the submission of filing of amendatory endorsements and exclusions.

### **Organized Intra-Department Communication**

State insurance departments are organized differently, but all perform a range of market regulation functions, from consumer assistance to producer licensing, from rate and form review to market conduct exams, and from investigations to enforcement. All of these functions, as well as financial regulation functions, generate useful information about market problems. An effective market analysis program must include clear procedures for regularly sharing data and other information among the various divisions of an insurance department. Recommended methods of sharing internal information include holding a monthly update meeting or emailing issues that may be of concern or interest to other sections.

**STANDARDS  
MARKETING AND SALES**

**Standard 1**

**All advertising and sales materials are in compliance with applicable statutes, rules and regulations.**

**Apply to:** All regulated entities

**Priority:** Essential

**Documents to be Reviewed**

\_\_\_\_\_ Applicable statutes, rules and regulations (Note: Reference applicable Interstate Insurance Product Regulation Commission (IIPRC) uniform standards for products approved by the IIPRC)

\_\_\_\_\_ All regulated entity advertising and sales materials, including radio and audiovisual items such as television commercials, telemarketing scripts, pictorial materials, social media or other electronic medium

\_\_\_\_\_ Policy forms as they coincide with advertising and sales materials

\_\_\_\_\_ Producer's own advertising and sales materials

\_\_\_\_\_ Regulated entity policies and procedures

**Others Reviewed**

\_\_\_\_\_  
\_\_\_\_\_

**NAIC Model References**

*Unfair Trade Practices Act* (#880)

*Advertisements of Life Insurance and Annuities Model Regulation* (#570), Section 3B

*Risk-Based Capital (RBC) for Insurers Model Act* (#312), Section 8B

*Life Insurance Disclosure Model Regulation* (#580), Section 8C

*Life and Health Insurance Guaranty Association Model Act* (#520), Section 19A

*Long-Term Care Insurance Model Act* (#640)

*Life Insurance Illustrations Model Regulation* (#582)

*Small Employer and Individual Health Insurance Availability Model Act* (#35)

*Model Regulation to Implement the Individual Accident and Sickness Insurance Minimum Standards Model Act* (#171), Section 7(H)(1)(a)(I)

*Advertisements of Accident and Sickness Insurance Model Regulation* (#40)

*Individual Health Insurance Portability Model Act* (#37), Section 5

*Title Insurers Model Act* (#628)

*Title Insurance Agent Model Act* (#230)

*Home Service Disclosure Model Act* (#920)

*Marketing Insurance Over the Internet White Paper*

*Group Health Insurance Standards Model Act* (#100)

*Medicare Supplement Insurance Minimum Standards Model Act* (#650)

*Model Regulation to Implement the NAIC Medicare Supplement Insurance Minimum Standards Model Act* (#651)

*The Use of Social Media in Insurance White Paper*

[\*Insurance Holding Company System Regulatory Model Act\* \(#440\), Section 8G](#)

## **IIPRC Uniform Standard References**

*IIPRC Standards for Individual Long-Term Care Advertising Materials* (applicable to individual long-term care (LTC) products and associated advertising materials submitted and/or approved by the IIPRC)

### **Review Procedures and Criteria**

Review advertising materials in conjunction with the appropriate policy form. If statistics are included, proper citation should be included in the documentation.

Materials should not:

- Misrepresent the dividends or share of the surplus to be received on any policy;
- Make a false or misleading statement as to the dividends or share of the surplus previously paid on the policy;
- Misrepresent any policy as being shares of stock;
- Misrepresent policy benefits forms or conditions by failing to disclose limitations, exclusions or reductions or use terms or expressions that are misleading or ambiguous;
- Make unfair or incomplete comparisons with other policies;
- Make false, deceptive or misleading statements or representations with respect to any person, regulated entity or organization in the conduct of insurance business; and
- Offer unlawful rebates or inducements.

Materials should:

- Disclose the name and address of insurer;
- Comply with applicable statutes, rules and regulations; and
- Cite the source of statistics used by the regulated entity.

Determine if the regulated entity approves producer sales materials and advertising. Determine if advertisements or lead-generating calls falsely project the image that they were sent by a government agency.

Review the regulated entity's and producer's websites with the following questions in mind:

- Does the website disclose who is selling/advertising/servicing for the website?
- Does the website disclose what is being sold or advertised?
- If required by statutes, rules or regulations, does the website reveal the physical location of the regulated entity/entities?
- Does the website reveal the jurisdictions where the advertised product is (or is not) approved, or use some other mechanism (including, but not limited to, identifying persons by geographic location) to accomplish an appropriate result?

For the review of Internet advertisements:

- Run an inquiry with the regulated entity's name;
- Review the regulated entity's home page;
- Identify all lines of business referenced on the regulated entity's home page;
- Research the ability to request more information about a particular product and verify the information provided is accurate; and
- Review the regulated entity's procedures related to producers advertising on the Internet and ensure the regulated entity requires prior approval of the producer pages, if the regulated entity name is used.

For the review of social media:

- Perform a search of social media sites with the regulated entity's name;
- Identify social media sites in which the regulated entity is active;

- Review identified social media sites and verify any product information provided by the regulated entity is accurate;
- Review the regulated entity’s policies and procedures to identify the personnel involved in monitoring the regulated entity’s marketing and sales-related social media activity;
- Review the regulated entity’s policies and procedures for tracking marketing and sales-related social media requiring regulated entity review; and
- If the regulated entity requires preapproval of producer advertising on the Internet, review the regulated entity’s preapproval procedures to determine whether the regulated entity identifies marketing and sales-related social media as also requiring regulated entity preapproval.

**Automation Tip:**

Enter a summary of all marketing materials of whatever description in an Excel spreadsheet. Capture the regulated entity’s name of the material; the form number, if any; the edition date, if any; source, if applicable; and media, such as Internet or direct mail. Include fields to note exceptions, such as unsupported statistics or possible misleading statements. Insert each possible violation/exception in a separate field.

Statistics and statements are likely to be repeated in more than one “piece” of marketing material. It is also possible that one piece of marketing material will contain more than one violation/exception.

The spreadsheet will make it easier to track any repeated statements and to identify any marketing material containing apparent multiple violations/exceptions.

**STANDARDS  
MARKETING AND SALES**

**Standard 1**

**The regulated entity’s mass marketing of property/casualty insurance is in compliance with applicable statutes, rules and regulations.**

**Apply to:** All regulated entities

**Priority:** Recommended

**Documents to be Reviewed**

\_\_\_\_\_ Applicable statutes, rules and regulations

\_\_\_\_\_ New business policy forms [and certificate of insurance \(such certificates will only be requested for lender-placed insurance policies\)](#)

\_\_\_\_\_ Advertising materials

\_\_\_\_\_ Disclosure materials

\_\_\_\_\_ Marketing complaints

\_\_\_\_\_ Underwriting guidelines

Others Reviewed

\_\_\_\_\_  
\_\_\_\_\_

**NAIC Model References**

*Mass Marketing of Property and Liability Insurance Model Regulation (#710)*

*Group Personal Lines Property and Casualty Insurance Model Act (#760)*

[Real Property Lender Placed Model Act \(#631\), Sections 5, 8 and 9](#)

**Review Procedures and Criteria**

Review documentation in new business policy files to determine a legitimate basis for the group. If not evident from the file, request additional documentation from the regulated entity to verify that the group is not fictitious.

Review underwriting guidelines, new business policy files, advertising materials, disclosure materials and complaints to verify:

- Compulsory participation not required for employment or group membership;
- Tie-in sales are not a condition of purchase; and
- Disclosures are provided, as required.

**STANDARDS  
UNDERWRITING AND RATING**

**Standard 4  
Verification of premium audit accuracy and the proper application of rating factors.**

**Apply to:** All regulated entities

**Priority:** Essential

**Documents to be Reviewed**

\_\_\_\_\_ Applicable statutes, rules and regulations

\_\_\_\_\_ Insurance department approved and/or filed rating plans, including risk modification plans

\_\_\_\_\_ Copies of cost containment certificates and loss improvement criteria to determine cost containment discount

\_\_\_\_\_ Final rate manual tables by classification codes applicable to the period under examination (tables maintained at the regulated entity level)

\_\_\_\_\_ Workers' Compensation Experience Modification Rating Sheets pertaining to the policy sample (experience modifiers as published by the NCCI and similar advisory organizations)

[For lender placed insurance, documentation showing regulated entity's separate rates for mortgage servicer obtained lender-placed insurance versus voluntary insurance on real estate owned property](#)

**Others Reviewed**

\_\_\_\_\_  
\_\_\_\_\_

**NAIC Model References**

*Property and Casualty Model Rating Guideline (File and Use Version) (#1775)*

*Property and Casualty Model Rating Guideline (Prior Approval Version) (#1780)*

*Property and Casualty Commercial Rate and Policy Form Model Law (Condensed) (#777)*

*Guidelines for the Filing of Workers' Compensation Large Deductible Policies and Programs (#1970)*

*Guidelines for Regulations and Legislation on Workers' Compensation Coverage for Professional Employer Organization Arrangements (#1950)*

[Real Property Lender Placed Model Act \(#631\), Section 9](#)

**Review Procedures and Criteria**

The purpose of this review is to determine that the final premium charged to the employer is being applied correctly, fairly and consistently.

The sample's premium audits should contain specific information on each policy. The sample's information should be compared to the NCCI unit statistical report and to the company's rating plan, to verify accuracy in the application and reporting of the following factors when applicable:

- Premiums by classification code;
- Payroll exposure;

- Schedule rating;
- Cost containment discount;
- Premium discounting;
- Designated medical provider discount;
- Expense loading;
- Application of the correct experience modifier;
- Small employer discount;
- Discount for rehiring previously disabled employees; and
- Any other rating elements.

The company documents should be reviewed. Any additional areas or lack of information should be discussed with company management.

It is recommended that separate samples be obtained for standard voluntary market, assigned risk (if applicable), large deductible (if applicable) and PEO accounts.

### STANDARDS UNDERWRITING AND RATING

<b>Standard 6</b> <b>Verification of loss reporting.</b>
---

**Apply to:** All workers' compensation examinations [and lender-placed insurance examinations, as applicable](#)

**Priority:** Essential

#### Documents to be Reviewed

- \_\_\_\_\_ Applicable statutes, rules and regulations
- \_\_\_\_\_ NCCI (and similar advisory organizations') rules governing the reporting of losses on unit statistical reports
- \_\_\_\_\_ Loss data pertaining to the policy sample and maintained by the regulated entity
- \_\_\_\_\_ Unit statistical reports pertaining to the policy sample and used to report regulated entity information to the NCCI (and similar advisory organizations)
- \_\_\_\_\_ [Applicable reports filed with the commissioner \(e.g., required reporting for insurers with at least \\$100,000 in direct written premium for lender-placed insurance, and required rate filing for insurers with an annual loss ratio of less than 35% in any lender-placed program, except with respect to lender-placed flood insurance, for two consecutive years\)](#)

#### Others Reviewed

- \_\_\_\_\_
- \_\_\_\_\_

#### NAIC Model References

- Property and Casualty Model Rating Guideline (File and Use Version) (#1775)*
- Property and Casualty Model Rating Guideline (Prior Approval Version) (#1780)*
- Property and Casualty Commercial Rate and Policy Form Model Law (Condensed) (#777)*

*Guidelines for the Filing of Workers' Compensation Large Deductible Policies and Programs (#1970)*  
*Guidelines for Regulations and Legislation on Workers' Compensation Coverage for Professional Employer Organization Arrangements (#1950)*  
[Real Property Lender Placed Model Act \(#631\), Section 9](#)

### Review Procedures and Criteria

Losses under each policy should be clearly and accurately maintained at the regulated entity, so that paid amounts, reserves, ~~and~~ deductibles [and, with respect to losses under lender-placed insurance policies, any excess amounts paid to the mortgagor](#) can be easily reviewed. The sample data should be compared to the unit statistical reports to verify accuracy of reporting of the following items:

- Paid losses;
- Paid loss adjustment expenses;
- Net of deductible reporting on the unit statistical reports;
- Adjustments to reserves and revised unit statistical reports; and
- Any other adjustments, such as subrogation.

