FEBRUARY 2026

FACTOR ANALYSIS

for Health-Related Sensitive Personal Information

NAI

PRIVACY, TRUST & ACCOUNTABILITY

Table of Contents

Introduction

Responsible health-related advertising creates informational benefits for consumers; but the health privacy landscape is becoming increasingly complicated.

Access to health-related information plays an important role in empowering consumers to understand and manage their health and well-being.[1] In a digital environment where people increasingly seek guidance about symptoms, treatments, and wellness options online, advertising can serve as one additional pathway for consumers to discover relevant resources. It can do so by helping to increase consumer awareness of health resources and options by delivering more relevant messages to interested audiences.[2] This, in turn, can help broaden consumer awareness of health topics, conditions, and care options that may be relevant to them. Health-related advertising can also contribute to a more informed public by enabling health organizations to disseminate educational messages at scale, such as public health campaigns, information about preventive care resources, and information about emerging health technologies.[3] Data-driven advertising can help disseminate timely, relevant public health information and resources, promoting earlier awareness and supporting consumers' ability to make informed health decisions.[4]

At the same time, it is imperative to protect consumer privacy when conducting health-related advertising. This imperative is recognized by a range of U.S. policies that classify certain personal information (PI) related to consumer health as "sensitive" and impose heightened requirements on the processing of this data, such as opt-in consent,[5] enhanced disclosures,[6] and risk assessments.[7] Together, this set of privacy laws,[8] regulations,[9] and enforcement actions[10] comprises a U.S. policy framework that is inconsistent and still evolving. **As a matter of terminology only, the NAI refers to these data types collectively as "health-related sensitive personal information" or HSPI. In doing so, the NAI is not providing its own, separate definition of health data it considers sensitive.[11] Instead, the use of the term HSPI in this factor analysis ("Factor Analysis") is a generalized term that relies on underlying state law and other relevant definitions for its specific meaning in context.**

The current maze of different approaches to classifying HSPI in the U.S. has challenged the ability of organizations to conduct health-related advertising. In the face of this complexity, there are a range of unintended consequences that the NAI has observed: (1) organizations attempting to force simplicity by over-classifying *any* information related to health or the human body as HSPI and therefore subjecting the data processing to heightened requirements; (2) the risk of organizations under-classifying what constitutes HSPI, or failing to recognize in certain cases where a broader category of data could meet various legal definitions; (3) organizations in some cases withdrawing altogether from jurisdictions where they perceive a risk that almost any data processing could trigger broad definitions of HSPI. Together, these trends are making it more challenging to realize the many benefits of responsible health-related advertising and health-related content supported by ads.

The goal of the NAI in providing this Factor Analysis is to prevent the unintended consequences noted above by helping companies and policymakers identify when information being processed to facilitate health-related advertising may, or may not, be sensitive, while retaining room for necessary nuance in making those classifications. Providing tools like this Factor Analysis to the marketplace enables the benefits of health-related advertising while helping to protect consumer

privacy. Assessing the sensitivity of health-related PI often requires nuanced analysis, particularly in close cases where different considerations are in tension and may cut against one another. Using this Factor Analysis can help make those nuanced considerations visible and explicit.

Specifically, this Factor Analysis is a tool for identifying PI that may implicate one of various definitions of "sensitive" information related to health in U.S. privacy laws. It works by singling out different factors that are present in those various definitions of sensitive data. This in turn enables companies to consider those factors individually and together as they deliberate about classifying, or not classifying, certain health-related PI as sensitive.

This document is divided into four sections:

- Section I explains the relationship of this Factor Analysis to existing legal and self-regulatory frameworks.
- Section II provides background and context on how HSPI is treated under various legal authorities in the United States.
- Section III introduces and explains the Factor Analysis, setting out five factors to consider for classifying HSPI and guidance on how to apply the factors.
- Section IV sets out hypothetical scenarios that illustrate the application of the five identified factors to specific use cases to help determine whether to classify PI being processed in those use cases as HSPI.

This Factor Analysis is provided for informational purposes only and does not constitute legal advice. Organizations should consult with qualified legal counsel regarding their specific data processing activities and compliance obligations.

I. Relationship of this Factor Analysis to U.S. legal and self-regulatory frameworks

A. Relationship to state and federal laws and regulations

This Factor Analysis sets out key factors to consider in assessing whether to classify personal information as HSPI. However, in doing so, the NAI is not presupposing or providing a determination as to whether any particular PI meets the definitional requirements of HSPI under a given consumer privacy law. Neither is it endorsing or challenging any definition of sensitive personal information or treatment by a regulator of certain PI as more sensitive. As discussed in more detail in section III below, differences in the types of data a company processes, how they process and use the data, what jurisdiction(s) they operate in, and other factors are expected to result in different conclusions as to whether personal information a company is processing also qualifies as HSPI.

B. Relationship to the NAI Framework

For decades, the Network Advertising Initiative (NAI) has promoted strong privacy practices among member companies engaged in data-driven digital advertising, particularly with regard to PI related to consumer health.[12] Consistent with this history, the NAI's current Self-Regulatory Framework ("Framework") binds NAI member companies to foundational privacy principles, including a Sensitive Personal Information (SPI) principle requiring additional safeguards for SPI.[13]

The Framework empowers the NAI to develop "non-binding privacy best practices, guidance, and tools" under the Framework that are "designed to assist NAI member companies in understanding and complying with their privacy obligations under U.S. law."[14] The NAI considers this Factor Analysis to be a **tool**, which NAI members and others in the industry are free to use to support their classification of HSPI and to drive consensus with partners regarding such classification.[15] Under the Framework, this tool may also be used to help NAI members to meet their self-regulatory obligation to assess whether PI they process may be sensitive in order to determine whether additional safeguards are called for.[16]

II. Background on the treatment of HSPI under U.S. privacy laws

State and federal jurisdictions in the U.S. vary widely in how they define HSPI and what factors they rely on in their definitions. These factors include the context of collection and processing, the type of data being processed, and how the data are used. However, as a threshold matter, U.S. privacy laws treat HSPI as a subset of PI. Accordingly, data must first qualify as PI before it can qualify as HSPI. In general, this means the data at issue must be linked or reasonably linkable to an identified or identifiable consumer.[17] Conversely, data that is not and cannot reasonably be linked to an identifiable consumer is generally considered to be de-identified data.[18] Companies often seek to de-identify HSPI before using it for advertising, analytics, and other purposes, since the resulting de-identified data would generally fall outside the scope of these laws.[19]

A. Treatment of HSPI under federal law

At the federal level, the Health Insurance Portability and Accountability Act (HIPAA)[20] regulates individually identifiable health information that is created or received by certain health-related entities, such as health care providers or health plans, when that information relates to an individual's health condition, care, or payment for care. In other words, HIPAA's protections are triggered primarily by the context of collection and handling of information, rather than the content of the information alone. Health information about an individual becomes Protected Health Information (PHI) only when it is maintained or transmitted *by a HIPAA-covered entity* or its business associate in connection with health care functions.[21]

This contextual limitation in HIPAA is a critical distinction, as it means that health-related data collected by entities *outside* of the health care system generally does not constitute PHI under HIPAA. For these entities, the classification of data as HSPI and the corresponding privacy obligations are instead governed, if at all, by other federal laws like the Federal Trade Commission (FTC) Act.[22]

Notably, outside the HIPAA context, neither the FTC Act nor other federal statutes provide a clear definition of HSPI similar to those definitions found under state privacy law. This absence of a clear definition has led to an unclear and inconsistent set of regulations and enforcement for HSPI. For example, the FTC has invoked its authority under section 5 of the FTC Act on numerous occasions to characterize certain health-related information as sensitive, even though the FTC Act does not define such information as "sensitive," and the information would not qualify as PHI under HIPAA. Recent enforcement actions in this vein have relied on "deception" claims sometimes in conjunction with more novel "unfairness" theories under the FTC Act.[23] The FTC also modified the Health Breach Notification Rule ("HBNR") in 2024, clarifying its applicability to

health apps and other similar technologies and expanding the information that covered entities must provide to consumers when notifying them of a breach of their health data.[24]

It is important to keep in mind, however, that both the HBNR and more expansive interpretation of the FTC's "unfairness" authority have been met with significant dissent from within the Commission, and it remains to be seen what course the current Commission will take on enforcement and policy. This represents significant legal uncertainty, as the scope of FTC unfairness authority over health data for entities outside the HIPAA-covered healthcare system is actively contested. Specifically, current FTC Chair Ferguson has argued that the FTC's unfairness authority should focus on the unlawful collection of data on the "front end" rather than regulating the "amorphous, backend" analysis or inferences drawn from that data[25] – and noting that "merely drawing a conclusion from lawfully obtained data does not violate Section 5."[26] In addition, Ferguson has joined in expressing skepticism about the Commission's ability to broadly define categories of sensitive data under the FTC Act, arguing that such definitions are the province of Congress, not the FTC.[27] Still, prior enforcement actions have shown the FTC's ability to consider factors such as the type of data and potential harm to consumers to determine whether information qualifies as sensitive, and to apply a more expansive interpretation of its Section 5 unfairness authority if it so chooses.

