[[1]](#footnote-1)

Addressing Information Asymmetry in the Health Insurance Market:

A Practical Framework for a Universal Health Insurance Database

Meredith Ludlow and Javier Saldaña Jr.

*Abstract*— A database is only as good as the data it collects. The health insurance industry like many other industries, is fragmented and only able to perform data analytics to the depth their database allows. There are legal and practical constraints which complicate the health insurance industry to strive for a collective database like the one privy to the property and casualty insurance industry. A framework for a collective relational database to be shared by health carriers is introduced. Along with the framework and design of the collective database, a roadmap for overcoming the constraints currently in place are also presented.

*Index Terms*—Database, Economics, Health Insurance, Information Asymmetry, Relational Database

# INTRODUCTION

O

NCE considered a competitive advantage, data analytics are now mandatory for a business to be able to survive in today’s market. The insurance industry, which has been hoarding data on its consumer base for decades, is starting to implement the use of data analytics into other facets of the business outside of their underwriting practices. However, the databases used by the carriers are only as powerful as the data they contain. Every time a potential customer is seeking new health insurance, they hold more information about their medical necessities than the carrier is privileged to know at inception. This is known as information asymmetry and causes carriers to mitigate health care costs by including a markup in the premiums issued.

The problem of information asymmetry could be addressed through the development of a collective relational database that would be shared amongst all health carriers and health care professionals. Although there are legal constraints which must be addressed, the benefits of a collective health insurance database would extend to insureds and insurers alike. This paper starts by further explaining information asymmetry, then proposes a theoretical universal healthcare database model that best deals with information asymmetry. The proposed model will be a relational database that combines information about policies, insureds, carriers, claims, and healthcare providers. This is followed by a discussion of the major data protection legislation that the database will need to comply with such as the Gramm-Leach-Bailey act and HIPPA. The major consequence of a universal database is that although healthier people will get lower premiums, people with pre-existing conditions and health problems will be charged more.

# Information Asymmetry & “Lemons”

Asymmetrical information is not a new concept in today’s markets. It has been around for centuries and has allowed false promises to take advantage of the placebo effect in order to survive. In an ideal market, the buyer and the seller are well informed about the product, which leads to the creation of the true market value for the product in a transaction. Unfortunately, this isn’t the case in many markets today, health insurance being one of them.

George Akerlof, an economist awarded with the 2001 Nobel Prize, argues that asymmetric information results in market failure [1]. He applies his model to the health insurance industry by pointing out how premiums are adjusted to accommodate for the asymmetrical information present in the market. Normally, the market would expect premiums to reflect the risk attributed to that one individual. However, the elderly individuals are having trouble getting new health care policies. If the premiums reflect the health risk of the individual, then the market would anticipate 65-year-old insured to have higher premiums but still be able to get a policy, nonetheless. Unfortunately, that is not the case. As you get older, it get increasingly difficult to get a health insurance policy. Once you turn 65-years-old, it becomes nearly impossible to get a new health insurance policy in the open market. The question posed by Akerlof is, if a 65-year-old needs health insurance, why can’t the premiums rise to match the risk?

The reason is that a 65-year-old is more than likely going to be a “lemon”. Akerlof argues every market has good products and “lemons,” which are essentially bad products. In the health insurance market, a healthy 20-year-old is more likely to be a good product while a 65-year-old is more likely to be a “lemon.” At the age of 65, people are more prone to illnesses and other medical complications, which normally result in frequent doctor’s visits. The health insurer is certain the insured will be submitting claims, thus making the insured unprofitable for the expected time the insured is anticipated to be in the books of business. The lack of profitability is the cornerstone as to why the elderly are unable to obtain new policies at that age and are viewed as “lemons” in the market. However, when the insured is 20-years-old, the insurer is taking a gamble that the individual is not likely to be submitting claims until decades later once the customer is deemed profitable. This gives the health insurer a higher probability to collect premiums well before the insured truly needs the health insurance.

The probability of whether the insured will be a good insured or a “lemon” is all based on asymmetric information. The insurer doesn’t know the true medical history of the 20-year-old (since there is a conflict of interest for the insured when disclosing health information to the insurer) and as a result, must take a gamble when writing the policy for the insured. Will the insured file a claim immediately after suddenly being diagnosed with a life-threatening illness (thus being classified a lemon)? Or will the insured pay premiums for decades and only start to file serious claims once the insured is then 50+ years old (thus making the insured a good insured)? This gamble however doesn’t exist with a prospective insured who is 65-years-old. At this age, the insurer is aware there is an enormously high probability that the new insured would be submitting claims much more frequently than a healthy 20-year-old insured. As a result, insurers view it safe to deem 65-year-olds as “lemons” in the market.