The regulated entity's documents should be reviewed. Any additional areas or lack of information should be discussed with the regulated entity's management.

## STANDARDS UNDERWRITING AND RATING

<b>Standard 8</b> <b>Underwriting, rating and classification are based on adequate information developed at or near inception of the coverage rather than near expiration, or following a claim.</b>
---

**Apply to:** All regulated entities

**Priority:** Essential

### Documents to be Reviewed

\_\_\_\_\_ Applicable statutes, rules and regulations

\_\_\_\_\_ Application

\_\_\_\_\_ Underwriting files

Others Reviewed

\_\_\_\_\_  
\_\_\_\_\_

### NAIC Model References

*Unfair Trade Practices Act (#880)*  
[Real Property Lender Placed Model Act \(#631\), Section 4](#)

### Review Procedures and Criteria

Decisions should be based on information that reasonably should have been developed at the inception of the policy or during initial underwriting and not, through audit or other means, [before the policy went into effect or](#) after the policy has expired.

Determine if the initial underwriting of a policy is based on the information obtained after a claim is submitted.

**STANDARDS  
UNDERWRITING AND RATING**

**Standard 13**

**The regulated entity does not engage in collusive or anti-competitive underwriting practices.**

**Apply to:** All regulated entities

**Priority:** Essential

**Documents to be Reviewed**

\_\_\_\_\_ Applicable statutes, rules and regulations

\_\_\_\_\_ Underwriting files

[For lender-placed insurers, books and records containing compensation, contingent commissions, profit sharing and other payments dependent on profitability or loss ratios](#)

[For lender-placed insurers, third party agreements for outsourced services](#)

Others Reviewed

\_\_\_\_\_

\_\_\_\_\_

**NAIC Model References**

*Unfair Trade Practices Act (#880)*

[Real Property Lender Placed Model Act \(#631\), Section 6](#)

**Review Procedures and Criteria**

Any practice suggesting anti-competitive behavior may involve legal considerations that should be referred to insurance department legal counsel. This would include engaging in collusive underwriting practices that may inhibit competition; e.g., entering into an agreement with other companies to divide the auto market within the jurisdiction by territory.

The examiner should be aware of unlawful pricing and other prohibited anti-competitive acts or practices.

## Chapter 24B—Conducting the Mental Health Parity and Addiction Equity Act (MHPAEA) Related Examination

### IMPORTANT NOTE:

The standards set forth in this chapter are based on established procedures and/or NAIC models, not on the laws and regulations of any specific jurisdiction. This handbook is a guide to assist examiners in the examination process. Since it is based on NAIC models, use of the handbook should be adapted to reflect each state's own laws and regulations with appropriate consideration for any bulletins, audit procedures, examination scope and the priorities of examination. Further important information on this and how to use this handbook is included in the Foreward section of the handbook.

### Introduction

The purpose of this chapter, Conducting the Mental Health Parity and Addiction Equity Act (MHPAEA) Related Examination, is to provide guidance for examiners when reviewing insurers whose business includes major medical policies offering mental health and/or substance use disorder coverage.

Federal law relies on state insurance regulators as the first-line enforcers of health reform provisions in the individual, small group and large group insurance markets. The examination standards in Chapter 24—Conducting the Health Examination of the *Market Regulation Handbook* provide guidance specific to all health carriers that may or may not include offering mental health and/or substance use disorder coverage. Chapter 24, Section G Claims, Standard 3 applies to examinations related to the Mental Health Parity and Addiction Equity Act (MHPAEA) of 2008 found at 42 U.S.C. § 300gg-26.

The guidance found in this chapter recognizes that when developing an examination or review plan related to MHPAEA compliance, it is important to consider examination standards as applicable from Chapter 24 and Chapter 24A—Conducting the Affordable Care Act (ACA) Related Examination, as well as Chapter 20.

Regardless of which chapter is used in the *Market Regulation Handbook*, the examiner will also need to reference Chapter 20—General Examination Standards for general examination standards that apply to all insurers.

## Mental Health and Substance Use Disorder Parity

### 1. Purpose

Mental health and substance use disorder parity compliance examinations should be designed to ensure that all health carriers are in compliance with all the requirements of the Mental Health Parity and Addiction Equity Act (MHPAEA) of 2008 (as amended by the Consolidated Appropriations Act of 2021) found at 42 U.S.C. § 300gg-26 and its implementing regulations found at 45 CFR § 146.136 and 45 CFR § 147.160.

These standards require health carriers to demonstrate compliance in terms of defining mental health or substance use disorder benefits, classifying benefits, financial requirements, quantitative treatment limitations (QTLs), nonquantitative treatment limitations (NQTLs), required disclosures and vendor coordination.

### 2. Definitions

For purposes of this Guide, except where the context clearly indicates otherwise, the following terms have the meanings indicated:

***Aggregate Lifetime Dollar Limit*** means a dollar limitation on the total amount of specified benefits that may be paid under a health plan (45 CFR § 146.136(a)).

**Annual Dollar Limit** means a dollar limitation on the total amount of specified benefits that may be paid in a 12-month period under a health plan (45 CFR § 146.136(a)).

**Classifications of benefits used for applying parity rules:**

**(1) Inpatient, In-network.** Benefits furnished on an inpatient basis and within a network of providers established or recognized under a plan or health insurance coverage (45 CFR § 146.136(c)(2)(ii)(A)(1)). See special rule for plans with multiple network tiers in paragraph (c)(3)(iii)(B) of 45 CFR § 146.136.

- a. If a plan provides benefits through multiple tiers of in-network providers (such as in-network preferred and in-network participating providers), the plan may divide its benefits furnished on an in-network basis into subclassifications that reflect network tiers, if the tiering is based on reasonable factors determined in accordance with the rules for NQTLs (such as quality performance, and market standards) and without regard to whether a provider provides services with respect to medical/surgical benefits for MH/SUD benefits. After the tiers are established, the plan may not impose any financial requirement or treatment limitation on MH/SUD benefits in any tier that is more restrictive than the predominant financial requirement or treatment limitation that applies to substantially all medical/surgical benefits in the tier.

**(2) Inpatient, Out-of-network.** Benefits furnished on an inpatient basis and outside any network of providers established or recognized under a plan or health insurance coverage. This classification includes inpatient benefits under a plan (or health insurance coverage) that has no network of providers (45 CFR § 146.136(c)(2)(ii)(A)(2)).

**(3) Outpatient, In-network.** Benefits furnished on an outpatient basis and within a network of providers established or recognized under a plan or health insurance coverage. See special rules for office visits and plans with multiple network tiers in paragraph (c)(3)(iii)(C) and (c)(3)(iii)(B) of 45 CFR § 146.136.

- a. A plan may divide its benefits furnished on an outpatient basis into two subclassifications: (1) office visits; and (2) all other outpatient items and services, for purposes of applying the financial requirement and treatment limitation rules.
- b. If a plan provides benefits through multiple tiers of in-network providers (such as in-network preferred and in-network participating providers), the plan may divide its benefits furnished on an in-network basis into subclassification the reflect network tiers, if the tiering is based on reasonable factors determined in accordance with the rules for NQTLs (such as quality, performance, and market standards) and without regard to whether a provider provides services with respect to medical/surgical benefits or MH/SUD benefits. After the tiers are established the plan may not impose any financial requirements or treatment limitation on MH/SUD benefits in any tier that is more restrictive than the predominant financial requirement or treatment limitation that applies to substantially all medical/surgical benefits in the tier.

**(4) Outpatient, Out-of-network.** Benefits furnished on an outpatient basis and outside any network of providers established or recognized under a plan or. This classification includes outpatient benefits under a plan (or health insurance coverage) that has no network of providers (45 CFR § 146.136(c)(2)(ii)(A)(4)). See special rule for office visits in paragraph (c)(3)(iii)(C) of 45 CFR § 146.136.

- a. A plan may divide its benefits furnished on an outpatient basis into two subclassifications: (1) office visits; and (2) all other outpatient items and services, for purposes of applying the financial requirement and treatment limitation rules.

**(5) Emergency Care.** Benefits for emergency care (45 CFR § 146.136(c)(2)(ii)(A)(5)).

**(6) Prescription Drugs.** Benefits for prescription drugs (45 CFR § 146.136(c)(2)(ii)(A)(6)).

**Coverage Unit** refers to the way in which a plan (or health insurance coverage) groups individuals for purposes of determining benefits, or premiums or contributions. For example, different Coverage Units include self-only, family, and employee plus-spouse (45 CFR § 146.136(a)).

**Cumulative Financial Requirements** are financial requirements that determine whether or to what extent benefits are provided based on accumulated amounts and include deductibles and out-of-pocket maximums. (However, cumulative financial requirements do not include aggregate lifetime or annual dollar limits because these two terms are excluded from the meaning of financial requirements.) (45 CFR § 146.136(a))

**Cumulative Quantitative Treatment Limitations** are treatment limitations that determine whether or to what extent benefits are provided based on accumulated amounts, such as annual or lifetime day or visit limits (45 CFR § 146.136(a)).

**Expected Plan Payments** are payments expected to be paid under the plan for the plan year (45 CFR § 146.136(c)(3)(i)(C)). Any reasonable method may be used to determine the dollar amount expected to be paid under the plan for medical/surgical benefits subject to a financial requirement or QTL (45 CFR § 146.136(c)(3)(i)(E)).

**Plan Payment** is the dollar amount of plan payments and is based on the amount the plan allows (before enrollee cost sharing) rather than the amount the plan pays (after enrollee cost sharing) because payment based on the allowed amount covers the full scope of the benefits being provided (45 CFR § 146.136(c)(3)(i)(D)).

**Financial Requirements** include deductibles, copayments, coinsurance, or out-of-pocket maximums. Financial requirements do not include aggregate lifetime or annual dollar limits (45 CFR § 146.136(a)).