Taken together, HIPAA anchors health-data sensitivity in the *context of collection for certain entities* (*i.e.* a HIPAA-covered entity or its business associate in connection with health care functions), while the FTC Act has the potential to be applied in a way that considers the *nature of the data* and *potential consumer harm*, with less emphasis on who collects it and in what context. Another emerging theme under federal law to watch is the classification of certain consumer information as sensitive in the national security context – including HSPI.[28]

B. Treatment of HSPI under state law

State consumer privacy laws vary widely in how they define sensitive PI related to health, yet most definitions still share two core elements. The first core element we characterize as the **relationship test**, which asks how the PI relates to a consumer's health (for example, whether it "reveals," "identifies," or "concerns" something about a consumer's health). The second core element we characterize as the **threshold test**, which specifies what kinds of health-related information meet a threshold to be considered sensitive (for example, whether the information pertains to a health diagnosis, treatment, or status). Putting the two elements together, one example formulation in state privacy laws states that HSPI is: (1) personal data **revealing** (relationship) (2) **mental or physical health diagnosis** (threshold).[29] We discuss different variations in these two elements below.

For the **relationship test**, multiple states prescribe that the PI at issue "**reveal**" something about a consumer's health to be considered HSPI.[30] The term "reveal" suggests that to pass the relationship test, the data at issue must itself convey or reflect some objective fact about the consumer's health. Accordingly, a company's collection or use of data is unlikely to "reveal" a fact about a consumer if the information is not accurate, or not known to be accurate, by the company.[31] For example, placing consumers aged 65+ in a segment to receive ads about heart disease medications is likely to include many consumers that do not have heart disease, as well as some consumers that may have heart disease. Placing this class of consumers in that segment based on a demographic factor does not appear to "reveal" that any particular consumer in the

segment has heart disease, even though the ads may be more relevant to that population segment as a whole. On the other hand, if a consumer were to self-report on a survey that they have heart disease, that information on its own is more likely to directly reveal a fact about that consumer's health (*i.e.*, that the consumer has heart disease).

The "reveal" version of the relationship test contrasts with formulations of the test used by other states that focus more on how data are used by a company, instead of what they may inherently reveal about a consumer. In Connecticut, for example, the PI must be "**used to identify**" a consumer's health status to qualify as HSPI.[32] This standard appears to speak more to how a company uses data (*i.e.*, whether data is **used** to associate a consumer with a health condition), even if the consumer does not actually have that condition and hence the condition itself could not be "revealed." In this case, placing a consumer in a segment labeled "Heart Disease" apparently could meet the Connecticut standard, even if the company is incorrectly guessing that the consumer has heart disease, because the segment is **used to identify** the consumer with a health condition. Similarly, in California, PI must be "**collected and analyzed**" and "concern" the consumer's health to pass the relationship test – but this does not appear to require that it reveal any objective fact about the consumer's health.[33] Some states have placed even further emphasis on how data are used by explicitly speaking to inferences made about a consumer's health based on non-health data, even if such inferences may be uncertain or incorrect.[34] A minority of states, such as Colorado and Washington, appear to have addressed this head-on by specifically indicating that health-related inferences drawn from PI are HSPI.[35]

States also vary with respect to the **threshold test**—even when they use the same formulation of the relationship test. For example, some states that are alike in using the "reveal" formulation of the relationship test point to different types of PI relating to a consumer's health that pass the threshold test. States with a stricter threshold test require that PI reveal a **diagnosis by a health care professional**, while states with a lower threshold allow that PI revealing a **health condition, treatment, or diagnosis** may be HSPI, even if it is not professionally verified.[36] States like Washington have relatively low thresholds for what types of data could be considered sensitive and include a long list of examples that may qualify as identifying a consumer's **past, present, or future physical or mental health status.**[37]

Taken together, the varied definitions of HSPI across state and federal laws differ not only in terminology but in the factors they emphasize—from context of collection, to use and purpose, to the degree of connection with consumer health. These differences make it challenging for companies, particularly in digital advertising, to develop uniform classification and compliance strategies. However, they also include recurring themes that provide the legal foundation for the five factors set out in the next section.

III. Factor Analysis

A. Scoping

What this factor analysis is intended to do:

The widely varied formulations used to classify HSPI under privacy laws in the U.S. have created practical challenges for companies conducting health-related advertising. The Factor Analysis set out below is intended to help companies navigate these challenges by providing discrete factors to consider when evaluating when PI being processed may qualify as HSPI. Processing of PI that implicates multiple factors likely deserves further scrutiny, while processing that does not implicate any of the identified factors is less likely to qualify as HSPI.

What this Factor Analysis does not do:

Applying this Factor Analysis does not necessitate a classification of PI as HSPI under any particular state or federal law. Although the implication of a factor may call for further scrutiny, the implication of any one factor alone does not force a conclusion without further analysis. Further, regulators and courts may reach different conclusions than companies do when applying these factors to similar fact patterns, particularly given ongoing legal uncertainty and jurisdictional variation in how HSPI is defined. Finally, while this Factor Analysis provides practical guidance for approaching data classification, it is not a substitute for qualified legal advice as it relates to the application of any specific legal definition to a company's data processing activities.

When to apply this Factor Analysis:

In some cases, companies may already have a high degree of confidence that the PI they are processing has no nexus to consumer health or the human body and is therefore not HSPI. In those circumstances, further detailed evaluation of that processing using this Factor Analysis may not be necessary. In addition, companies may have existing policies and review processes in place for determining whether PI they are processing is sensitive and may have already conducted detailed review and analysis to make those classifications. In those circumstances, additional review and application of this Factor Analysis may not be efficient.

However, where companies have not already implemented a review process for relevant PI to identify and classify it as HSPI (or not HSPI), this Factor Analysis provides a model for how to do so. In addition, this Factor Analysis may serve to supplement existing review processes for use cases that present novel or nuanced issues where analysis along different factors can help companies arrive at sound classifications. It is not intended to dictate how companies must classify data, but to provide a structured way to reason through the relevant considerations.

B. Five factors relevant to HSPI

The different treatments of HSPI under U.S. state and federal legal frameworks discussed above present a number of different criteria that go toward whether health-related PI is sensitive. Considering the different formulations of HSPI used in the U.S., this Factor Analysis distills five key factors companies can use to help evaluate whether certain PI they are processing, or combinations thereof, might align with characteristics that regulators have associated with HSPI. Those factors are:

(1) the source of the PI being processed (*e.g.*, whether from a health care provider);
(2) the contents of the PI;
(3) the intended use of the PI;
(4) heightened consumer expectations of privacy; and
(5) whether there is a heightened risk that processing the PI could result in consumer harm, considering both the *severity* and *likelihood* of harm.

It is important to consider all factors when conducting this analysis as in many cases no single factor is dispositive.

Factor 1: The source of the PI being processed

Review the source of the PI to determine whether it is associated with services relating to health**.** For example, this factor is more likely to be implicated by PI collected from certain sources, such as health records, medical practitioners, medical insurance claims, and health-related mobile applications.

Sometimes, the source of PI alone can significantly influence how it is treated under privacy laws. Specific **sources** of PI about consumers that may implicate this factor include medical records from health care providers[38] or information resources or tools specific to a consumer's health (*e.g.*, apps intended for tracking reproductive health such as menstrual cycle or pregnancy status).