In economics, this gamble is viewed as an opportunity cost and could be given a financial value to determine whether the decision rendered is a financially beneficial decision. Considering the ultimate interest of a company lies in its shareholders, a company will do everything in its power to offset the losses from the “lemons” in order to remain profitable. In order to do so, the insurer spreads out the risk of acquiring “lemons” to all of its book of business by maintaining a market risk that attaches to all policies, regardless of whether the customer is a good customer or a “lemon.” This normally translates into higher premiums for the “good” insured even though the premium being paid by the good customer doesn’t reflect their true health risk.

# The Relational Database Design

In order to address the issue regarding information asymmetry in the health insurance industry, we proposed a relational database design.

## Benefits of a Relational Database Design

A relational design is ideal for this proposal considering the data gathered by the health insurance industry is extremely structured. By keeping the data inserted into the database structured, the entities feeding the data into the database are forced to abide by the rules set in place by the database. This helps maintain uniformity and allows for quicker analysis of the data. We are aware that outside of the forms and test results, medical information can be very unstructured such as x-ray images, physician notes/recommendations, etc. These are all elements which, while beneficial for medical professionals, are not very helpful for insurance adjusters. Licensed (if required by the regulatory state) adjusters are not medical professionals, which means they are not able to review medical records and conduct a diagnosis or even recommend a treatment plan. Health/bodily injury adjusters use recommended treatments/tests and billing formulas to adjust a medical claim. Even though a physician may recommend 9 weeks of treatment, the adjuster would argue that injury normally incurs only 6 weeks of treatment based on the recommended guidelines put forth by the United States Department of Health and Human Services. The discrepancy between treatment duration is then dealt with in court when a medical expert is brought in to challenge/defend the treatment plan.

Aside from structured data, a relational database will ensure the database remains true to the ACID principles. A database of this magnitude should have the capability to handle an immense amount of transactions, all the while maintaining consistency. The ACID principles guarantee the data being displayed to the underwriter/adjuster/medical professional/etc. is the most current and true to the patient’s history. For a medical professional, this means being able to view a patient’s history regardless of where the visit occurred. For an underwriter, this equates into seeing the true medical history of the individual in order to asses a relative risk score and properly assess premiums. For an adjuster, this means being to view prior medical treatments and if any injuries may have been aggravated.

Since relational database systems have been around much longer than the up-and-coming NoSQL databases, the data being kept by many self-employed medical professionals is likely to be kept in a standard relational database system. This would allow for a much simpler normalization and data migration into the larger collective database, which they would also have access to for their business practice.

## Proposed Framework of the Database Design

The proposed framework is a relational model. The figure

for a basic version of it is in the appendix. The central table of the database is the policy table. This table will contain information about the policy, including policy number, the type of policy, the effective dates of the policy, and whether the policy is currently active. The policy table is connected to the primary insured table, the claim table, and the carrier table. The carrier table will have information about the insurance carrier that hold the policy. The relationship will be one-to-many as a carrier has many policies.

The primary insured table will have information about the primary insured that is insured under a given policy. It will also contain their medical history prior to them being entered into the database, including pre-existing conditions, previous claims and family health history. The relationship between primary insured and policy is one to many, as a single insured can have multiple policies over the course of their life and the database will have all of them in it. The primary insured table will be connected to the household members table which will have information on the members of an insured’s household and their medical history. As the database continues, there will be less and less need to include people’s medical history as separate fields of the database. Household members will break off and get their own policies with their own household members and their medical histories and that of their families will already be in the database.

Finally, the policy table will be connected to the claims table. The claims table will hold information about all claims that are filed on any of the policies in the database. The relationship will be one to many as a single policy can have more than one claim. The claim table will connect to the healthcare provider table that will carry information about what healthcare provider is providing the medical care for the claim. Finally, the claim table is connected to the claim payment table. This table gives payment information about the pay out for each claim.

# Constraints

The main challenges of creating a centralized database for healthcare insurance and claim information are federal regulations and privacy concerns. There are many federal regulations that dictate how health care data is protected and shared. This paper will focus on the Gramm-Leach-Bailey Act and HIPPA.

The Gramm-Leach-Bailey Act requires that financial institutions, including insurance companies, must share their data sharing practices with their customers and obligates an insurance company to protect its customer data [2]. Giving this data to a universal healthcare database could conflict with staying in compliance with this rule. Since insurance companies must share with customers their data sharing practices, customers would also have the option to opt out of sharing with the universal database. This could lead to a lack of data in the database if large amounts of people choose to opt out.