**Medical/Surgical Benefits** means benefits with respect to items or services for medical conditions or surgical procedures, as defined under the terms of the plan or health insurance coverage and in accordance with applicable Federal and State law, but does not include mental health or substance use disorder benefits. Any condition defined by the plan or coverage as being or as not being a medical/surgical condition must be defined to be consistent with generally recognized independent standards of current medical practice (for example, the most current version of the International Classification of Diseases (ICD) or State guidelines) (45 CFR § 146.136(a)).

**Mental Health Benefits** means benefits with respect to items or services for mental health conditions, as defined under the terms of the plan or health insurance coverage and in accordance with applicable Federal and State law. Any condition defined by the plan or coverage as being or as not being a mental health condition must be defined to be consistent with generally recognized independent standards of current medical practice (for example, the most current version of the Diagnostic and Statistical Manual of Mental Disorders (DSM), the most current version of the ICD, or State guidelines) (45 CFR § 146.136(a)).

**Substance Use Disorder Benefits** means benefits with respect to items or services for substance use disorders, as defined under the terms of the plan or health insurance coverage and in accordance with applicable Federal and State law. Any disorder defined by the plan as being or as not being a substance use disorder must be defined to be consistent with generally recognized independent standards of current medical practice (for example, the most current version of the DSM, the most current version of the ICD, or State guidelines) (45 CFR § 146.136(a)).

**Treatment Limitations** include limits on benefits based on the frequency of treatment, number of visits, days of coverage, days in a waiting period, or other similar limits on the scope or duration of treatment. Treatment limitations include both quantitative treatment limitations (QTLs), which are expressed numerically (such as 50 outpatient visits per year), and nonquantitative treatment limitations (NQTLs), which are not expressed numerically but otherwise limit the scope or duration of benefits for treatment

under a plan or coverage. A permanent exclusion of all benefits for a particular condition or disorder, however, is not a treatment limitation for purposes of this definition (45 CFR § 146.136(a)).

### **3. Techniques**

To evaluate compliance with MHPAEA, examiners must request that the carrier submit the analyses and other underlying documentation that it has performed to determine that it meets all of the standards of MHPAEA. There must be specific documentation of how mental health conditions, substance use disorders and medical/surgical conditions were defined and how they were assigned to benefit classifications. There are specific mathematical analyses that the carrier must have performed in order to determine that it satisfies the MHPAEA requirements for financial requirements and quantitative treatment limitations QTLs. There are separate analyses the carrier must have performed in order to determine that it satisfies the MHPAEA requirement for NQTLs, which entail analyses for the “as written” component and analyses for the “in operation” component.

### **4. Standards and the Regulatory Tests**

The mental health and substance use disorder parity review includes, but is not limited to, the following standards related to MHPAEA. The sequence of the standards listed here does not indicate priority of the standard.

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 1**

**The health carrier shall define all covered services as mental health or substance use disorder benefits or as medical or surgical benefits. Mental health benefits or substance use disorder benefits must be defined to mean items or services for the treatment of a mental health condition or substance use disorder, as defined by the terms of the health plan and applicable state and federal law. Any definition of a condition or disorder as being or as not being a mental health condition or substance use disorder must be consistent with generally recognized independent standards of current medical practice or state guideline. (45 CFR § 146.136(a)).**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

- \_\_\_\_\_ Applicable state and federal statutes, rules, regulations and published sub-regulatory guidance
- \_\_\_\_\_ Type of generally recognized independent standards of current medical practice, state law or guidance, used to define mental health conditions, substance use disorders and medical/surgical conditions (e.g., the most current version of the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Statistical Classification of Diseases and Related Health Problems (ICD code), etc.)
- \_\_\_\_\_ List of specific mental health conditions or substance use disorders by diagnosis excluded from coverage as stated in the policy documents
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical claim files
- \_\_\_\_\_ Health carrier complaint/grievances/appeals records concerning mental health and/or substance use disorders (supporting documentation, including, but not limited to: written and phone records of inquiries, call center scripts, complaints, complainant correspondence and health carrier response)
- \_\_\_\_\_ Internal department appeals/grievance files
- \_\_\_\_\_ Applicable external appeals register/logs/files, external appeal resolution and associated documentation

**Other References**

Enforcement of the Public Health Services Act  
42 U.S. Code § 300gg–22

Preemption relating to the Public Health Services Act  
42 U.S. Code § 300gg–23

Mental Health Parity and Addiction Equity Act of 2008  
42 U.S. Code § 300gg–26

Publication of summary plan description  
ERISA 104(b) (29 U.S.C. § 1024(b))

The Federal Parity Self Compliance Tool: <https://www.dol.gov/sites/dolgov/files/EBSA/laws-and-regulations/laws/mental-health-parity/self-compliance-tool.pdf>

[U.S. Department of Labor Frequently Asked Questions Guidance: https://www.dol.gov/agencies/ebsa/laws-and-regulations/laws/mental-health-and-substance-use-disorder-parity](https://www.dol.gov/agencies/ebsa/laws-and-regulations/laws/mental-health-and-substance-use-disorder-parity)

### **Review Procedures and Criteria**

Review definitions in the health carrier's policy forms and/or certificates of coverage for compliance with the definitions in 45 CFR § 146.136(a) and included in the definitions section of this chapter.

Review the health carrier's description of the independent standards it used to define mental health conditions, substance use disorders and medical/surgical conditions. These independent standards must be generally recognized independent standards of current medical practice such as the Diagnostic and Statistical Manual (DSM) or the International Classification of Diseases (ICD), or state guidelines.

Review exclusions in the health carrier's policy forms and/or certificates of coverage to identify those that involve a mental health or substance use disorder condition or diagnosis and compare it to the list of mental health and substance use disorder conditions excluded from coverage provided by the health carrier.

Verify that exclusions in the health carrier's policy forms and/or certificates of coverage identified as not a mental health or substance use disorder condition comply with state law and are consistent with generally recognized independent standards such as the International Classification of Diseases (ICD) or the Diagnostic and Statistical Manual of Mental Disorders (DSM).

Review any attestations required by the state and submitted by the health carrier.

For services the health carrier has determined are both medical/surgical and mental health/substance use disorders, review the explanation of how they determine the correct expected dollar amount for these services (e.g., nutritional counseling, occupational therapy).

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 2**

**The health carrier must apply the same standards to medical/surgical benefits and to mental health or substance use disorder benefits in determining the classification in which a particular benefit belongs (or applicable sub-classification) (45 CFR § 146.136(c)(2)(ii)(A)).**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

- \_\_\_\_\_ Applicable state and federal statutes, rules, regulations and published sub-regulatory guidance
- \_\_\_\_\_ All policy documents (e.g., if group or association, request master policy and a sample of each certificate type issued during the examination scope)
- \_\_\_\_\_ Documentation as to how the carrier demonstrates assignment to the six classifications of benefits (and applicable sub-classifications) and the standard used
- \_\_\_\_\_ Company and vendor claim procedure manuals and bulletins/communications (if a carrier uses a behavioral health claims vendor for processing MH/SUD claims or for providing utilization management services)
- \_\_\_\_\_ Internal company claim audit reports for both mental health or substance use disorders and medical/surgical services
- \_\_\_\_\_ Provider contracts, instructions, communications and similar documents regarding coding instructions, code changes, etc.
- \_\_\_\_\_ Utilization review and managed care guidelines and procedure manuals
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical claim files
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical complaint and grievance files

**Other References**

Enforcement of the Public Health Services Act  
42 U.S. Code § 300gg–22

Preemption relating to the Public Health Services Act  
42 U.S. Code § 300gg–23

Mental Health Parity and Addiction Equity Act of 2008  
42 U.S. Code § 300gg–26

Publication of summary plan description  
ERISA 104(b) (29 U.S.C. § 1024(b))

The Federal Parity Self Compliance Tool: <https://www.dol.gov/sites/dolgov/files/EBSA/laws-and-regulations/laws/mental-health-parity/self-compliance-tool.pdf>

[U.S. Department of Labor Frequently Asked Questions Guidance: https://www.dol.gov/agencies/ebsa/laws-and-regulations/laws/mental-health-and-substance-use-disorder-parity](https://www.dol.gov/agencies/ebsa/laws-and-regulations/laws/mental-health-and-substance-use-disorder-parity)

### **Review Procedures and Criteria**

Review the health carrier's list that specified the classification or sub-classification to which each benefit was assigned.

Determine whether the health carrier uses permissible sub-classifications for any benefits.

Please note that the only permissible sub-classifications are: multiple tiers for prescription drugs benefits that are based on reasonable factors<sup>1</sup> (45 CFR § 146.136(c)(3)(iii)(A)); multiple network tiers that are based on reasonable factors within the inpatient in-network and outpatient in-network classifications (45 CFR § 146.136(c)(3)(iii)(B)); outpatient office visits and outpatient other services within the outpatient in-network and outpatient out-of-network classifications (45 CFR § 146.136(c)(3)(iii)(C)). Reasonable factors include cost, efficacy, generic versus brand name, and mail order versus pharmacy pick-up (45 CFR § 146.136(c)(3)(iii)(A)).

Review the standard used by the health carrier to determine which classification of benefits (or applicable sub-classification) a particular benefit was assigned to and verify that the same standards were used for assigning medical/surgical benefits and mental health or substance use disorder benefits.

Review the health carrier's documentation that demonstrates that mental health or substance use disorder benefits are covered in each classification in which medical/surgical benefits are covered.

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 3**

**The health carrier shall not apply any financial requirement on mental health or substance use disorder benefits in any classification (or applicable sub-classification) that is more restrictive than the predominant financial requirement of that type applied to substantially all medical/surgical benefits in the same classification (or applicable sub-classification) (45 CFR § 146.136(c)(2)(i)).**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

- \_\_\_\_\_ Applicable state and federal statutes, rules, regulations and published sub-regulatory guidance
- \_\_\_\_\_ Health carrier list of all financial requirements applied to mental health or substance use disorder benefits and medical/surgical benefits in each classification (or applicable sub-classification) (this will include schedules of benefits and other policy documents)
- \_\_\_\_\_ Health carrier documentation of the reasonable method used to determine expected plan payments for medical/surgical benefits within each classification of benefits (or applicable sub-classification), including documentation and communications with vendors engaged to provide assistance with analyses
- \_\_\_\_\_ Documentation demonstrating the predominant and substantially tests performed by the health carrier for each applicable financial requirement applied to all benefits
- \_\_\_\_\_ Internal company claim audit reports specific to mental health or substance use disorders
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical claim files
- \_\_\_\_\_ Health carrier complaint records concerning mental health and/or substance use disorder (supporting documentation, including, but not limited to: written and phone records of inquiries, complaints, complainant correspondence and health carrier response)
- \_\_\_\_\_ Internal department appeals/grievance files concerning mental health and/or substance use disorders
- \_\_\_\_\_ Applicable external appeals register/logs/files related to concerning mental health and/or substance use disorder, external appeal resolution and associated documentation

**Other References**

Enforcement of the Public Health Services Act  
42 U.S. Code § 300gg–22

Preemption relating to the Public Health Services Act  
42 U.S. Code § 300gg–23

Mental Health Parity and Addiction Equity Act of 2008  
42 U.S. Code § 300gg–26

Publication of summary plan description  
ERISA 104(b) (29 U.S.C. § 1024(b))

The Federal Parity Self Compliance Tool: <https://www.dol.gov/sites/dolgov/files/EBSA/laws-and-regulations/laws/mental-health-parity/self-compliance-tool.pdf>

## **Review Procedures and Criteria**

Financial requirements include deductibles, copayments, coinsurance and out-of-pocket maximums (45 CFR § 146.136(c)(1)(ii)). A financial requirement is considered to apply to substantially all medical/surgical benefits in a classification of benefits if it applies to at least two-thirds of all medical/ surgical benefits in that classification (45 CFR § 146.136(c)(3)(i)(A)). The level of the financial requirement that is considered the predominant level of that type in a classification of benefits is the level that applies to more than one-half of medical/surgical benefits in that classification subject to the financial requirement (45 CFR § 146.136(c)(3)(i)(B)). The determination of the portion of medical/ surgical benefits in a classification of benefits subject to a financial requirement (or subject to any level of a financial requirement) is based on the dollar amount of all plan payments for medical/surgical benefits in the classification expected to be paid under the plan for the plan year (or for the portion of the plan year after a change in plan benefits that affects the applicability of the financial requirement) (45 CFR § 146.136(c)(3)(i)(C)).