On the other hand, the same *type* of data may be collected from different sources, resulting in different outcomes. For example, some PI related to a consumer's health may be provided directly by the consumer in a marketing survey, in which case the source of data is unlikely to implicate this factor, while the same type of data collected from a health record could implicate this factor.

Factor 2: The contents of the PI being processed

Review the contents of the PI being used to determine whether the data has a direct connection to health. In some cases, the contents of the PI on its own—regardless of the context of its collection or how it may be used—is suggestive of information about a consumer's health and is therefore more likely to implicate this factor.

For example, if a company is processing PI about a consumer's specific health condition, diagnosis, or treatment, this factor is implicated. While this type of information is often collected in a health care context (implicating Factor 1 above), a company might collect this type of information outside of a health care context, such as through a survey or questionnaire for market research. Consider

the following questions about the content of the PI you are processing to help determine if the **type** of data weighs in favor of it being HSPI.

a. Does the PI relate to a treatment for a specific health condition?

Query whether the PI relates to treatments, medications, prescriptions, or therapeutics that are specific to an identified health condition. Notably, PI identifying a treatment **prescribed** to treat an illness, injury, or other specific health condition more strongly implicates this factor compared to PI relating to over-the-counter medication.[39] For example, PI associating a consumer with the purchase or use of over-the-counter medications such as pain relievers may relate to certain general symptoms like aches or pains without revealing a specific health condition and, therefore, is less likely to implicate this factor. On the other hand, PI associating a consumer with a condition-specific prescription treatment is more likely to implicate this factor.

b. Does the PI relate to a specific health condition or diagnosis?

PI indicating that a consumer has a specific health condition or diagnosis would likely implicate this factor. The strength of this factor could range from data *directly* revealing a health condition or diagnosis, such as a diagnostic code in a health record, or data that *weakly or indirectly* relates to or indicates a health condition or diagnosis.

For example, in its 2025 *Healthline* settlement, the California attorney general's office scrutinized the processing of PI indicating that a consumer viewed certain health-related article titles, such as "Newly Diagnosed with Ulcerative Colitis? Here's What to Know," and expressed concern about what the article title might reveal about a consumer's health.[40] Even though reading an article does not *directly* reveal a health condition (and the individual accessing the article may not have the condition at all and may be reading the article for informational or other purposes), the focus on this issue in California militates in favor of weighing this type of indirect information as a factor, even when it is only weakly or indirectly related to an actual consumer health diagnosis.

Further, how heavily to weigh information such as a suggestive article title also depends highly upon the specific facts and circumstances of the surrounding data processing. The article titles cited in *Healthline* referred to diagnosis with specific conditions, while many article titles related to specific conditions do not refer to a consumer's diagnosis with that condition. The NAI recognizes that distinguishing between that level of nuance in article titles (*e.g.*, articles referring to those 'Newly Diagnosed' with a condition vs. 'General Trends' in diagnosis) at scale presents significant operational challenges. In those circumstances when the content of such data can be ambiguous, the **intended use (Factor 3)** of the data becomes even more salient. For example, from an intended use perspective, processing article titles solely for real-time contextual advertising (where the data is used to match an ad to the page content immediately and not retained to build a user profile) is less likely to rise to the level of HSPI compared with using the same article titles to segment or profile a consumer as having a specific condition, particularly where there is ambiguity about the specific article title at issue beyond certain keywords. Further, even if an article title is suggestive of diagnosis with a specific condition, the transient nature of processing for contextual advertising—where data is discarded immediately after the ad decision—significantly reduces the privacy impact compared to profiling.

In weighing this factor, also note that there was no allegation in *Healthline* that the article titles at issue were "sensitive personal information" under the CCPA; instead, the California attorney general's office advanced the point of view that those specific articles may carry heightened expectations of privacy for consumers reading them (see also Factor 4 below).[41]

c. *Does the PI include measurements of a vital sign or relate to a specific bodily function?*

Although these do not necessarily reveal a specific health condition, vital signs or measurements may directly relate to an individual's health and are likely to implicate this factor. Examples may include PI about an individual's heart rate, blood pressure, menstrual cycle, or weight. On the other hand, some PI, although related to health, is more general in nature and is less suggestive of specific bodily functions, such as information about overall fitness level or diet.

Factor 3: The intended use of the PI

Evaluate the intended use of the PI being processed. Even PI that is not otherwise related to health may implicate this factor if it is used to impute a health condition to an individual consumer.

Various non-health-related PI may be processed together to predict or infer a consumer's health status or diagnosis, or to identify a consumer with a condition. As such, also evaluate and consider the volume of data being processed about any individual consumer. Larger datasets may be apt for inferring information about a consumer's health, even if isolated datapoints in that set would not be.[42]

How data is used in practice determines whether this factor is implicated. For example, a company processing product purchase data to retarget ads and measure conversions in a content-neutral manner is unlikely to implicate this factor. However, if the same company were to assign a "pregnancy prediction score" to consumers using the same product purchase data, that particular use of the purchase data and accompanying inferences would likely implicate this factor.[43] By contrast, relying solely on demographic factors (as opposed to individual behavioral history) such as age and gender does not implicate this Factor 3 in the same way. As detailed in the NAI's *Demographic Health Advertising Best Practices*,[44] there is a critical distinction between targeting based on general demographic correlations (*e.g.*, age and gender as population-level markers for relevance) and targeting based on individualized health profiling. While a prediction score attempts to determine if a *specific individual* has a condition, a demographic segment relies on population-level public health correlations to reach a broader audience without ascribing a health status to any single user. Consequently, the use of broad demographic proxies generally avoids implicating Factor 3, unless a company *actually uses* those proxies to impute a health condition to an individual consumer.

Similarly, as noted above in Factor 2, an ad-tech company may collect and process article titles from webpage URLs that relate to health topics in order to serve contextual advertising and frequency cap those contextual ads without making any inference about the user visiting the URL; however if the same ad-tech company adds that user to a segment that identifies them with a specific health condition, that activity is more likely to implicate this factor. The NAI's prior guidance on refraining from placing users, without opt-in consent, in segments that identify them with sensitive health conditions like cancer, a mental health condition, or an STD continues to be apt for assessing this factor with respect to interest segments.[45]

Factor 4: Whether consumers have a heightened expectation of privacy

Evaluate the social and contextual factors that may shape an individual's expectations of privacy. In some cases, consumers may have a heightened expectation of privacy with respect to information related to health.[46] This may especially be the case for reproductive or sexual health, mental health, or particularly serious health issues or conditions.[47] If a consumer is likely to have a heightened expectation of privacy with respect to PI related to those topics, that would implicate this factor.

In addition to social and contextual factors, also consider any disclosures presented to the consumer regarding personal data processing, their level of detail, and their prominence, which may affect consumer expectations.[48] In this context, consider what may be material to an average consumer's understanding – for example, disclosure of primary *purposes* of the processing (relevant advertising) rather than a detailed description of complex technical mechanisms that may be involved. Companies may consider whether a reasonable person would view the data as sensitive, considering context and consumer understanding. The other factors detailed in this analysis may also aid companies in evaluating consumer expectations of privacy (*e.g.* the source and context of collection may heavily influence consumer expectations of privacy).

Factor 5: The risk of consumer harm

Consider whether the processing of health-related PI increases the likelihood of consumer harm, and the severity of that harm. Harm may come in different forms. Some are more objective and severe, such as unlawful discrimination, economic harm, or physical harm.[49] For example, if PI related to a consumer's health is processed in a way that could impact an eligibility determination for a consumer, by raising costs or preventing the consumer from obtaining a benefit like health coverage or employment, this could rise to the level of an economic harm or an instance of unlawful discrimination based on a protected health status, such as a disability.[50] Other potential harms are more subjective, such as embarrassment. Some authorities have even suggested the collection alone of some forms of PI may be harmful,[51] but this point of view has not been widely adopted and even argued against by other authorities.[52]

In addition, companies may weigh the risk and severity of harm to the consumer against the benefits the consumer may receive from the processing.[53] Within the context of advertising, the benefit to the consumer is often the information about products and services the consumer finds useful. In some cases, these benefits can be considerable, such as informing consumers about a product or service that can improve or protect their health.

