

PROPOSAL FOR HEALTH INSURANCE COST PREDICTION APP:
CareCompass

Prepared by Angelina Cottone

UWP 104AY
University of California, Davis
November 11, 2025

Table of Contents

Executive Summary.....	2
Introduction.....	2
The Problem.....	3
Patients.....	3
Insurers.....	3
Providers.....	3
Recommendation.....	4
Core Features.....	4
Technical Framework.....	4
Data and Security.....	4
Modeling Approach.....	5
Benefits.....	5
Patients.....	5
Insurers.....	5
Providers.....	5
Risks and Mitigation Strategies.....	6
Implementation Plan.....	6
Phase 1: Research and Data Collection (Months 1-4).....	6
Phase 2: App Development and Testing (Months 5-10).....	6
Phase 3: Beta Launch and User Feedback (Months 11-14).....	7
Phase 4: Full Rollout and Marketing (Months 15-18).....	7
Projected Financial Cost.....	8
Evaluation Plan.....	9
Conclusion.....	10
References.....	10

Executive Summary

Patients in the United States often face unexpected or unclear medical costs due to limited transparency and a lack of individualized cost prediction in the U.S. healthcare system. This unpredictability contributes to financial stress and may discourage individuals from seeking essential preventative care. **CareCompass** is a data-driven mobile application that leverages predictive analytics, generalized linear models (GLMs), and machine learning to forecast individualized healthcare costs. The app provides personalized procedure and annual cost estimates, confidence intervals to communicate uncertainty, real-time deductible and copay tracking, provider comparisons, and a visual forecast dashboard.

The development of this app would help patients gain clarity and plan for healthcare expenses proactively, reducing stress. Insurers would benefit from improved customer satisfaction, retention, and reduced billing disputes. Providers would experience streamlined billing and enhanced patient communication with reduced payment delays. CareCompass will be developed through a phased approach in collaboration with insurance companies, healthcare providers, and a health analytics development team. Phases include research and data collection, app development and testing, beta launch, and full rollout.

Introduction

Rising healthcare costs in the United States have led to growing confusion and financial stress for patients. As of 2021, hospitals are required to publish machine-readable files or online displays of “shoppable services”, enabling consumers to estimate the cost of care before receiving treatment. (*Hospital price transparency: CMS*) Despite this requirement, unexpected and rising healthcare costs continue to be a burden to patients nationwide. Nearly 20% of patients undergoing in-network elective surgeries or childbirth received surprise bills for out-of-network services, even when they believed their care was fully covered. (*Pollack, 2022*)

Significant price discrepancies also exist among insurance companies for identical procedures at the same hospital. For example, at University of Mississippi hospitals, a colonoscopy was priced at \$2144 for Aetna plan patients, \$1463 for Cigna plan members, and \$782 for patients without insurance. (*Pollack, 2022*) This variability among plans highlights how inconsistent pricing structures limit patients’ ability to make informed choices about their care. While moves have been made to decrease price discrepancies and the lack of transparency in healthcare, most patients lack financial literacy or the analytical tools to interpret and apply the information effectively. Medical debt is also widespread issue, as around 20 million Americans owed medical bills in 2024. (*Rakshit et al., 2024*)

To address this gap, this proposal recommends the development of CareCompass, a mobile tool designed to estimate out-of-pocket costs and forecast annual healthcare spending for individuals.

By integrating healthcare cost transparency with statistical forecasting, the app will assist patients in understanding their expected financial responsibility before receiving services. This innovation aligns with ongoing reforms like the 2022 No Surprises Act, which aims to protect consumers from out-of-network surprise billing but still leaves gaps. (*J et al., 2023*) The proposed app would build upon these policy efforts by offering a practical and personalized solution that helps patients plan for healthcare expenses and make informed treatment decisions.

The Problem

A core issue in the U.S. healthcare system is that patients often discover the true cost of medical care only after receiving treatment, which can result in significant financial hardship. This lack of transparency not only creates stress for patients but also discourages preventative care, ultimately undermining trust in healthcare systems. This has the following effects:

Patients

- Difficulty predicting out-of-pocket costs.
- Avoidance of preventative care.
- Confusion around deductibles, copays, and coverage.
- Financial stress and doubt.

Insurers

- Lower customer satisfaction and retention.
- High volume of billing disputes and delayed payments.
- Inefficient resource allocation.

Providers

- Delayed payments.
- Administrative burden from patient confusion.
- Strained provider-patient trust.

Current tools and insurance apps provide some information on procedure costs; however, they often fail to account for individualized variables such as deductible progress, comorbidities, or in-network vs. out-of-network coverage. Most applications also lack the presence of confidence intervals or probability distributions to indicate the range of likely out-of-pocket expenses, and generally are unable to forecast future expenses over time based on patients' treatment patterns and conditions, leaving user with incomplete financial insight.

Therefore, the problem is both a data-accessibility issue and a health-literacy issue. Pricing data exists, but is often not present in meaningful, personalized, and easily understandable ways.

Predictive analytics and modeling provide a practical solution to this challenge, offering the ability to estimate individualized costs and project future spending patterns.