Review the health carrier's methodology for performing its analysis that determines expected plan payments within each classification of benefits (or applicable sub-classification) for medical/surgical benefits. Note: A health carrier must always use appropriate and sufficient data to perform the analysis in compliance with applicable Actuarial Standards of Practice (ACA FAQ 34 Q3).

Review the health carrier's documentation that demonstrates that any type of financial requirement applied to mental health or substance use disorder benefits in a classification (or applicable sub-classification) applies to at least two-thirds of expected plan payments on medical/surgical benefits within that classification (or applicable sub-classification) (45 CFR § 146.136(c)(3)(i)(A)). Note: If the financial requirement applies to all medical/surgical benefits in the classification, no cost analysis is required. No financial requirements shall apply only to mental health or substance use disorder benefits.

Determine whether the health carrier's documentation supports that the level of financial requirement applied to mental health or substance use disorder benefits in a classification (or applicable sub-classification) is comparable and no more restrictive than the level of financial requirement that applies to more than one-half of expected plan payments that are subject to the financial requirement within that classification for medical/surgical benefits (45 CFR § 146.136(c)(3)(i)(B)(1)).

If no single level of the financial requirement applies to more than one-half of medical/surgical benefits in the classification, determine whether the health carrier can demonstrate that it has satisfied this test (45 CFR § 146.136(c)(3)(i)(B)(2)).

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 4**

**The health carrier shall not apply any QTL on mental health or substance use disorder benefits in any classification (or applicable sub-classification) that is more restrictive than the predominant QTL of that type applied to substantially all medical/surgical benefits in the same classification (or applicable sub-classification) (45 CFR § 146.136(c)(2)(i)).**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

\_\_\_\_\_ Applicable state and federal statutes, rules, regulations and published sub-regulatory guidance

\_\_\_\_\_ Health carrier list of all QTLs applied to mental health or substance use disorder benefits and medical/surgical benefits in each classification (or applicable sub-classification) (this will include schedules of benefits and other policy documents)

\_\_\_\_\_ Health carrier documentation of the reasonable method used to determine expected plan payments for medical/surgical benefits within each classification of benefits (or applicable sub-classification), including documentation and communications with vendors engaged to provide assistance with analyses

\_\_\_\_\_ Documentation demonstrating the predominant and substantially tests performed by the health carrier for each applicable financial requirement applied to all benefits

\_\_\_\_\_ Internal company claim audit reports

\_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical claim files

\_\_\_\_\_ Health carrier complaint, grievance and appeals records (supporting documentation, including, but not limited to: written and phone records of inquiries, complaints, call center scripts, complainant correspondence and health carrier response)

**Other References**

Enforcement of the Public Health Services Act  
42 U.S. Code § 300gg–22

Preemption relating to the Public Health Services Act  
42 U.S. Code § 300gg–23

Mental Health Parity and Addiction Equity Act of 2008  
42 U.S. Code § 300gg–26

Publication of summary plan description  
ERISA 104(b) (29 U.S.C. § 1024(b))

The Federal Parity Self Compliance Tool: <https://www.dol.gov/sites/dolgov/files/EBSA/laws-and-regulations/laws/mental-health-parity/self-compliance-tool.pdf>

## Review Procedures and Criteria

QTLs include annual, episode, and lifetime day and visit limits. (45 CFR § 146.136(c)(1)(ii)). A QTL is considered to apply to substantially all medical/surgical benefits in a classification of benefits if it applies to at least two-thirds of all medical/ surgical benefits in that classification (45 CFR § 146.136(c)(3)(i)(A)). The level of the QTL that is considered the predominant level of that type in a classification of benefits is the level that applies to more than one-half of medical/surgical benefits in that classification subject to the QTL (45 CFR § 146.136(c)(3)(i)(B)). The determination of the portion of medical/ surgical benefits in a classification of benefits subject to a quantitative treatment limitation (or subject to any level of a quantitative treatment limitation) is based on the dollar amount of all plan payments for medical/surgical benefits in the classification expected to be paid under the plan for the plan year (or for the portion of the plan year after a change in plan benefits that affects the applicability of the quantitative treatment limitation) (45 CFR § 146.136(c)(3)(i)(C)).

Review the health carrier’s methodology for performing its analysis that determines expected plan payments within each classification of benefits (or applicable sub-classification) for medical/surgical benefits. Note: A health carrier must always use appropriate and sufficient data to perform the analysis in compliance with applicable Actuarial Standards of Practice (ACA FAQ 34 Q3).

Review the health carrier’s documentation that demonstrates that any type of QTL applied to mental health or substance use disorder benefits in a classification (or applicable sub-classification) applies to at least two-thirds of expected plan payments on medical/surgical benefits within that classification (or applicable sub-classification) (45 CFR § 146.136(c)(3)(i)(A)). Note: If the quantitative limitation applies to all medical/surgical benefits within the classification, no cost analysis is required. No quantitative treatment limitations shall apply only to mental health or substance use disorder benefits.

Determine whether the health carrier’s documentation supports that the level of QTL applied to mental health or substance use disorder benefits in a classification (or applicable sub-classification) is no more restrictive than the level of QTL that applies to more than one-half of expected plan payments that are subject to the quantitative treatment limitation within that classification for medical/surgical benefits (45 CFR § 146.136(c)(3)(i)(B)(1)).

If no single level applies to more than one-half of medical/surgical benefits in the classification, determine whether the health carrier can demonstrate that it has satisfied this test (45 CFR § 146.136(c)(3)(i)(B)(2)).

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 5**

**The health carrier shall apply non-quantitative treatment limitations (NQTLs) to mental health or substance use disorder benefits within a classification of benefits (or applicable sub-classification) so that any processes, strategies, evidentiary standards, or other factors used to apply a limitation, 1) as written and 2) in operation, are comparable to the processes, strategies, evidentiary standards, or other factors used to apply the limitation to medical/surgical benefits within the classification (or applicable sub-classification) (45 CFR § 146.136(c)(i)). The health carrier shall perform and document comparative analyses of the design and application of NQTLs in accordance with 42 U.S.C. § 300gg-26(a)(8)(A).**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage.

**Priority:** Recommended

**Documents to be Reviewed**

\_\_\_\_\_ Applicable state and federal statutes, rules, regulations and published sub-regulatory guidance

\_\_\_\_\_ A list of all NQTLs imposed upon mental health or substance use disorder benefits within each classification of benefits (or applicable sub-classification), including the methodology used to determine those NQTLs. A state may focus its review on a subset of NQTLs rather than all NQTLs. (See reference link to DOL Self-Compliance Tool for a non-exhaustive list) Note: Due to the significant number of potential NQTLs, it is advised that the examiner selects a targeted subset or sample of NQTLs based on examination resources, state specific concerns, company common practices, etc. to avoid the review of hundreds of service variations. Additional NQTLs can be phased into the review as appropriate.

\_\_\_\_\_ Utilization management manuals and utilization review documents such as: utilization review criteria; criteria hierarchies for performing utilization review; case management referral criteria; initial screening scripts and algorithms; policies relating to reviewer discretion; processes for identifying and evaluating clinical issues and utilizing performance goals

\_\_\_\_\_ Notes and/or logs kept during utilization review, such as those describing: peer clinical review; telephonic consultations with attending providers; consultations with expert reviewers; clinical rationale used in approving or denying benefits; the selection of information deemed reasonably necessary to make a medical necessity determination; adherence to utilization review criteria and criteria hierarchy; professional judgment used in lieu of utilization review criteria; actions taken when incomplete information is received from attending providers

\_\_\_\_\_ Company claim procedure manuals and bulletins/communications

\_\_\_\_\_ Claims processor and customer services MHPAEA training materials

\_\_\_\_\_ Company fraud, waste, and abuse policies and procedures

\_\_\_\_\_ Internal company claim audit reports

\_\_\_\_\_ Prescription drug formulary for each product/plan design

\_\_\_\_\_ Prescription drug utilization management documentation

- \_\_\_\_\_ Fail-first policies or step therapy protocols
- \_\_\_\_\_ Network development/contracting policies and procedures
- \_\_\_\_\_ Standards for provider admission to participate in a network, including credentialing requirements
- \_\_\_\_\_ Standards for determining provider reimbursement rates
- \_\_\_\_\_ Samples of provider/facility contracts in use during the exam period
- \_\_\_\_\_ Plan methods for determining usual, customary and reasonable charges for each product/plan design
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical claim files.
- \_\_\_\_\_ Mental health and/or substance use disorder and medical/surgical utilization review procedures
- \_\_\_\_\_ Complaint files, logs and disposition notes
- \_\_\_\_\_ Documentation, including but not limited to comparative analyses, demonstrating that within each of the 6 classifications of benefits (and applicable sub-classifications), the as written and in operation processes, strategies, evidentiary standards, or other factors used in applying a NQTL are comparable to and applied no more stringently to mental health or substance disorder benefits than to medical/surgical benefits in the classification.

#### Other References

Enforcement of the Public Health Services Act  
42 U.S. Code § 300gg–22

Preemption relating to the Public Health Services Act  
42 U.S. Code § 300gg–23

Mental Health Parity and Addiction Equity Act of 2008  
42 U.S. Code § 300gg–26

Publication of summary plan description  
ERISA 104(b) (29 U.S.C. § 1024(b))

The Federal Parity Self Compliance Tool: <https://www.dol.gov/sites/dolgov/files/EBSA/laws-and-regulations/laws/mental-health-parity/self-compliance-tool.pdf>

FAQs about Mental Health and Substance Use Disorder Parity Implementation and the Consolidated Appropriations Act, 2021 Part 45 (ACA FAQ 45): <https://www.dol.gov/sites/dolgov/files/EBSA/about-ebbsa/our-activities/resource-center/faqs/aca-part-45.pdf>

#### Review Procedures and Criteria

Review the list of all NQTLs imposed on mental health/substance use disorders and choose a sample.