IV. Hypothetical scenarios

The following examples illustrate the application of this Factor Analysis to help determine whether the PI being processed in each scenario might be HSPI. We review three hypothetical scenarios:

> (1) a retargeting campaign for running shoes relying on add-to-cart events to identify consumers that may be interested in purchasing running shoes;
> (2) an ad campaign for newborn diapers relying on "pregnancy prediction scores" extrapolated from purchase data; and
> (3) a campaign for a new diabetes medication relying on HIPAA de-identified insurance claims data to model the demographic characteristics that are most common for diabetics in the relevant population.

A. Relationship of hypotheticals to legal frameworks defining HSPI

The following hypothetical scenarios are designed to demonstrate how the factors described above can be applied to practical data-processing situations. They illustrate an analytical process rather than the interpretation of any law, though the approach reflects common themes found across U.S. privacy frameworks. However, these examples are purely illustrative and do not represent NAI's conclusion as to whether any specific data set or processing activity constitutes HSPI under any legal definition.

B. Hypotheticals

Hypothetical # 1

A sportswear retailer uses an ad-tech provider to run a retargeting campaign for running shoes. The retailer markets products to a general audience without catering to a specific health condition or diagnosis of the consumer. When a consumer visits a running shoe product page and clicks "Add to Cart," the campaign records this event and identifies the consumer as interested in purchasing running shoes. The campaign uses cookies to retarget the consumer for a period of time and records a purchase of running shoes as a conversion event. Both the retailer's privacy policy and the ad-tech company's privacy policy include disclosures describing how they collect PI from online activity to provide advertisements more likely to be relevant to a consumer's interests.

Application of Factor Analysis:

Factor 1: The source of the data being used. The source of the PI is a general online sports retailer serving a consumer in the market for running shoes. Although data collected from a general online sports retailer does not implicate this factor, note that if the retailer instead specialized in selling orthopedic shoes to treat specific podiatric conditions, that context would be important to consider in analyzing this factor. Here, the retailer markets sporting goods products without considering the health condition or diagnosis of the consumers visiting the website. As such, Factor 1 is not implicated by the processing activity.

Factor 2: The contents of the data being used. The contents of the PI collected are a specific cookie ID associated with a consumer adding running shoes to a shopping cart, suggesting intent to

purchase. The purchase of running shoes is not directly related to a specific health condition, treatment, or diagnosis. Neither do running shoes indicate any specific vital sign or relate to a specific bodily function. Instead, they at most indicate an interest in a general fitness or recreational activity like running.[54] As such, Factor 2 is not implicated by the processing activity.

Factor 3: The intended use of the data. Here, a single "add to cart" event is used to identify consumers that may be interested in purchasing running shoes and add them to an audience eligible to receive ads for those shoes. This process does not involve the generation of any additional inferences about the consumer's characteristics or interests, even at a general level. In other words, the audience segmentation is content-neutral. This is distinct from other scenarios where the consumer's action of viewing running shoes might cause them to be added to an interest segment such as "running enthusiast" which are more likely to be viewed as making an inference about the consumer (regardless as to whether that inference is related to the consumer's health, which should be evaluated under Factor 2). As such, Factor 3 is not implicated by the processing activity.

Factor 4: Heightened consumer expectations of privacy. Interactions with an online general sports retailer are less likely to have a heightened expectation of privacy compared to certain other settings, such as interactions with a health care provider, where expectations of privacy may be enhanced. In addition, the type of data being processed (shopping events for running shoes) does not relate to a health condition or status that carries heightened sensitivity, such as sexual or mental health status. Finally, the inclusion in the relevant privacy policies of disclosures that the consumer's PI will be processed to provide advertisements reduces the likelihood that a consumer would expect heightened privacy protections when viewing a product page. As such, Factor 4 is not implicated by the processing activity.

Factor 5: The risk of consumer harm. A retargeting campaign for running shoes is unlikely to result in any concrete harm to a consumer. For example, it does not appear to create a risk of economic harm such as denial of a benefit; or discrimination against the consumer based on a protected health status. Further, considering potential benefits, a consumer that completes a purchase of running shoes after expressing interest in them based on an ad view benefits from having obtained the shoes they wanted. As such, Factor 5 is not implicated by the processing activity.

Conclusion: After considering each factor and taking them together, the ad-tech company concludes that processing "add to cart" and purchase event data for the purpose of a retargeting ad campaign for running shoes does not meaningfully implicate any of the five factors associated with HSPI.

<u>Hypothetical # 2</u>

A large, general-purpose retailer markets and sells a variety of consumer packaged goods and is seeking to run an advertising campaign to increase sales of baby diapers. The company's market research indicates that pregnant or expectant consumers are more likely to purchase the company's diapers. To identify prospective customers meeting those criteria to target with its ads, the company plans to use data gathered through its loyalty card program to assign a pregnancy prediction score to consumers. In doing so, it aims to identify those consumers who are more likely to be pregnant and, thus, may soon be in the market to purchase baby diapers. When consumers

sign up to join the loyalty program, they agree to terms of service describing how purchase data may be used for marketing purposes.

Factor 1: The source of the data being used. The source of the PI is a loyalty program of a large general retailer. Large, general-purpose retailers are not specifically associated with services relating to health. As such, Factor 1 is not implicated by the processing activity.

Factor 2: The contents of the data being used. The contents of the PI originally collected by the retailer consist of retail transaction information collected through a loyalty program. This includes information about purchase of items such as groceries and toiletries that are not directly related to a specific health condition, treatment, or bodily function. The contents of these data do not on their own appear likely to implicate this factor. However, considering that the company intends to use the contents of the PI collected through the loyalty program to infer the likelihood that consumers are pregnant by assigning them a pregnancy prediction score (*see also* Factor 3 below), the resulting inference that a consumer is likely to be pregnant *does* relate to pregnancy, a specific condition of the human body. The inference about pregnancy implicates Factor 2.

Factor 3: The intended use of the data. Here, general purchase information from consumers participating in the loyalty program is used to assign a pregnancy prediction score to identify consumers who are more likely to be pregnant. As PI may qualify as HSPI if it is used to infer information about a consumer's health, using non-sensitive PI to assign a pregnancy prediction implicates Factor 3.

Factor 4: Heightened consumer expectations of privacy. The data processing here includes a pregnancy prediction score, which relates to a consumer's reproductive or sexual health. The private and intimate nature of reproductive and sexual health is likely to increase a consumer's expectation of privacy for PI related to those topics. Even though consumers should expect that their purchase data may be used for advertising purposes as disclosed in the loyalty program's terms of service, this is weighed against the heightened expectation of privacy associated with reproductive health. As such, Factor 4 is likely implicated by this scenario.

Factor 5: The risk of consumer harm. Here, the retailer's use of loyalty program data, and even a pregnancy prediction score, is unlikely to result in concrete harm to a consumer. However, if such data were inadvertently released or misused, knowledge of a consumer's pregnancy status could lead to a risk of harm. For example, pregnancy is a protected class under some state and federal laws, which implicates a risk of unlawful discrimination. Further, some state laws may be implicated if a consumer's pregnancy status changes. Although a consumer might also benefit from loyalty card ads or offers related to diapers, the potential risks associated with misuse of a pregnancy prediction score implicate Factor 5.

Conclusion: After conducting this analysis, the company concludes that its use of general purchase information to infer a pregnancy status implicates several relevant factors and flags the use case for further consideration and legal counsel to determine whether it is HSPI.