Recommendation

To address the challenge of unpredictable healthcare costs, the development of the predictive health cost app CareCompass is proposed, a data-driven mobile tool designed to provide personalized forecasts of out-of-pocket expenses and future healthcare spending patterns. CareCompass leverages actuarial modeling, generalized linear models (GLMs), and machine learning techniques to predict costs accurately and provide actionable insights for patients, insurers, and care providers.

Core Features

1. Personalized Cost Forecasts

- GLMs, regression models, and machine-learning algorithms estimate procedure-specific and annual costs.
- Predictions incorporate claims history, demographic information, comorbidities, insurance plan details, and deductible status.

2. Confidence Intervals and Risk Bands

- Users receive both point estimates and uncertainty ranges (e.g., 90% prediction intervals).

3. Deductible and Copay Tracking

- Real-time updates on remaining deductible balances and expected out-of-pocket spending.
- Integration with insurer APIs ensures accuracy.

4. Provider Comparisons

- In-network providers ranked by historical cost, predicted out-of-pocket expense, and patient satisfaction ratings.

5. Forecast Dashboard

- Visual display of predicted annual spending, trends, and upcoming cost spikes.

Technical Framework

CareCompass will integrate securely with insurance APIs and utilize de-identified claims data in accordance with HIPAA policy and standards.

Data and Security

- Encrypted connections to insurer APIs.
- Use of de-identified historical claims data.
- Third-party security audits and penetration testing.

Modeling Approach

Predictive models include Gamma or lognormal GLMs and machine learning algorithms such as gradient boosting, Artificial Neural Networks (ANNs), Ridge regression, and LASSO variable selection. Empirical research supports the effectiveness of these models:

- A study utilized University of Utah health plans data found that gradient boosting had the best predictive performance for low and medium cost patients, while ANNs and Ridge regression performed best for high cost patients. (*Morid et al., 2018*)
- Analysis of the 2003 Medical Expenditure Panel Survey (MEPS) demonstrated that GLMs with Lasso variable selection accurately predict outpatient expenditures, including both the likelihood of incurring costs and expenditure amounts. (*Ge, 2024*)

Different predictive models work well for different populations because of cost distributions. Low-cost patients who mostly receive routine care have more stable and predictable costs that can be modeled through gradient boosting, while high-cost patients have skewed distributions which are better modeled through neural networks and ridge regression. CareCompass will select models based on performance and recalibrate regularly with new data.

Benefits

CareCompass offers measurable advantages for all stakeholders:

Patients

- Clear understanding of expected medical spending.
- Reduced risk of surprise bills.
- Increased financial literacy and confidence.
- More likely to seek and keep up with preventative care.

Insurers

- Strengthened trust and member satisfaction.
- Fewer billing disputes.
- Streamlined communication and a competitive advantage.

Providers

- Fewer payment delays.
- Improved patient trust and understanding.
- Reduced administrative burden from billing questions.

By combining predictive analytics, modeling, and user-centered design, CareCompass addresses current limitations in insurance cost estimation tools. Unlike traditional applications that provide static, generalized estimates, CareCompass adapts predictions to individual users, forecasts spending patterns, and communicates any uncertainties clearly and effectively, empowering patients to make informed healthcare decisions while also supporting insurers and providers.

Risks and Mitigation Strategies

Risk	Mitigation
Data Security	End-to-end encryption, HIPAA compliance, and penetration testing.
Prediction Error	Display confidence intervals, provide disclaimers, recalibrate models regularly.
Insurance Resistance	Begin with progressive insurers for pilot programs.
Low User Adoption	Partner with insurers for integrated enrollment and education programs.

Implementation Plan

Phase 1: Research and Data Collection (Months 1-4)

- **Objectives:** Gather anonymized claims data from insurance companies and identify key predictors of healthcare costs, such as demographics, comorbidities, prior claims, deductible status, etc.
- **Steps:**
 - Establish data sharing agreements with insurers and hospitals ensuring HIPAA compliance.
 - Perform exploratory data analysis to understand distribution of data, missing values, and potential bias.
 - Conduct literature review and consult experts to validate predictor selection.
 - Create a comprehensive database of historical claims for training predictive models.
- **Deliverables:** Data dictionary, preliminary analytics report, and identified variables to feed predictive models.

Phase 2: App Development and Testing (Months 5-10)

- **Objectives:** Build a secure and user-friendly mobile application and integrate predictive models.
- **Steps:**

- Design user interface with input from UX/UI specialists to ensure accessibility and clarity for users.
- Develop backend infrastructure to handle secure data storage, API integration with insurers, and predictive analytics computation.
- Integrate predictive models (GLMs, machine learning algorithms) into app workflow, enabling real-time personalized cost forecasts.
- Conduct internal testing to check accuracy, security, and performance.
- Perform HIPAA compliance audits and vulnerability assessments to ensure compliance.
- **Deliverables:** Functional app prototype with integrated predictive models, compliance documentation, internal test reports.