Review the health carrier’s comparative analyses to verify that within any classification of benefits, as written and in operation, the process, strategies, evidentiary standards, or other factors used in applying an NQTL to mental health or substance disorder benefits are comparable to, and are applied no more stringently than those used in applying the limitation with respect to medical/surgical benefits in the classification. The comparative analyses

shall include the following, for each NQTL applied to mental health or substance use disorder benefits, separately for each classification of benefits (42 U.S.C. § 300gg-26(a)(8)(A)):

- The specific coverage terms or other relevant terms regarding the NQTL and a description of all mental health or substance use disorder and medical or surgical benefits to which such NQTL applies in each respective benefits classification;
- The factors used to determine that the NQTL will apply to mental health or substance use disorder benefits and medical or surgical benefits;
- The evidentiary standards used for the factors identified, when applicable, provided that every factor shall be defined, and any other source or evidence relied upon to design and apply the NQTL to mental health or substance use disorder benefits and medical or surgical benefits;
- The comparative analyses demonstrating that the processes, strategies, evidentiary standards and other factors used to apply the NQTL to mental health or substance use disorder benefits, as written and in operation, are comparable to, and are applied no more stringently than, the processes, strategies, evidentiary standards, and other factors used to apply the NQTL to medical or surgical benefits in the benefits classification; and
- The specific findings and conclusions reached by the health carrier with respect to the health insurance coverage, including any results of the analyses described in 42 USC 300gg-26(a)(8)(A) that indicate that the health carrier is or is not in compliance with 45 CFR 146.136(c)(4).

The health carrier's analyses must contain the following, at a minimum (ACA FAQ 45 Q2):

1. A clear description of the specific NQTL, plan terms and policies at issue;
2. Identification of the specific mental health or substance use disorder and medical/surgical benefits to which the NQTL applies within each benefit classification, and a clear statement as to which benefits identified are treated as mental health or substance use disorder and which are treated as medical/surgical;
3. Identification of any factors, evidentiary standards or sources, or strategies or processes considered in the design or application of the NQTL and in determining which benefits, including both mental health or substance use disorder benefits and medical/surgical benefits, are subject to the NQTL. Analyses should explain whether any factors were given more weight than others and the reason(s) for doing so, including an evaluation of any specific data used in the determination;
4. To the extent the health carrier defines any of the factors, evidentiary standards, strategies, or processes in a quantitative manner, it must include the precise definitions used and any supporting sources;
5. The analyses, as documented, should explain whether there is any variation in the application of a guideline or standard used by the health carrier between mental health or substance use disorder and medical/surgical benefits and, if so, describe the process and factors used for establishing that variation;
6. If the application of the NQTL turns on specific decisions in administration of the benefits, the health carrier should identify the nature of the decisions, the decision maker(s), the timing of the decisions and the qualifications of the decision maker(s);
7. If the health carrier's analyses rely upon any experts, the analyses, as documented, should include an assessment of each expert's qualifications and the extent to which the health carrier ultimately relied upon each expert's evaluations in setting recommendations regarding both mental health or substance use disorder and medical/surgical benefits;
8. A reasoned discussion of the health carrier's findings and conclusions as to the comparability of the processes, strategies, evidentiary standards, factors and sources identified above within each affected classification, and their relative stringency, both as applied and as written. This discussion should include citations to any specific evidence considered and any results of analyses indicating that the health carrier is or is not in compliance with MHPAEA; and
9. The date of the analyses and the name, title and position of the person or persons who performed or participated in the comparative analyses.

The health carrier shall avoid the following practices and procedures when responding to a request for comparative analyses (ACA FAQ 45 Q3):

1. Production of a large volume of documents without a clear explanation of how and why each document is relevant to the comparative analysis;

2. Conclusory or generalized statements, including mere recitations of the legal standard, without specific supporting evidence and detailed explanations;
3. Identification of processes, strategies, sources and factors without the required or clear and detailed comparative analysis;
4. Identification of factors, evidentiary standards and strategies without a clear explanation of how they were defined and applied in practice;
5. Reference to factors and evidentiary standards that were defined or applied in a quantitative manner, without the precise definitions, data, and information necessary to assess their development or application; and
6. Analysis that is outdated due to the passage of time, a change in plan structure, or for any other reason.

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 6**

**The health carrier shall ensure that it complies with all availability of plan information and related disclosure obligations regarding: 1) criteria for medical necessity determinations; 2) reasons for denial of services; 3) information relevant to medical/surgical, mental health and substance use disorder benefits; 4) rules regarding claims and appeals, including the right of claimants to free reasonable access to and copies of documents, records and other information including information on medical necessity criteria for both medical/surgical benefits and mental health and substance use disorder benefits, as well as the processes, strategies, evidentiary standards and other factors used to apply a NQTL with respect to medical/surgical benefits and mental health or substance use disorder benefits under the plan, including any analyses performed by the carrier as to how the NQTL complies with MHPAEA.**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

- \_\_\_\_\_ Plan policies and procedures for responding to participant requests for medical necessity criteria for either or both mental health and substance use disorder services and medical/surgical services
- \_\_\_\_\_ Plan policies and procedures for responding to requests for information on the processes, strategies, evidentiary standards and other factors used to apply a NQTL with respect to medical/surgical benefits and mental health or substance use disorder benefits under the plan
- \_\_\_\_\_ Sample adverse benefit determination letters
- \_\_\_\_\_ Sample letters responding to disclosure requests for medical necessity criteria and information on NQTLs
- \_\_\_\_\_ Policies and procedures for classifying denials as administrative or medical necessity
- \_\_\_\_\_ Internal and external appeals files for mental health and substance use disorder services adverse benefit determinations
- \_\_\_\_\_ Log of disclosure requests, including date requested, date responses was provided, samples of documents sent in response

**Other References**

- 45 CFR § 146.136(d)
- ERISA 104
- 29 CFR § 2520.104b-1
- 29 CFR § 2560.503-1
- 29 CFR § 2590.715-2719

**Review Procedures and Criteria**

Review the health carrier's method for providing to any current or potential participant, beneficiary, or contracting provider upon request the medical necessity criteria used to make mental health or substance use disorder determinations (45 CFR § 146.136(d)(1)).

Review the health carrier's letters providing the reason for any denial of reimbursement for mental health or substance use disorder benefits and verify that the letters are dated within 30 days of the request (45 CFR § 146.136(d)(2)).

Review the health carrier's policy & procedure for responding promptly to requests for all documents, records and other information relevant to an adverse benefit determination, including medical necessity criteria and the comparative analysis required under (42 USC 300gg-26(a)(8)(A)), disclosures referenced above (45 CFR § 146.136(d)(3)) as referenced in ACA FAQ 45-Q6.

Document that the health carrier's claims processing and disclosure regarding adverse benefit determinations complies with the federal claims and appeals regulations (45 CFR § 147.136).

**STANDARDS**  
**Mental Health and Substance Use Disorder Parity Compliance**

**Standard 7**

**The health carrier as the entity is responsible for parity compliance. The health carrier shall ensure that management of mental health and substance use disorder benefits coverage as a whole complies with the applicable provisions of MHPAEA, including any vendor relationships. The carrier shall provide or require sufficient information in terms of plan structure and benefits to or from any vendor to ensure that the mental health and substance use disorder benefits are coordinated with the medical/surgical benefits for purposes of compliance with the requirements of MHPAEA.**

**Apply to:** Certain group and individual health carriers offering mental health and substance use disorder coverage

**Priority:** Recommended

**Documents to be Reviewed**

\_\_\_\_\_ Contractual agreements between the carrier and vendors having administrative, claims and/or medical management responsibilities

\_\_\_\_\_ Policies and procedures for ensuring availability of health carrier information needed for vendor analysis of compliance with MHPAEA

\_\_\_\_\_ A narrative summary outlining how the vendor and the carrier coordinate benefit design and application to ensure compliance with MHPAEA

\_\_\_\_\_ Select written communications relevant to mental health and substance use disorder benefits between the carrier and the vendor

**Other References**

29 CFR § 2590.712(e).

75 FR § 5426

78 FR § 68250

**Review Procedures and Criteria**

Review the contractual agreements between the health carrier and any vendors providing administrative, claims and/or medical management responsibilities.

Review the health carrier's protocols and procedures to document that any contracted vendors are collaborating with the health carriers to satisfy compliance with MHPAEA. This shall include explanation of how both the design of benefits and the application of benefits, in operation, are compliant with MHPAEA.

Review any audits the health carrier has completed of its vendors to ensure compliance with MHPAEA.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

## **The Market Information Systems Research and Development (D) Working Group Review of Artificial Intelligence Techniques in Market Analysis**

### **Executive Summary**

This report fulfills the Market Information Systems Research and Development (D) Working Group charge to evaluate the potential benefits of artificial intelligence (AI) in relation to market analysis. After careful consideration, the Working Group concluded that there may be possible benefits to improve analysis techniques. Several caveats are discussed as well. AI may not be suitable for data currently available to state insurance regulators. In addition, some of the techniques perform complex data mining operations, which can produce results that lack a clear interpretation. Lastly, AI techniques are designed for, and many require, very large datasets. As such, AI should be contemplated in the context of a long-range plan, beginning with repairing known issues with existing data, and employing more rigorous traditional statistical techniques to assess predictive accuracy of analytical tools. Subsequently, state insurance regulators can consider the acquisition of data appropriate to AI.

### **Introduction**

In early 2021, the Market Information Systems Research and Development (D) Working Group received a charge from the Market Information Systems (D) Task Force to explore possible applications of artificial intelligence (AI) methods in market analysis. An early difficulty encountered by the Working Group is that the term “AI” itself has a variety of contested meanings. In addition, private sector entities have adopted the term as a marketing concept and inappropriately apply the label to products simply as a selling point. As such, the term has come to acquire a variety of meanings and is an “essentially contested concept.”<sup>1</sup>

At its most general level, the term “AI” implies machine capacities that mimic or are analogous to processes of human reasoning and learning and entail some degree of machine autonomy in which learning occurs without significant human intervention. Beyond this general description, the Working Group did not feel that an attempt to define the term more strictly would be fruitful. Rather, the term is employed simply as a shorthand reference for a collection of various techniques that algorithmically seek patterns in data that are predictive of some future outcome. Common methods include machine learning, neural networks, and decision tree analysis. These processes are often contrasted to the traditional hypothetical-deductive methods of model specification associated with classical statistics. However, there does not appear to be a bright line of demarcation so that a particular technique can be firmly fixed within either category.

In addition, the Working Group focuses on what is commonly called “narrow AI,” in which machine algorithms are employed for narrowly defined and limited tasks. More advanced systems, called

---

<sup>1</sup> The term “essentially contested concept” was coined by W.B. Gallie in the seminal presentation to the Aristotelian Society in 1956.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

“general AI,” possess generalized autonomous problem-solving capacities that are comparable to the processes of the human brain, and they are able to adapt to novel situations or information (Macnish et al., 2019).