Hypothetical # 3

A pharmaceutical company plans to run an ad campaign for a diabetes medication. The company enlists the help of an ad-tech partner to provide an audience for whom the medication is more

likely to be relevant. To assist the pharmaceutical company, the ad-tech company analyzes de-identified insurance claims data to better understand the demographic characteristics of the population who has already purchased the drug and sought reimbursement or payment through insurance. This analysis results in insights that the population that has already purchased the drug is more likely to be over age 65 and have an annual income <$25,000, making that demographic an audience that the drug is more likely to be relevant to.[55] The ad-tech company then obtains a set of targetable device IDs associated with consumers who match those demographic criteria from a data partner that lawfully obtained that information from public records and survey data and runs the ad campaign, showing ads for the diabetes medication to users of the identified devices.[56]

Factor 1: The source of the data being used. The source of the data used to model the targetable audience is de-identified insurance claims data. PI collected from records covered by HIPAA, such as insurance claims records, clearly implicate this factor because the source is closely tied to consumer health. However, as the data at issue here is robustly de-identified in accordance with applicable legal standards (*e.g.*, HIPAA expert determination and any applicable state-law requirements), it is not reasonably linkable to any particular consumer and, as such, would not be treated as PI (much less HSPI) under U.S. legal frameworks. As such, even though sourcing *PI* from insurance claims data would implicate this factor, because the data here is *not PI* (it is de-identified), it does not implicate this factor. However, the demographic data being used to create a targetable audience is not de-identified and instead associates particular consumers or devices with demographic characteristics (age 65+ and income <$25,000). However, because the demographic data were obtained from public records and general surveys – *sources* that do not have any nexus to consumer health – Factor 1 is not implicated for that dataset either (Note: The *use* of this demographic data is assessed separately under Factor 3).

Factor 2: The contents of the data being used. The specific data elements being processed to define the audience are age and income. On their face, these data points are generic and do not reveal any specific health condition, diagnosis, or treatment. Unlike *e.g.*, a prescription record, the *content* of generic demographic data points such as these is health-neutral. Therefore, the contents of the PI in this example do not implicate Factor 2. In considering this factor, the ad-tech company also notes that while the approximately 40% prevalence of diabetes in this demographic segment is significantly higher than the approximately 15% baseline for the general adult population,[57] it still represents a minority of the segment overall. However, if the prevalence of diabetes in the demographic segment were so high that membership in the segment becomes functionally equivalent to a diagnosis (*e.g.*, effectively treating the segment as a proxy for the condition), this could raise questions about whether membership in that segment by itself effectively *reveals* a consumer's health status. If that were the case, this factor *could* be implicated. However, at 40% prevalence, this factor is not implicated. (Note: While the content of the demographic data in this example is health-neutral, the intended *use* of the same data to build a health-related segment is assessed separately under Factor 3).

Factor 3: The intended use of the data. Factor 3 examines how data is actually used in practice. Here, the critical question is whether the ad-tech company uses the demographic data to make an individual-level inference that specific consumers have diabetes. If the company were to use the 40% prevalence rate to ascribe diabetes to individual members of the segment (*e.g.*, assigning a health attribute or probability score to each user), that use would implicate Factor 3. However, if the company instead relies only on the population-level observation that diabetes medication is statistically more relevant to this demographic without making individual health inferences, that

use does not implicate Factor 3. In this hypothetical, the ad-tech company does the latter—it uses demographic correlations to improve ad relevance without imputing health status to individual consumers.

More specifically, note that in this example the prevalence of diabetes in the demographically defined audience is assumed to be 40%, compared to roughly 15% of the general adult population. Under this assumption, the ad campaign is more than twice as likely to be relevant to a member of the demographically defined audience, which is a helpful result for an ad campaign and more likely to lead to discovery of a useful treatment for affected consumers. However, to achieve this increase in efficiency in the ad campaign, the ad-tech company does not impute a diabetes diagnosis to all (or any particular) members of the audience segment. Indeed, it would be unreasonable and unnecessary for the ad-tech company to infer that any one member of this segment has diabetes based on 40% population-level prevalence alone.

Still, there is technically nothing *preventing* the ad-tech company from making an inferential leap from an individual consumer's membership in a demographic segment with 40% diabetes prevalence to a likely inaccurate conclusion that the consumer actually has diabetes (60% inaccurate, to be exact). And if the ad-tech company were to *use* population-level prevalence to make this inference – however poor that inference is – then this factor *would* be implicated. In this hypothetical, though, the ad-tech company does not use the combination of prevalence statistics and consumer demographic data in that way, and, as such, its use of demographic data in this case does not implicate Factor 3.

Factor 4: Heightened consumer expectations of privacy. Consumers should not have a heightened expectation of privacy with respect to de-identified insurance claims data because such data are not linked or reasonably linkable to them. Indeed, whenever any form of PI is successfully de-identified, consumers should no longer have a heightened expectation of privacy for it because it is no longer *their* PI, or even PI at all. With respect to the demographic characteristics of age and income, consumers should not have any heightened expectation of privacy for those data because they are often publicly available. As such, the use of demographic data in this case does not implicate Factor 4.

Factor 5: The risk of consumer harm. The risk of harm to consumers with respect to de-identified insurance claims data is minimal because such data are not linked or reasonably linkable to any identified or identifiable consumer. With respect to the demographic characteristics, while age may be a protected characteristic in some contexts and create a risk of unlawful discrimination if misused, inclusion or exclusion from a demographically defined advertising audience is unlikely to present those risks. As such, the use of demographic data in this case does not implicate Factor 5.

Conclusion: After conducting this analysis, the company concludes that its use of de-identified insurance claims data along with demographic features of individual consumers to target ads for a diabetes medication does not meaningfully implicate any of the five factors associated with HSPI.

End Notes

[1] *See generally* Alma Taya and Ying-Chih Chuang, *Internet use for health information, health service utilization, and quality of care in the U.S.*, BMC Health Services Research (May 8, 2025), https://link.springer.com/article/10.1186/s12913-025-12807-5, (concluding that using the internet to obtain health information for discussions with health care providers has a positive impact on perceived care quality).

[2] *Cf.* Mia L. A. Lustria et al., *A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions*, 18 J. Health Commc'n 1039 (2013), https://www.tandfonline.com/doi/abs/10.1080/10810730.2013.768727; Seth M. Noar, Christina N. Benac & Melissa S. Harris, *Does Tailoring Matter? Meta-Analytic Review of Tailored Health Behavior Change Interventions*, 133 Psychol. Bull. 673 (2007), https://pubmed.ncbi.nlm.nih.gov/17592961/.

[3] *See generally* Kathryn J. Aikin et al., Dep't of Health & Hum. Servs., U.S. Food & Drug Admin. Ctr. for Drug Eval. and Res., *Patient and Physician Attitudes and Behaviors Associated With DTC Promotion of Prescription Drugs* at 31-32 (2004), https://www.fda.gov/files/drugs/published/Patient-and-Physician-Attitudes-and-Behaviors-Associated-With-DTC-Promotion-of-Prescription-Drugs-Final-Report.pdf (finding that direct-to-consumer advertising prompts consumers to ask physicians about previously undiagnosed conditions and empowers patients to have more informed treatment discussions). *See also Bates v. State Bar of Arizona*, 433 U.S. 350, 364 (1977) ("[C]ommercial speech serves to inform the public of the availability, nature, and prices of products and services, and thus performs an indispensable role in the allocation of resources in a free enterprise system."); *In re Cal. Dental Ass'n*, 121 F.T.C. 190, 296 (1996), *vacated on other grounds*, 526 U.S. 756 (1999), https://www.ftc.gov/sites/default/files/documents/commission_decision_volumes/volume-121/ftc_volume_decision_121_january_-_june_1996pages_291-378.pdf, ("We believe in the basic premise, as does the Supreme Court, that by providing information advertising serves predominantly to foster and sustain competition, facilitating consumers' efforts to identify the product or provider of their choice and lowering entry barriers for new competitors.").

[4] *See generally* Susan Athey et al., *Digital Public Health Interventions at Scale: The Impact of Social Media Advertising on Beliefs and Outcomes Related to COVID Vaccines*, 120 Proc. Nat'l Acad. Sci. e2208110120 (2023), https://www.pnas.org/doi/10.1073/pnas.2208110120; R. Michael Schwartz et al., *Long-term engagement in smoking cessation campaign: A mixed methods randomized trial*, PLOS ONE (2025), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0318160.