The second major regulation is HIPPA, or the Health Insurance Profitability and Accountability Act. HIPPA has quite a few parts that involve healthcare insurance practices, but the relevant section for this paper is Title II: Administrative Simplification. This section of HIPPA has a privacy rule that requires health insurance carriers to protect patient health information, specifically, information that makes a patient identifiable[3]. A proposed universal healthcare database could potentially violate the HIPPA privacy clause depending on what data was shared with the universal database. If the data given to the database was such that it didn’t violate a customer’s privacy, it might be so sparse that it could lessen the benefits of even having a universal database at all.

Another major challenge that arises in building a universal database is the danger of having all healthcare data in one place. While fragmented databases, each piece with a different carrier, are less beneficial for analytics, the data ends up being safer. If a carrier’s data where to be hacked, the hacker would only have access to a portion of the nation’s healthcare information. If a universal database where hacked, the magnitude of people whose information has the potential to be stolen increases. There is a balanced that needs to be found in the benefits of analytical ability and data privacy.

# Unintended Consequence

Addressing the information asymmetry in the health insurance industry will then correct premiums for many individuals. The premium would reflect a true risk factor tied to the insured’s health. However, this would result dangerous consequences. When people are shopping for a used vehicle, they will not purchase a certain vehicle if they know that vehicle will not last more than their worthwhile. For the health insurers, this would be the same as refusing to insure an insured if they know or believe that insured will not remain in their book of business to become a profitable insured. Afterall, the interest of health insurers is not to the public health but to their shareholders. In other words, removing the information asymmetry from the transaction would penalize the individuals who need the medical insurance most.

However, this is not a new phenomenon. Texas tort law, like many other states in the United States, requires a licensed driver carry at least a minimum mandated limit of liability insurance for themselves when operating a motor vehicle. Carriers run background checks on drivers prior to issuing policies in order to determine if the driver is a “good” driver or a “lemon.” In this instance, an insurer may refuse to write a policy. When a driver has been refused by more than two carrier for a voluntary auto insurance policy, they qualify for the Texas Automobile Insurance Plan Association (TAIPA) insurance coverage. TAIPA is a state-sponsored reserved for high-risk drivers in Texas who are unable to attain a voluntary automobile insurance. TAIPA reviews and then assigns the high-risk insured to an auto insurance company based on their given market share of policies in the state.

A similar approach could be applied to the health insurance industry in order to fill in the hole that would be left by information asymmetry. High risk insureds would be allocated to health insurers across the country based on their market share. This would ensure everyone is able to get a health insurance policy. However, it still doesn’t address the issue pertaining to high premiums. Insurers would be able to provide true risk premiums, which would mean these individuals may end up with premiums they can’t afford. Since that issue is now a public health and welfare concern, that burden would be bestowed upon the public to provide for those who are unable to provide for themselves through expanded and/or reformed welfare programs.

# Conclusion

By empowering the health insurers and addressing the information asymmetry through the collective relational database, the premiums should then change to more accurately reflect a true premium based on risk of the insured. Healthier individuals would then be rewarded with lower premiums, but high-risk individuals could be penalized unintentionally.

Nonetheless, a collective relational database shared between insurers and health care professionals would give insurers the ability to minimize (if not eliminate) transactions based off asymmetric information. The stability of a relational database coupled with the transparency that allows markets to thrive would give the health insurance market the jump start it needs into a fair market, which could see health insurance become extremely affordable for many individuals across the country.

References

1. J. E. Stiglitz, “The market for lemons: quality uncertainty and the market mechanism,” *Market Failure or Success*, vol. 84, no. 3, pp. 488–500, Aug. 1970*.*
2. “Gramm-Leach-Bliley Act.” Federal Trade Commission, [www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act](http://www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act)
3. “What Is HIPAA (Health Insurance Portability and Accountability Act) ? - Definition from WhatIs.com.” SearchHealthIT, searchhealthit.techtarget.com/definition/HIPAA

1. Submitted for review on Mar. 17, 2019.

   M. Ludlow is a graduate student at Southern Methodist University in Dallas, TX. She is currently employed by \_\_\_ Insurance Agency in \_\_\_, TX as a Licensed Insurance Agent (e-mail: [mludlow@mail.smu.edu](mailto:mludlow@mail.smu.edu)).

   J. Saldaña is a graduate student at Southern Methodist University in Dallas, TX. He is employed by Allstate Insurance in Irving, TX as a Senior Data Consultant (e-mail: [saldanaj@smu.edu](mailto:saldanaj@smu.edu)). [↑](#footnote-ref-1)