It is important to emphasize the ways in which AI modeling techniques contrast to the standard scientific model employed in classical or traditional statistics:

**Classical Statistics:** Method of hypothetical-deductive reasoning in which hypotheses are clearly and narrowly specified *prior* to data testing, often with a prior understanding of the underlying causal nature of the relationships between variables. **Purpose:** To further causal understanding.

**AI:** Often employs a type of “data mining” in which a machine pattern-seeking algorithm is released “into the wild” to identify possible correlations between variables that may be predictive of some independent variable. Hypotheses are not specified prior to data analysis, and the algorithm may very well identify correlations that would not have occurred to an analyst and whose causal relationship is constructed post-hoc (to the degree that AI users are concerned with causality at all). **Purpose:** Predict future outcomes or events.

The difference between these two approaches is not trivial, and significant disagreements about the advantages and disadvantages of AI remain. It is of note that AI did not emerge principally from university statistics departments, but rather from the field of computer science. Many statisticians remain skeptical of the techniques and have offered up a variety of caveats for their use. For example, recently the American Statistical Society (ASA) reacted to the “reproducibility crisis” afflicting some disciplines that have discovered, with much consternation, that a large volume of published works could not be replicated. The concern was that increasingly less rigorous statistical methods departing from the hypothetical-deductive approach were becoming more prominent in a variety of fields, undermining confidence on research findings. Remarking on departures from a rigorous hypothetical-deductive approach with “data mining” and like methods in which pattern seeking is largely ceded from a researcher to a machine, the ASA warned about improper inferences that might result from such techniques. The ASA centered its discussion on the p-value, related to the probability that some observed relationship occurred by chance along. A low p-value is often employed to minimize the probability that chance relationships will be misinterpreted as a relationship that is a meaningful, non-random outcome:

“Conducting multiple analyses of the data and reporting only those [analyses] with certain p-values...renders the reported p-values essentially uninterpretable. Cherry-picking promising findings, also known by such terms as data dredging, significant chasing, significance questions, selective inference and a ‘p-hacking’ leads to a spurious excess of statistically significant results...and should be vigorously avoided” (Wasserstein & Lazar, 2016).

To translate the ASA’s statement into more easily understood and less technical terms, the ASA is warning against *false positives* in which an analysis produces random or chance correlations between items that are not meaningfully related—that is, where a chance relationship is mistaken for a true causal relationship. That AI largely jettisons causal understanding as its primary goal (to the degree that causality is a concern at all) increases the probability that statistical results may be uninterpretable in any meaningful sense. This is clearly evinced by the increasing debate among state insurance

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

regulators and insurers regarding the meaning of statistical relationships appearing in predictive models that lack intuitive or, in many cases, even plausible explanations. See Appendix A for further discussion of the ASA statement.

The discussion above is not intended to sway state insurance regulators one way or the other with respect to AI. The purpose is simply to proffer some caveats shared by many statisticians. A final caveat is the AI techniques were developed to analyze very large data sets consisting of millions of records and possibly thousands or tens of thousands of variables. It is said to have an advantage in that algorithms can perform a large volume of analyses across different constellations of variables in a way that would be highly impractical employing traditional (and manual) model building. For small data sets, such as the limited data currently available to market analysts, it is unclear whether the expense associated with developing AI techniques can be justified, nor whether AI is at all superior to traditional model building methods. This is not an unimportant point and is discussed in more depth elsewhere in this recommendation.

### **Current Status of Market Analysis**

Quantitative market analysis relies on just a handful of data sources:

**The Complaint Database System (CDS):** The NAIC compiles complaints against insurers received by state insurance regulators. Thus, each state has access to a national-level database. Complaint indices are “normalized” by expressing the volume of complaints to premium, compared with the overall industry total.

**The Regulatory Information Retrieval System (RIRS):** Regulatory actions in relation to insurance entities are captured in the RIRS database. Actions range from intervention in financially troubled entities to violations of producers and insurance carriers. Each record identifies the cause of the action, as well as any orders, fines, or restitution amounts. The RIRS database is currently being substantially revised to capture significantly more detail.

**The Market Actions Tracking System (MATS):** The MATS database captures information pertaining to market conduct exams, as well as actions short of exams. Data captured include area of scrutiny (claims, underwriting, etc.) and the outcome of the market action (order, fine, etc.). By matching MATS actions with RIRS, additional detail about the nature of the violation can be assessed.

**The Market Conduct Annual Statement (MCAS):** The MCAS was developed to capture data with the primary purpose of assessing an insurer’s market performance and identify potential market irregularities. The data focus primarily on claims handling and underwriting, and data are scrutinized with respect to claims processing times and denials, nonrenewal and cancellation practices, and overall turnover in a book of business. Data are captured by line and coverage. To date, MCAS data are collected for life and annuities, private automobile, homeowners, health (both on and off the federally facilitated marketplace [FFM]), long-term care (LTC), lender-placed insurance, disability income, and private flood.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

**Miscellaneous Data Sources:** Some financial data has been incorporated into market information systems. Insurers that are under financial stress, or that rapidly expand into or contract out of a line of business, or that exhibit high defense or other adjudication costs, may be subjected to additional analysis. While financial indicators are only indirect or proxy measures of potential market issues, and by themselves may have no clear market-based interpretation, interpretation within the context of a host of other indicators may be reflective of the present of a market-relevant issue.

The NAIC, in conjunction with state insurance regulators, has developed a broad scope “market score” that incorporates much of the data referenced above, which is made available to regulators via the Market Analysis Prioritization Tool (MAPT). One such data are “normalized” by the premium volume and scope of company operations as necessary. For example, several RIRS-based ratios express the volume of RIRS actions in relation to premium volume, the number of states in which they have significant premium, and a composite ratio that incorporates both premium and scope. Each ratio is given a score, and their contribution to the overall score weighted according to their perceived predictive relevance. For example, financial ratios are accorded significantly less weight than complaints, as their relationship to market misconduct is considered more speculative and indirect.

An important caveat is that predictive analytics is not well developed in market regulation. The ratios employed in the Market Analysis Review System (MARS) have not been subjected to rigorous statistical tests that demonstrate their analytic utility. While some work has been performed in this regard, such work is significantly hampered by a dearth of appropriate data. For example, future RIRS actions are often employed as the dependent variable (the outcome of interest to be predicted). However, this presents all manner of statistical challenges. While it is certainly reasonable to use prior outcomes (past RIRS actions) to predict future outcomes (the RIRS actions to be predicted), employing RIRS actions as both dependent and independent variable introduces significant complexities in the interpretation of any observed relationship between the two. One can imagine, for example, that the use of RIRS actions in market analysis invites greater scrutiny to a given insurer, and that in turn generates future regulatory actions precisely because the company received additional scrutiny. Companies that have no “prior offenses” fail to attract regulatory scrutiny, so that any infractions may escape regulatory action *for precisely that reason*. This problem is certainly not insurmountable, but it must be explicitly recognized in any model building exercise, whether with AI or with more conventional statistical techniques.

In general, the paucity of rich data sources has significantly hampered the adoption of more rigorous analytical techniques. To return to RIRS, these data are not rich sources of detailed information. Schematics are not well designed “from the ground up.” Essential data are missing, such as line of business.

Any consideration of AI or any other analytical techniques must necessarily view the utility of such techniques within the context of available data. Regardless of the validity of a technique in general, it will have limited utility if data are themselves limited. Any recommendation to employ such methods must therefore at the same time recommend a thorough review of available data.

Importantly, results of quantitative analysis are always treated as merely suggestive and tentative and are regarded as at most a precursor to more qualitative analysis. It currently is employed to prioritize

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

entities that may merit additional scrutiny and to narrow focus on a much more limited subset of companies out of a larger pool of companies. It therefore primarily prioritizes limited regulatory resources.

State insurance regulators avail themselves of the formal analytical processes adopted by the NAIC. Quantitative or “baseline” analysis identifies entities with anomalous indicators that significantly depart for industry-wide values. A “level 1” analysis may be pursued, in which an analyst devotes additional scrutiny to such things as complaint trends, common reasons complaints are lodged against an insurer, similarities in RIRS actions, etc. If concern still remains (or additional concerns are identified) subsequent to level 1 analysis, a structured level 2 analysis may be performed. A level 2 analysis requires a much greater commitment of time and resources. For example, rather than just manually reviewing complaint data to identify patterns, an analyst may manually review actual complaint documentation to garner a more detailed understanding of the nature of complaints.

As a preliminary to the following discussion, AI/statistical analysis may have two primary functions within the context of the current market analysis structure:

1. More accurately identify companies that merit the additional expenditure of resources necessary to perform the more labor-intensive level 1 and level 2 analyses. Analysis processes that more efficiently identify problem companies for this purpose are by definition more effective and more effectively target resources by avoiding “false positives” (for lack of a better word).
2. Potentially, AI methods could assume many of the functions that are currently performed manually. For example, many of the pattern-seeking analysis performed by analysts in a level 1 review could conceivably be more efficient if automated. Potentially, AI could identify patterns that might elude a human analysis. A very advanced level of AI could perhaps assume complex analysis involved with manually reviewing complaint files and documents. However, while the possibility is raised here, it is not further pursued. That level of AI suitable for tasks may not even exist as yet, or if it does, it may be so specialized that it may not be available to state insurance regulators. Even if available, the likely enormous costs themselves would render them highly impractical.

Whether such AI exists, is available at a practical cost, and can actually out-perform more conventional analyses are questions that the Market Information Systems Research and Development (D) Working Group is simply unable to satisfactorily address. The Working Group merely suggests initially limiting the scope of ambitions to a few methods that are commonly, if not universally, recognized as AI, such as machine learning or neural networks. More expansive or ambitious efforts may result in a fruitless search for “unobtainium.”<sup>2</sup>

Given very large data sets, well beyond what is currently available to market analysts, AI may have clear advantages to more conventional approaches. The slow, methodical, hypothetical-deductive

---

<sup>2</sup> A tongue-in-cheek term originating among engineers in the 1950s. It is defined by Wikipedia as “... any hypothetical, fictional, or impossible material, but it can also mean a tangible but extremely rare, costly, or reasonably unobtainable material. Less commonly, it can refer to a device with desirable engineering properties for an application, but which are exceedingly difficult or impossible to achieve.”

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

approach that forms the core of conventional statistics may have advantages in terms of generating valid causal conclusions. However, AI may have certain advantages with respect to confronting the enormity of modern data. As AI is well-suited to performing much more expansive analysis and pattern-seeking routines over vast quantities of data, it may well identify predictive patterns that would have escaped conventional analysis or that are counterintuitive such that some hypotheses may never have occurred to an analyst employing a standard hypothetical-deductive approach. However, there are distinct disadvantages as well, and they are shared by other approaches often termed “data mining.” The fact is that patterns may lack an intuitive meaning, and the manner in which such patterns are identified and rendered for interpretation may be unclear. Additionally, patterns may generate numerous “false positives,” apparent patterns or correlations that are purely random and possess no meaning or any real predictive power whatsoever. This is not fatal for AI techniques, but it introduces much in the way of caveats and requires significant remedial measures to be employed. This problem is so significant that it merits a much fuller discussion in a separate section below.