[5] *E.g.* Colo. Rev. Stat. § 6-1-1308(7) ("A controller shall not process a consumer's sensitive data without first obtaining the consumer's consent[.]"); Wash. Rev. Code § 19.373.030(1)(c) ("[T]he request for consent must clearly and conspicuously disclose: (i) The categories of consumer health data collected or shared; (ii) the purpose of the collection or sharing of the consumer health data, including the specific ways in which it will be used; (iii) the categories of entities with whom the consumer health data is shared; and (iv) how the consumer can withdraw consent from future collection or sharing of the consumer's health data.");

[6] *E.g.* Cal. Civ. Code § 1798.100(a)(2) (requiring businesses to inform consumers of the categories of sensitive personal information they collect and the purposes for which the categories of such information are collected or used, and whether that information is sold or shared).

[7] *E.g.* Conn. Gen. Stat. § 42-522(a) (2025) (requiring controllers to conduct data protection assessment for several processing activities including the processing of sensitive data).

[8] *E.g.* Cal. Civ. Code § 1798.140(ae)(2)(B) ("Sensitive personal information means… [p]ersonal information collected and analyzed concerning a consumer's health.") (quotations removed); Wash. Rev. Code § 19.373.010(8) ("Consumer health data means personal information that is linked or reasonably linkable to a consumer and that identifies the consumer's past, present, or future physical or mental health status.")

(quotations removed); Tex. Bus. & Com. Code § 541.001(29)(A) (defining "Sensitive data" to include "personal data revealing… mental or physical health diagnosis[.]").

[9] *E.g.* Colo. Code Regs. 904-3, Rule 2.02, (defining "Sensitive Data Inference" to include inferences extrapolated from non-sensitive PI to indicate a consumer's mental or physical health condition or diagnosis, sex life, or sexual orientation, amongst other sensitive attributes).

[10] *Cf.* Decision and Order, *In re BetterHelp, Inc.*, FTC Docket No. C-4796 at 6 (July 14, 2023) https://www.ftc.gov/system/files/ftc_gov/pdf/2023169betterhelpfinalorder.pdf (defining "Treatment Information" to include "individually identifiable information relating to the past, present, or future physical or mental health or condition(s) of a consumer[.]").

[11] Prior to the development of state privacy frameworks offering varying definitions for what types of health-related information they deem to be sensitive, the NAI Code of Conduct offered its own definition for "Sensitive Information" that included "[i]nformation, including inferences, about sensitive health or medical conditions or treatments[.]" Network Advert. Initiative, *2020 NAI Code of Conduct*, https://thenai.org/wp-content/uploads/2021/07/nai_code2020.pdf.

[12] The NAI's legacy Code of Conduct defined Sensitive Information to include a consumer's "sensitive health or medical conditions or treatments," and required opt-in consent for the use of such information for covered purposes. *See id*. In 2023, the NAI released its *Demographic Health Advertising Best Practices*, which sets out methods for creating demographic audience segments that are relevant to different health statuses without relying on health-related sensitive personal information. *See* Network Advert. Initiative, *2023 Demographic Health Advertising Best Practices*, https://thenai.org/wp-content/uploads/2023/11/NAI-Health-Targeting-Best-Practices-Document-Final.pdf.

[13] *See* Network Advert. Initiative, *NAI Self-Regulatory Framework* at 6 (hereinafter "NAI Framework"), https://www.thenai.org/wp-content/uploads/2025/03/NAI-Framework_March-2025.pdf. ("Each member company shall limit its processing of sensitive personal data to disclosed purposes, and purposes consented to by the consumer as required by applicable laws, and shall provide additional safeguards when processing such data.").

[14] *Id.* at 3.

[15] *Compare with* Network Advert. Initiative, *2023 Demographic Health Advertising Best Practices*, https://thenai.org/wp-content/uploads/2023/11/NAI-Health-Targeting-Best-Practices-Document-Final.pdf. (providing voluntary, but normative standards in the form of best practices for NAI Members creating and using demographic audience segments).

[16] The NAI reviews whether members have implemented a process that allows them to assess whether personal data is sensitive. *See* NAI Framework, *supra* note 13, at 8.

[17] Every state with a comprehensive privacy law defines PI as data that is "reasonably linkable" or could "reasonably be linked" to an identifiable individual. *See, e.g.*, Cal. Civ. Code § 1798.140(v)(1) ("Personal information means information that… is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household.") (quotations removed); Tex. Bus. & Com. Code § 541.001(19) ("Personal data means any information… that is linked or reasonably linkable to an identified or identifiable individual.").

[18] Every state with a comprehensive privacy law excludes data that has been de-identified from the definition of personal information. *See, e.g.*, Cal. Civ. Code § 1798.140(v)(3) ("Personal information does not

include consumer information that is deidentified[.]") (quotations removed); Colo. Rev. Stat. § 6-1-1303(17)(b) ("Personal data… [d]oes not include de-identified data[.]") (quotations removed).

[19] State laws substantially agree on the standard for de-identification with only slightly different formulations in some cases. *Compare* Cal. Civ. Code § 1798.140(m) (requiring that measures be taken to ensure the information cannot be associated with a consumer or household) *with* Colo. Rev. Stat. § 6-1-1303 (11) (requiring that measures be taken to ensure the information cannot be used to infer information about, or otherwise be linked to, an identified or identifiable individual, or a device linked to such as individual); *and with* Tex. Bus. & Com. Code § 541.001(12) (defined to be data that cannot reasonably be linked to an identified or identifiable individual, or a device linked to that individual). State approaches to defining "de-identified data" sometimes reference to a *household* or *device linked to an individual*, whereas the HIPAA de-identification standard is limited to the individual and whether that information can be used to identify the individual. *See* 45 C.F.R. § 164.514(a). Additionally, state definitions of "de-identified data" do not reflect certain allowances in the HIPAA De-identification Standard, which includes two specific recognized methods for de-identifying protected health information: the expert determination method and the safe harbor method. *See* 45 C.F.R. § 164.514(b).

[20] 42 U.S.C. § 1320d *et seq*.

[21] *See* 45 C.F.R. § 160.103 (defining "Health Information," "Individually Identifiable Health Information," and "Protected Health Information"); 45 C.F.R. § 164.502(a) (limiting HIPAA's use and disclosure requirements to covered entities and business associates). *See also* U.S. Dep't of Health & Human Servs., *Summary of the HIPAA Privacy Rule* (noting that "the Privacy Rule applies only to covered entities and their business associates, and not to other persons or organizations that may hold health-related information"), available at https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.

[22] 15 U.S.C. § 45 (prohibiting unfair or deceptive acts or practices in or affecting commerce).

[23] *See, e.g.*, Complaint, *In re BetterHelp, Inc.*, FTC Docket No. C-4796 (July 7, 2023) at 17, https://www.ftc.gov/system/files/ftc_gov/pdf/2023169betterhelpcomplaintfinal.pdf (alleging that the defendant's failure to obtain consumers' affirmative express consent before collecting, using, and disclosing to third parties those consumers' health information likely caused substantial injury to those consumers and, therefore, constitute unfair acts or practices in violation of Section 5 of the FTC Act); Compl. at 2, *United States v. GoodRx Holdings*, No. 23-cv-00460 (N.D. Cal. Feb. 1, 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/goodrx_complaint_for_permanent_injunction_civil_penalties_and_other_relief.pdf (alleging that it is a deceptive practice for an entity to promise consumers it will never share their health-related sensitive personal information with advertisers when in practice, the entity does indeed share this information with advertising platforms).

[24] Health Breach Notification Rule, 16 C.F.R. § 318.2; *see also Complying with FTC's Health Breach Notification Rule*, Federal Trade Commission (July 2024), https://www.ftc.gov/business-guidance/resources/complying-ftcs-health-breach-notification-rule-0, ("Incidents of unauthorized access, including a company's disclosure of covered information without a person's authorization, triggers notification obligations under the Rule.").

[25] Concurring and Dissenting Statement of Comm'r Andrew N. Ferguson, *In re Social Media & Video Streaming Servs. Report*, FTC File No. P205402 (Sept. 19, 2024) at 6, https://www.ftc.gov/system/files/ftc_gov/pdf/ferguson-statement-social-media-6b.pdf ("Policymakers should focus on protecting consumer data privacy on the front end [collection] rather than on implementing the sort of amorphous, backend advertising regulations [inferences] that the report recommends.")