## **The Work of Market Information Systems Research and Development (D) Working Group**

The Working Group solicited input from various parties. Two parties delivered presentations to the Working Group:

1. On June 16, 2021, the Working Group discussed a presentation regarding AI methods currently being explored by NAIC staff to predict which insurers are likely to experience financial stress, including insolvency. Beginning in January 2021, an outside consulting group was retained to develop both AI as well as more traditional statistical techniques to construct predictive models of insolvency risk. The efforts are ongoing at the time of writing. Presenters believed the methods were promising and could significantly advance financial risk surveillance. Among AI and statistical models explored were decision tree analysis, generalized linear models (GLMs), and logistic regression.
2. During the Working Group’s June 21, 2021, meeting, Birny Birnbaum (Center for Economic Justice—CEJ) encouraged the Working Group to adopt a long-term perspective and develop a multiyear plan to explore AI techniques that might be beneficial to market analysis. He also indicated that state insurance regulators have to date failed to acquire granular transactional data that could be exploited by AI methods to afford a much more robust surveillance system to reduce consumer harm to the extent possible.

After the meeting, the Working Group convened a subject-matter expert (SME) group with the intent of creating a draft recommendation to be submitted to the Working Group.

## **Recommendations**

The Working Group recommends developing a long-range plan, in a sequence of five steps.

### **I. Existing Market Analysis Data**

As noted above, market analysis suffers from a paucity of detailed data. Some movement in expanding data and remedying deficiencies was made with a complete redesign of the RIRS data, which will facilitate analysis of factors related to an entity sanctioned by state insurance regulators. If

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

implemented, RIRS will also capture much more detailed data related to the specific misconduct that garnered a regulatory response. The RIRS proposal is currently under discussion with the Market Information Systems (D) Task Force, to which Working Group reports.

The remainder of available data also suffers from significant deficiencies. Insurers employ a variety of definitions to produce MCAS data. Even such a fundamental concept as a “claim” is reported differently by different insurers, making market-wide analysis challenging. For example, the MCAS defines a claim in the conventional sense of “a demand for payment.” Investigation by the Missouri Department of Commerce & Insurance (DCI) has determined that the definition is interpreted in wildly divergent ways across the industry that simply makes meaningful comparison impossible and renders key market indicators or ratios largely meaningless. Some insurers set up a claim on a coverage that is reasonably related to the facts of the incident as relayed by a claimant. Other insurers set up all possible coverages on a policy as a claim in their internal systems regardless of whether those coverages might be reasonably implicated in a claim. As might be imagined, those carriers have significantly higher ratios of claims closed without payment. This and other issues remain with the MCAS and significantly impair market analysis.

**Recommendation 1:** Survey currently available market analysis data, and identify substantive deficiencies based on the nature and substance of the data elements collected. Ensure that all data are consistently reported across insurers to the degree practical and ensure adherence to definitions of data elements.

## II. Existing Methods of Market Analysis

Current quantitative methods of market analysis are large based on *ad hoc* and *intuitive* understanding of how data indicators might be related to market misconduct. For example, one of the earliest indicators developed are complaints received by state insurance regulators regarding insurers. It is probably not unreasonable to interrogate complaint data to identify trends over time, as well as just overall complaint volume, to attempt to identify potential problems in a market. Similar indices consider the volume of RIRS actions, as well as the gravity of infractions in terms of potential consumer harm. It is the opinion of many state insurance regulators that such indicators possess a rational relationship to market misconduct and are relevant to identify market actors that might benefit from a heightened level of regulatory scrutiny.

While the Working Group agrees with the rationale behind such market indicators, analytical tools have not to date been subjected to more rigorous statistical methods to clearly identify the predictive power and assess their relative importance or weight. For example, the MAPT, maintained by the NAIC and available to state insurance regulators, employs overall insurer scores based on various indicators. However, the weight of these indicators employed in the score were assigned by state insurance regulators based on experience, as well as assessment of whether a likely relationship have a clear rational meaning. For example, complaint ratios are weighted significantly more heavily than things like financial indicators. The Working Group believes subjecting the scoring system to rigorous statistical analysis could yield significant benefits in identifying problem market actors.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

**Recommendation 2:** In conjunction with recommendation 1 (assess data quality), state insurance regulators should adopt a much more rigorously statistical approach to identify the predictive power of market scoring systems, assess how each variable should be weighted in terms of its unique contribution to productiveness, and drop those that lack analytic utility. In addition, effort should be made to integrate data into a single overall analysis. For example, the MAPT does not incorporate MCAS data, which is typically subject to a separate analysis. The Working Group believes that a “piecemeal” approach is likely less effective than a more integrated approach.

It is noted that the current state of data will likely prove limiting and that such efforts may not make much progress until additional data are made available (such as the proposed revisions to the RIRS data, currently subject to NAIC discussion).

### III. Available Approaches: Exploring AI

In addition to more traditional statistical tools, such as various types of regression models and correlation analyses, AI may offer additional benefits. Some commercial statistical packages have incorporated AI methods. The statistics package SAS, which is widely used in both the private and public sectors, makes some AI techniques available in its standard statistical module.<sup>3</sup> In addition, SAS has developed a module called Enterprise Miner, which incorporates both data mining and some lower-level AI routines. (For those familiar with the terms, it performs such things as decision-tree analysis, neural networks, and like forms of analyses). Other modules make machine learning available—a potentially powerful type of analysis that modifies prior predictive algorithms as new data become available.

**Recommendation 3:** In undertaking recommendation 2, incorporate various promising AI modes of analyses, as well as traditional statistical modeling. Constantly assess the precision of model outcomes relative to objectives such as identifying potential market issues.

### IV. Qualitative Analysis

The current model of market analysis incorporates a multistage hierarchical structure. First, quantitative analysis such as that produced by the MAPT identifies potential market problems and narrows focus to entities that appear to exhibit potential areas of regulatory concern. Having narrowed down the focus of analysis to a much more limited pool of candidates, market analysts in the states engage in more manual or qualitative analysis of additional information sources. For example, an analyst may review a selection of complaint files to identify additional patterns of market behavior to better understand their nature and substance.

---

<sup>3</sup> SAS is markets in “modules,” each consisting of a different suite of capabilities that can be tailored to a user’s need. For example, “base SAS” provides standard data handling programs. A “statistics module” provides a wide-ranging set of analytical routines.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

As noted above, AI techniques such as text analysis could potentially expand such exercises and improve the identification of concerning patterns at a deeper level, as well as assess ways to improve the efficiency of other qualitative tasks.

**Recommendation 4:** Assess ways AI can improve both the efficiency of *qualitative* analysis and facilitate pattern recognition across larger volumes of textual evidence, including most especially complaints, but perhaps other textual sources. For example, the “level 1” analysis formalized in NAIC market system may include a review of the “management discussion and analysis” of the financial annual statement.

## V. Longer-Range Planning

As noted above, data mining and AI techniques were developed primarily as tools to analyze large volumes of data. For data past a certain magnitude, including especially those containing many hundreds or even thousands of variables, the traditional hypothetical-deductive cornerstone that is the cornerstone of traditional statistical inference may be ill-suited as well as cost-prohibitive in terms of time and resources. If the purpose is solely prediction as opposed to causal understanding, AI can fine-tune predictive algorithms by testing relationships that may be unlikely to occur to a statistician employing causal modeling.

Currently, such large volumes of data are unavailable to market analysts, though they could potentially be obtained. More granular data pertaining to claims, underwriting, and other areas of company operations are routinely collected via the “standard data requests” adopted as a supplement to the *Market Regulation Handbook* and commonly employed in market conduct exams.

However, AI and data mining can churn up counterintuitive statistical relationships that defy ready interpretation. In addition, it is likely to detect proxy relationships that are not understood. Proxy relationships, in which a third variable is substituted for an underlying variable of interest, are often employed in statistical models. This is often due to the accessibility or cost of obtaining data of the actual causal variable of interest. However, when employed in traditional statistical analysis, the nature of the relationship between the proxy variable and the actual variable of interest is generally well understood. This is not true of AI techniques that employ or resemble data mining.

The techniques are also likely to generate some number of purely chance relationship, where a correlation is generated by random chance. Inferential statistics seek to minimize mistaking a chance relationship for a meaningful association. Typically, the use of a p-value requirement of 0.05 or less limits the probability of accepting a random relationship to no more than 5% of occurrences. However, a 5% threshold means that over time, false, or chance relationships will be misinterpreted of a true correlation.

This fact is not fatal for the use of AI in market analysis, but it does represent a strong caveat for those employing the techniques, at least those that share elements with data mining. Careful interpretations of p-values should recognize an increased possibility of false positives. Observed relationships should be assessed and validated over time to ensure correlations are stable. In addition, once relationships

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

are identified via AI and found useful, standard statistical models should also be employed to test whether different techniques yield superior predictive power. Additional discussion of caveats is presented in the appendix.

That said, there is much potential of AI in market analysis, *assuming that additional, more granular, data are available*. As noted, such techniques are most suited for large datasets whose very size would make a standard statistical approach impractical just given the sheer number of possible correlations available for testing.

**Recommendation 5:** Systematically explore potential data sources suitable for AI techniques, with an eye for discovering patterns and relationships in relation to some well-defined outcome one is attempting to predict. This may be identifying entities that may merit additional regulatory scrutiny in a way that is currently done by the less sophisticated methods employed in the MAPT or with the MCAS. Larger volumes of data, such as the standard data requests, can be subjected to AI to identify problematic claims handling, underwriting, and other insurer practices.

## Summary of Recommendations

**Recommendation 1:** Survey currently available market analysis data, and identify substantive deficiencies based on the nature and substance of the data elements collected. Ensure that all data are consistently reported across insurers to the degree practical, and ensure adherence to definitions of data elements.

**Recommendation 2:** In conjunction with recommendation 1 (assess data quality), state insurance regulators should adopt a much more rigorously statistical approach to identify the predictive power of market scoring systems, assess how each variable should be weighted in terms of its unique contribution to productiveness, and drop those that lack analytic utility. In addition, effort should be made to integrate data into a single overall analysis. For example, the MAPT does not incorporate MCAS data, which is typically subject to a separate analysis. The Working Group believes that a “piecemeal” approach is likely less effective than a more integrated approach.

**Recommendation 3:** In undertaking recommendation 2, incorporate various promising AI modes of analyses, as well as traditional statistical modeling. Constantly assess the precision of model outcomes relative to objectives, such as identifying potential market issues.

**Recommendation 4:** Assess ways AI can improve both the efficiency of *qualitative* analysis and facilitate pattern recognition across larger volumes of textual evidence, including most especially complaints, but perhaps other textual sources. For example, the “level 1” analysis formalized in NAIC market system may include a review of the “management discussion and analysis” of the financial annual statement.

**Recommendation 5:** Systematically explore potential data sources suitable for AI techniques, with an eye for discovering patterns and relationships in relation to some well-defined outcome one is attempting to predict. This may be identifying entities that may merit additional regulatory scrutiny in

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

a way that is currently done by the less sophisticated methods employed in the MAPT or with the MCAS. Larger volumes of data, such as the standard data requests, can be subjected to AI to identify problematic claims handling, underwriting, and other insurer practices.

**Appendix: Caveats**

Recently, some fields of scientific inquiry have experienced much consternation and hand-wringing due to the so-called “replicability crisis” resulting from the realization that many studies published in top-tier journals could not be replicated. In 2015, Open Science Collaboration published research into the replicability of psychological studies. Of the 100 studies that were subjected to testing, replications yielded statistically significant results in only 36% compared to 97% of the original publications (Open Science Collaboration, 2015). Similar reproducibility issues were found in other fields.

Attention was directed at quantitative methods, particularly those made possible by modern computing power. Researchers can run countless variations of models, including multiple different variables, cross-effects, and other tweaks, until they eventually produce positive or statistically significant results. The inevitable outcome of the lack of rigor of such methods is that many chance correlations will be mistaken for meaningful relationships.

Think of it this way. The probability of obtaining all heads from 10 flips of a fair coin is  $1/1024$ . So, if a researcher actually performed the experiment 1,024 times and obtained 10 heads at least once, it would obviously be improper to infer that the coin was a two-headed coin. Without knowledge of the total number of trials, one might reject the “null hypothesis” that the coin is fair, and results would be “statistically significant” with a p-value of  $(1/1,024) = 0.00098$ , well below the 0.05 maximum threshold to establish statistical significance. But the true p-value can only be calculated with knowledge of the total number of trials prior to obtaining the recorded result, such that the true p-value is well above the maximum threshold.

There are no allegations of willful misconduct so much as careless and sloppy methods, producing much introspection about how statistics methods are taught to scientists at colleges and universities. The problem is so significant that the following year, the American Statistical Association (ASA) released a statement regarding misuse of p-values and practices known as “p hacking” or “data dredging.” A letter from the ASA is reprinted below, with a link to the full statement (used with permission).

Really, this is a warning for state insurance regulators not to adopt a casual attitude about apparent relationships turned up by the methods. When such methods are employed, modelers should be on constant guard against mechanical interpretations of model outputs. It is important to fully understand what is going on in the “black box” of an AI algorithm, the results of all statistical tests performed, and the totality of processes generating final results.

A high number of false positives that prompt regulatory follow-up can risk draining away regulatory resources going down blind allies.

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES

*Provides Principles to Improve the Conduct and Interpretation of Quantitative Science*

March 7, 2016

The American Statistical Association (ASA) has released a “Statement on Statistical Significance and P-Values” with six principles underlying the proper use and interpretation of the  $p$ -value [<http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108#.Vt2XIOaE2MN>]. The ASA releases this guidance on  $p$ -values to improve the conduct and interpretation of quantitative science and inform the growing emphasis on reproducibility of science research. The statement also notes that the increased quantification of scientific research and a proliferation of large, complex data sets has expanded the scope for statistics and the importance of appropriately chosen techniques, properly conducted analyses, and correct interpretation.

Good statistical practice is an essential component of good scientific practice, the statement observes, and such practice “emphasizes principles of good study design and conduct, a variety of numerical and graphical summaries of data, understanding of the phenomenon under study, interpretation of results in context, complete reporting and proper logical and quantitative understanding of what data summaries mean.”

“The  $p$ -value was never intended to be a substitute for scientific reasoning,” said Ron Wasserstein, the ASA’s executive director. “Well-reasoned statistical arguments contain much more than the value of a single number and whether that number exceeds an arbitrary threshold. The ASA statement is intended to steer research into a ‘post  $p < 0.05$  era.’”

“Over time it appears the  $p$ -value has become a gatekeeper for whether work is publishable, at least in some fields,” said Jessica Utts, ASA president. “This apparent editorial bias leads to the ‘file-drawer effect,’ in which research with statistically significant outcomes are much more likely to get published, while other work that might well be just as important scientifically is never seen in print. It also leads to practices called by such names as ‘ $p$ -hacking’ and ‘data dredging’ that emphasize the search for small  $p$ -values over other statistical and scientific reasoning.”

The statement’s six principles, many of which address misconceptions and misuse of the  $p$ -value, are the following:

1. *P-values can indicate how incompatible the data are with a specified statistical model.*
2. *P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.*
3. *Scientific conclusions and business or policy decisions should not be based only on whether a  $p$ -value passes a specific threshold.*

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

4. *Proper inference requires full reporting and transparency.*
5. *A  $p$ -value, or statistical significance, does not measure the size of an effect or the importance of a result.*
6. *By itself, a  $p$ -value does not provide a good measure of evidence regarding a model or hypothesis.*

The statement has short paragraphs elaborating on each principle.

In light of misuses of and misconceptions concerning  $p$ -values, the statement notes that statisticians often supplement or even replace  $p$ -values with other approaches. These include methods “that emphasize estimation over testing such as confidence, credibility, or prediction intervals; Bayesian methods; alternative measures of evidence such as likelihood ratios or Bayes factors; and other approaches such as decision-theoretic modeling and false discovery rates.”

“The contents of the ASA statement and the reasoning behind it are not new—statisticians and other scientists have been writing on the topic for decades,” Utts said. “But this is the first time that the community of statisticians, as represented by the ASA Board of Directors, has issued a statement to address these issues.”

“The issues involved in statistical inference are difficult because inference itself is challenging,” Wasserstein said. He noted that more than a dozen discussion papers are being published in the ASA journal *The American Statistician* with the statement to provide more perspective on this broad and complex topic. “What we hope will follow is a broad discussion across the scientific community that leads to a more nuanced approach to interpreting, communicating, and using the results of statistical methods in research.”

#### *About the American Statistical Association*

The ASA is the world’s largest community of statisticians and the oldest continuously operating professional science society in the United States. Its members serve in industry, government and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare. For additional information, please visit the ASA website at [www.amstat.org](http://www.amstat.org).

For more information:

Ron  
Wasserstein

#### **Citations**

Macnish, K., Ryan, M. & Stahl, B. (2019). Understanding ethics and human rights in smart information systems: A multi-case study approach. *The Orbit Journal*, 2(2), 1–34.

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science*, 349(6251). <https://science.sciencemag.org/content/349/6251/aac4716>

*Adopted by the Market Information Systems Research and Development (D) Working Group, Oct. 14, 2021*

*Adopted by the Market Information Systems (D) Task Force, June 16, 2022*

Wasserstein, R.L., & Lazar, N.A. (2016). The ASA statement on p-values: Context, process and purpose. *The American Statistician*, 70(2), 129–133.

Draft: 3/22/22

*Adopted by Market Regulation and Consumer Affairs (D) Committee, \_\_\_\_, \_\_\_\_  
Adopted by Producer Licensing (D) Task Force, May 5, 2022*

***DRAFT FOR DISCUSSION***  
**GUIDELINES FOR AMENDING THE UNIFORM LICENSING APPLICATIONS**

The mission of the Producer Licensing (D) Task Force includes the development and implementation of uniform standards with a primary emphasis on encouraging the use of electronic technology. As part of this mission, the Task Force has appointed a Producer Licensing Uniformity (D) Working Group to “review and update, as needed, the NAIC’s uniform producer licensing applications and uniform appointment form.” In support of this mission and charge, the Producer Licensing (D) Task Force recognizes the importance of having stable, streamlined, and consistent NAIC’s Uniform Producer Licensing Applications, which comply with the statutes and regulations of the NAIC Membership and encourage the use of electronic technology in the most efficient manner.

In support of this mission and the importance of maintaining stable and consistent NAIC Uniform Licensing Applications, the Producer Licensing (D) Task Force will use the following guidelines for substantive changes to the NAIC’s Uniform Licensing Applications.

1. On a biennial basis, the Producer Licensing (D) Task Force will send an email communication by Dec. 1, to members of the Producer Licensing (D) Task Force, interested regulators, interested parties, and state producer licensing directors, asking for proposed changes to the NAIC Uniform Licensing Applications. The requested changes are to be submitted as a Word document using the NAIC Uniform Licensing Application Change Request form. The form should be completed in its entirety, attached to an email message, and directed and submitted to the NAIC staff providing primary support for the Producer Licensing (D) Task Force. All requests should be submitted by Feb. 1.
2. If the Producer Licensing (D) Task Force recommends further analysis of the request, the Task Force will assign the request to the Producer Licensing Uniformity (D) Working Group by the close of the NAIC Spring National Meeting. The Working Group will review the request using the following guiding questions:
  - a. Does the proposed change maintain the NAIC Membership’s mission of uniform licensing standards with a primary emphasis on encouraging the use of electronic technology?
  - b. Does the proposed change serve the regulatory purpose of strengthening consumer protection while maintaining an efficient licensing process for producer applicants? This should include documentation on why the existing Uniform Applications do not meet these objectives.
  - c. Does the proposed change comply with the statutes and regulations of the NAIC Membership and encourage the use of the NAIC’s Uniform Applications in all jurisdictions?
3. The initial comment period on exposure drafts issued by the Producer Licensing Uniformity (D) Working Group should be 30 calendar days. The Working Group may consider additional exposure periods of less than 30 days for revisions to the same draft.
4. Revisions to the NAIC’s Uniform Applications should be adopted by the Producer Licensing Uniformity (D) Working Group and the Producer Licensing (D) Task Force by the close of the NAIC Summer National Meeting.<sup>1</sup>
5. If the Producer Licensing Uniformity (D) Working Group recommends a requested change not be pursued, the request will be updated with that decision and filed for future reference. A copy of the recommendation and decision will be provided to the requestor.
6. If the Producer Licensing Uniformity (D) Working Group recommends proceeding with a requested change, NAIC staff providing primary support for the Producer Licensing (D) Task Force will coordinate with NIPR and States, including back-office system support vendors, during the next 45 days, conduct an analysis culminating in the provision of a time

---

<sup>1</sup>The dates and meetings set forth herein pertain only to the year in which the Producer Licensing (D) Task Force solicits proposed changes to the Uniform Licensing Applications as described in item 1.

and cost estimate for the Producer Licensing (D) Task Force's review and prioritization. Using staff analysis, the Producer Licensing (D) Task Force will identify an appropriate implementation date.

7. Revisions to the NAIC's Uniform Applications should be adopted by the Market Regulation and Consumer Affairs (D) Committee by Oct. 15, and the Executive Committee and Plenary by the conclusion of the NAIC Fall National Meeting.

## NAIC Uniform Application Change Request

**Date Submitted:** \_\_\_\_\_

**Name:** \_\_\_\_\_

**State:** \_\_\_\_\_

**E-Mail:** \_\_\_\_\_

**Phone:** \_\_\_\_\_

**Change Request to Following NAIC Uniform Application (Check all that apply)**

- Uniform Application for Individual License/Registration
- Uniform Application for Individual License Renewal/Continuation
- Uniform Application for Business Entity Licensing Registration
- Uniform Application for Business Entity License Renewal/Continuation

**Provide Concise Description of Proposed Change**

**Provide Reason for the Proposed Change**

**Provide Supporting Information Related to the Proposed Change**

**To Be Completed by NAIC Staff**

Change Request ID #	Date Received	Estimated Hours	Recommendation