[26] *See* Concurring and Dissenting Statement of Commissioner Andrew N. Ferguson, *In re Mobilewalla, Inc.*, FTC File No. 202-3196, & *In re Gravy Analytics, Inc.*, FTC File No. 202-3191 (Dec. 3, 2024) at 3 fn. 24,

https://www.ftc.gov/system/files/ftc_gov/pdf/gravy_-mobilewalla-ferguson-concurrence.pdf; *id.* at 5. *See also* Dissenting Statement of Commissioner Melissa Holyoak, *In re Mobilewalla, Inc.*, FTC File No. 202-3196 (Dec. 3, 2024) at 2, https://www.ftc.gov/system/files/ftc_gov/pdf/commissioner-holyoak-dissent-mobilewalla.pdf ("The Majority erroneously declares Mobilewalla's collection of consumer information from the RTB exchanges is unfair. Specifically, the Complaint alleges that the practice of collecting data was unfair in part because it caused or is likely to cause substantial injury. But the Complaint's allegations are remarkably sparse when it comes to establishing how the collection itself caused substantial injury, and its related allegations do not otherwise satisfy what Section 5 requires for unfairness.")

[27] *See* Dissenting Statement of Commissioner Melissa Holyoak, Joined by Commissioner Andrew N. Ferguson, *In re Health Breach Notification Rule*, FTC File No. P205405 (Apr. 26, 2024), https://www.ftc.gov/system/files/ftc_gov/pdf/p205405_hbnr_mhstmt_0.pdf (arguing that the Commission's "capacious" and "astonishingly broad" definitions of health-related entities and data exceed the bounds Congress established).

[28] *See, e.g.*, Data Security Program, 28 C.F.R. § 202.249(a) ("The term sensitive personal data means covered personal identifiers, precise geolocation data, biometric identifiers, human 'omic data, personal health data, personal financial data, or any combination thereof.") (emphasis added); Protecting Americans' Data from Foreign Adversaries Act, 15 U.S.C. § 9901(c)(7) (defining "Sensitive Data" to include biometric information, genetic information, and information identifying the sexual behavior of an individual).

[29] *See, e.g.* Va. Code 59.1-575(1) (defining sensitive data).

[30] *E.g.* Tex. Bus. & Com. Code § 541.001(29) ("Sensitive data means a category of personal data. The term includes: (A) personal data revealing . . . mental or physical health diagnosis[.]"); Colo. Rev. Stat. § 6-1-1303(24); Conn. Gen. Stat. § 42-515(38) (2025); *but see* Cal. Civ. Code § 1798.140(ae) (defining Sensitive Personal Information to include PI "collected and analyzed concerning a consumer's health.") (emphasis added).

[31] *Cf. Am. Hosp. Ass'n v. Becerra*, No. 4:23-cv-1110-P (N.D. Tex.) (June 20, 2024), slip op. at 18, 24 ("A user's intent in visiting a[n] [unauthenticated public webpage] is unknowable . . .The closest the Proscribed Combination gets to [Individually Identifiable Health Information] is a speculative inference extrapolated from (but unsubstantiated by) collected metadata.").

[32] *See* Conn. Gen. Stat. § 42-515(9) (2025) (defining consumer health data to include personal data a controller "uses to identify a consumer's physical or mental health condition, diagnosis or status) (emphasis added). *See also* RCW 19.373.010(8)(a) (defining "consumer health data" to include PI that "identifies the consumer's past, present, or future physical or mental health status."); Nev. Rev. Stat. § 603A.430 (2024) ("Consumer health data means personally identifiable information… that a regulated entity uses to identify the past, present or future health status of the consumer.") (internal quotations removed).

[33] *See* Cal. Civ. Code § 1798.140(ae) (defining Sensitive Personal Information to include PI "collected and analyzed concerning a consumer's health.").

[34] *See* Wash. State Office of the Att'y Gen., *Protecting Washingtonians' Personal Health Data and Privacy*, https://www.atg.wa.gov/protecting-washingtonians-personal-health-data-and-privacy, ("The definition of consumer health data includes information that is derived or extrapolated from nonhealth data when that information is used by a regulated entity or their respective processor to associate or identify a consumer with consumer health data… any inferences drawn from purchases could be consumer health data.").

[35] *See* Colo. Code Regs 904-3, Rule 2.02 ("'Revealing' as referred to in C.R.S. § 6-1-1303(24)(a) includes Sensitive Data Inferences."); *id.* ("'Sensitive Data Inference' or 'Sensitive Data Inferences' means inferences

made by a Controller based on Personal Data, alone or in combination with other data, which are used to indicate an individual's racial or ethnic origin; religious beliefs; mental or physical health condition or diagnosis; sex life or sexual orientation; or citizenship or citizenship status."). *See also generally* Daniel J. Solove, *Data is what Data Does: Regulating Based on Harm and Risk Instead of Sensitive Data*, 118 Nw. U. L. Rev. 1081, 1105 (2024) ("[N]onsensitive data, such as mundane purchases for lotion, soap, and cotton balls, can be used to infer sensitive data about health.").

[36] *See, e.g.*, N.J. Stat. Ann. § C.56:8-166.4 ("Sensitive data means personal data revealing . . .; mental or physical health condition, treatment, or diagnosis[.]") (emphasis added).

[37] *See* RCW 19.373.010(8)(a) (defining "Consumer Health Data" to include any data processed to associate or identify a consumer's past, present, or future physical or mental health status).

[38] *See* 45 C.F.R. § 160.103 (scoping Health Information under HIPAA to information created or received by a "health care provider, health plan, public health authority, employer, life insurer, school or university, or health care clearinghouse[.]").

[39] *See* Wash. State Office of the Att'y Gen., *Protecting Washingtonians' Personal Health Data and Privacy*, https://www.atg.wa.gov/protecting-washingtonians-personal-health-data-and-privacy, ("[The My Health My Data Act] defines consumer health data to include the 'use and purchase of prescribed medication.' Non-prescription data is only considered consumer health data if the regulated entity draws an inference about a consumer's health status from its purchase of non-prescription medication.").

[40] *See* Final Judgement and Permanent Injunction at 2, *People v. Healthline Media LLC*, No. CGC-25-626794 (Cal. Super. Ct. S.F. Cnty. July 28, 2025) (hereinafter "Healthline"), https://oag.ca.gov/system/files/attachments/press-docs/Signed%20Judgment_0.pdf, (defining Diagnosed Medical Condition Article to mean an article with a title or URL that indicates the consumer visiting the article has already been diagnosed with a medical condition and restricting Healthline from selling this data).

[41] *See id.* at 8 (arguing that it may be unlawful to share data "of a more intimate nature" with third parties if consumers would not expect that to happen).

[42] *See generally* Daniel J. Solove, *Data is what Data Does: Regulating Based on Harm and Risk Instead of Sensitive Data*, 118 Nw. U. L. Rev. 1081, 1105 (2024) ("With Big Data and powerful machine learning algorithms, most nonsensitive data give rise to inferences about sensitive data.").

[43] *Id.* at 1104-05 (summarizing how Target used purchase data to generate a "pregnancy prediction score").

[44] *See* Network Advert. Initiative, *2023 Demographic Health Advertising Best Practices*, https://thenai.org/wp-content/uploads/2023/11/NAI-Health-Targeting-Best-Practices-Document-Final.pdf. ("The NAI does not consider an audience segment that is created based only on demographic attributes such as age, gender, education level, presence of children, or region to reveal, or to be an inference about, a health condition, treatment, or diagnosis of any specific individual in the audience segment.").

[45] *See* Network Advert. Initiative, *2020 NAI Code of Conduct* at 23 ("[S]ensitive health segments, which require Opt-In Consent under the Code, include, but are not limited to, categories such as: drug addiction, all sexually transmitted diseases (such as AIDS, HIV, HPV), all types of mental health conditions (such as generalized anxiety disorder, schizophrenia, Alzheimer's, depression, anorexia/bulimia), pregnancy termination, all conditions predominantly affecting or associated with children that are not treated by over-the-counter medications, as well as cancer.").

[46] *See, e.g.* Complaint, *In re BetterHelp, Inc.*, FTC Docket No. C-4796 (July 7, 2023) at 13, https://www.ftc.gov/system/files/ftc_gov/pdf/2023169betterhelpcomplaintfinal.pdf ("Following the February 2020 publication of news reports that Respondent was sharing consumers' health information with third parties, including Facebook, numerous Users contacted Respondent and voiced their anger about the disclosures. For example, one individual noted: "I learned that you sell yet more private information to Facebook. This is disgusting. This information makes clients easily identifiable and your platform takes 100% control of its dissemination. I have no ability to decide where that information is sent. Only you do." Another stated: "I have not given ANY consent to share my information with ANYONE. ESPECIALLY ads targeting my mental health 'weakness.'").

[47] *E.g.* Network Advert. Initiative, *2020 NAI Code of Conduct*, https://thenai.org/wp-content/uploads/2021/07/nai_code2020.pdf, (defining Sensitive Information to include information, including inferences, about any past, present, or potential future health or medical conditions or treatments and information about sexual orientation).

[48] However, privacy disclosures alone are unlikely to be the only consideration in setting consumer expectations. *See* Cal. Code Regs. tit. 11 § 7002(b) (limiting the purposes for which personal information is collected to be consistent with the "reasonable expectations of the consumer(s) whose personal information is collected or processed" based on five factors, including the "type, nature, and amount of personal information that the business seeks to collect or process" and the "source of the personal information and the business's method for collecting or processing it."); *see also* Healthline, *supra* note 40, at 8 ("Thus, the law provides that invisibly sharing data of a more intimate nature to third parties, briefly alluded to in a privacy policy, may be unlawful when consumers would not expect that to happen. The law further provides that even detailed privacy disclosures regarding other intended uses of data may violate the principle if the disclosed purposes differ substantially from the consumer's reasonable expectations.").

[49] *See, e.g.*, Complaint, *In re BetterHelp, Inc.*, FTC Docket No. C-4796 (July 7, 2023) at 16, https://www.ftc.gov/system/files/ftc_gov/pdf/2023169betterhelpcomplaintfinal.pdf (alleging that "Respondent's collection, use, and disclosure of millions of Visitors' and Users' health information without reasonable privacy practices or safeguards has caused or is likely to cause them substantial injury."); Compl. at 20, *United States v. GoodRx Holdings*, No. 23-cv-00460 (N.D. Cal. Feb. 1, 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/goodrx_complaint_for_permanent_injunction_civil_penalties_and_other_relief.pdf (alleging that "Disclosure of [personal and health] information without authorization is likely to cause GoodRx users stigma, embarrassment, or emotional distress, and may also affect their ability to obtain or retain employment, housing, health insurance, disability insurance, or other services.").

[50] Some state privacy laws would also likely treat this type of processing as a form of "profiling" requiring additional disclosures and the right for consumers to opt out. *See, e.g.*, Colo. Rev. Stat. § 6-1-1303(20) (defining "profiling"); § 6-1-1306(1)(a)(I)(C) (providing consumers with the right to opt out of "[p]rofiling in furtherance of decisions that produce legal or similarly significant effects[.]"); Colo. Code Regs 904-3, Rule 6.03(A)(1)(c) (requiring a privacy notice that provides consumers with a meaningful understanding of how their personal data will be used, including whether it will be sold or used for profiling). California's regulations go further, defining "significant decision" as a "decision that results in the provision or denial of financial or lending services, housing, education enrollment or opportunities, employment or independent contracting opportunities or compensation, or healthcare services" and requiring a risk assessment be performed when using automated decision-making technology (ADMT) to make a significant decision concerning a consumer. *See* Cal. Code Regs. tit. 11 § 7001(ddd) (defining "Significant Decision"); § 7150(b)(3) (requiring a risk assessment be conducted when using ADMT for a significant decision). Additionally, the Americans with Disabilities Act acknowledges this harm by, among other things, restricting employers from requiring medical examinations or making disability-related inquiries about employees, and limiting inquiries only to questions about the applicant's ability to perform job-related functions. *See* 42 U.S.C. § 12112(d)(2)(A)–(B).

[51] *See generally* Complaint, *In re Mobilewalla, Inc.*, FTC Docket No. 202-3196 (Dec. 3, 2024), https://www.ftc.gov/system/files/ftc_gov/pdf/2023196mobilewallacomplaint.pdf (alleging Unfair Collection of Consumer Information from RTB Exchanges) (hereinafter "Mobilewalla Complaint").

[52] *See* Dissenting Statement of Commissioner Melissa Holyoak, *In re Mobilewalla, Inc.*, FTC File No. 202-3196 (Dec. 3, 2024) at 2, https://www.ftc.gov/system/files/ftc_gov/pdf/commissioner-holyoak-dissent-mobilewalla.pdf ("The Majority erroneously declares Mobilewalla's collection of consumer information from the RTB exchanges is unfair. Specifically, the Complaint alleges that the practice of collecting data was unfair in part because it caused or is likely to cause substantial injury. But the Complaint's allegations are remarkably sparse when it comes to establishing how the collection itself caused substantial injury, and its related allegations do not otherwise satisfy what Section 5 requires for unfairness.").

[53] *See* 15 U.S. Code § 45(n) (measuring unfairness by considering whether the act or practice causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition); *See also* Mobilewalla Complaint, *supra* note 51, at 10, https://www.ftc.gov/system/files/ftc_gov/pdf/2023196mobilewallacomplaint.pdf, (alleging that the selling of sensitive location information is likely to cause substantial injury to consumers and is not outweighed by countervailing benefits to consumers or competition).

[54] *See* Wash. State Office of the Att'y Gen., *Protecting Washingtonians' Personal Health Data and Privacy*, https://www.atg.wa.gov/protecting-washingtonians-personal-health-data-and-privacy, ("Information that does not identify a consumer's past, present, or future physical or mental health status does not fall within the Act's definition of consumer health data."); *but see, e.g.*, Mike Hintze, *Washington My Health My Data Act - Part 2: The Scope of "Consumer Health Data"* (Apr. 12, 2023), https://hintzelaw.com/blog/2023/4/12/wa-my-health-my-data-act-pt-2-scope-of-consumer-health-data (suggesting that PI indicating a consumer purchased running shoes could indicate that the consumer is seeking health care services covered by Washington's My Health My Data Act).

[55] While this example is purely hypothetical, it is consistent with diabetes prevalence data derived from the Centers for Medicare & Medicaid Services (CMS) Medicare Current Beneficiary Survey (MCBS), as analyzed in the Commonwealth Fund report. *See* Janet Sutton et al., *The Health Care Experiences of People Dually Eligible for Medicare and Medicaid* (2024), https://www.commonwealthfund.org/sites/default/files/2024-07/PDF_Sutton_dual_eligibles_comparing_TM_MA_chartpack.pdf. That report indicates a prevalence rate of 39.3% specifically for beneficiaries eligible for both Medicare and Medicaid ("dual eligible"). In federal health policy analysis, dual eligibility is the standard proxy for low-income seniors, as qualification requires meeting strict federal poverty guidelines (typically an income below 100% of the Federal Poverty Level and limited assets). This cohort consistently exhibits higher rates of chronic conditions compared to Medicare-only beneficiaries. *See also* Ctrs. for Disease Control & Prevention, *National Diabetes Statistics Report* (2024), https://stacks.cdc.gov/view/cdc/148231 (indicating that the overall prevalence of diabetes in the United States is approximately 11.6% across all ages, rising to 14.7% among adults aged 18 and older).

[56] *See* Network Advert. Initiative, *2023 Demographic Health Advertising Best Practices*, https://thenai.org/wp-content/uploads/2023/11/NAI-Health-Targeting-Best-Practices-Document-Final.pdf, ("The NAI does not consider an audience segment that is created based only on demographic attributes such as age, gender, education level, presence of children, or region to reveal, or to be an inference about, a health condition, treatment, or diagnosis of any specific individual in the audience segment.").

[57] *See* Ctrs. for Disease Control & Prevention, *National Diabetes Statistics Report* (2024), https://stacks.cdc.gov/view/cdc/148231 (indicating that the overall prevalence of diabetes in the United States is approximately 11.6% across all ages, rising to 14.7% among adults aged 18 and older).