Phase 3: Beta Launch and User Feedback (Months 11-14)

- **Objectives:** Test the app with a limited number of users to identify issues and refine predictive accuracy.
- **Steps:**
 - Partner with selected insurers to recruit pilot users representing diverse demographics and insurance plans.
 - Collect user feedback via surveys, focus groups, and in-app analytics on usability, clarity of cost forecasts, and dashboard functionality.
 - Monitor predictive model performance against actual claims to identify gaps or necessary recalibrations.
 - Make improvements to UX/UI, model outputs, and notification systems based on feedback.
- **Deliverables:** Beta test report, updated app with improvements, validated predictive model performance metrics.

Phase 4: Full Rollout and Marketing (Months 15-18)

- **Objectives:** Launch app for general use, advertise, and provide ongoing support to users.
- **Steps:**
 - Collaborate with insurers to integrate the app into their member platforms and marketing campaigns.
 - Conduct educational campaigns to inform patients about the app's features, benefits, and secure handling of data.
 - Provide training for insurance customer support staff and healthcare providers to address user inquiries.
 - Monitor app usage, predictive accuracy, and customer satisfaction continuously to identify potential enhancements.
- **Deliverables:** Full app deployment, marketing materials, support manuals, usage reports.

Projected Financial Cost

Developing and deploying the CareCompass app requires investment in data acquisition, app development, predictive modeling, and rollout.

1. Development Costs

- a. App Engineering and Infrastructure (**\$80,000 - \$150,000**)
 - Full-stack developers, mobile developers (iOS/Android), backend engineers.
 - Secure database development and cloud infrastructure (AWS/GCP with HIPAA compliance).
 - API integration with insurers and providers.
 - Authentication and authorization systems.
- b. Predictive Modeling and Data Science (**\$40,000 - \$100,000**)
 - Data scientists to develop GLMs, machine learning models, and testing frameworks.
 - Actuarial consultants to validate cost models.
 - Ongoing model recalibration during development.
- c. UX/UI Design (**\$15,000 - \$40,000**)
 - Accessibility testing.
 - User-flow mapping and interactive app prototype development.
- d. Security, Compliance, and Legal (**\$30,000 - \$90,000**)
 - HIPAA compliance audits.
 - Penetration testing.
 - Secure data-handling documentation.
 - Legal review of insurer contracts.

2. Data Acquisition and Integration Costs

- a. Claims Data Access (**\$10,000 - \$30,000**)
 - Costs related to obtaining anonymized or de-identified claims data from insurers.
 - Data normalization and preprocessing.
- b. API Development with Partners (**\$5,000 - \$20,000**)
 - Custom connections with insurer systems (deductible tracking, plan rules, etc.).
 - Maintenance of integration endpoints.

3. Testing and Pilot Launch

- a. User Testing
 - Recruiting diverse pilot users in partnership with insurers.
 - Running focus groups, surveys, and model-validation analyses.
- b. Model Performance Monitoring (**\$10,000 - \$30,000**)
 - Compared predicted vs. actual claims over pilot period.

- Adjustments to improve MAE and interval accuracy.

4. Full Rollout, Marketing, and Training

- a. Marketing and Patient Education (**\$75,000 - \$150,000**)
 - Digital ads, onboarding materials, explainer videos.
 - Integration into insurer marketing channels.
- b. Provider and Insurer Staff Training (**\$25,000 - \$75,000**)
 - Documentation and training sessions for support staff.
- c. Customer Support Setup (**\$50,000 - \$125,000**)
 - Hiring and training support staff.
 - Creating help-center resources.

Evaluation Plan

1. User Adoption Rates

- Track the percentage of policyholders actively using CareCompass monthly.
- Segment adoption rates by demographics, plan type, and region to identify areas for increased engagement.
- Aim for an adoption target of 30% of users within the first 6 months post-launch.

2. Accuracy of Cost Predictions

- Continuously compare predicted costs with actual claims to evaluate model performance.
- Track key performance metrics: Mean Absolute Error (MAE), prediction intervals coverage.
- Implement recalibration of models if performance falls below 90% accuracy.

3. Customer Satisfaction Surveys

- Conduct pre-implementation surveys to assess transparency, trust, and stress levels regarding healthcare costs.
- Follow-up surveys 3-6 months post-launch to measure improvements in user experience, confidence in cost estimates, and financial planning satisfaction.
- Use feedback for feature enhancements in app updates.

4. Reduction in Billing Disputes and Late Payments

- Collaborate with insurance companies and healthcare providers to monitor the number of billing disputes and late payments before and after implementation.
- Measure trends and assess whether CareCompass users experience fewer disputes compared to non-users.
- Use this data to refine predictive models, improve deductible and payment tracking, and enhance provider communication.

Conclusion

CareCompass addresses the transparency gap in healthcare costs by providing an all-in-one, personalized, and data-driven health insurance dashboard for patients. By leveraging predictive analytics, GLMs, and machine learning, the app enables patients to plan financially, empowers insurers to improve engagements, and assists providers in streamlining billing processes. For insurers, CareCompass offers a competitive advantage through improved member satisfaction and retention. For providers, it streamlines billing communication and reduces disputes. Most importantly, for patients, it restores agency and reduces anxiety when it comes to healthcare. The adoption of CareCompass will foster trust, improve financial literacy, and create a more transparent and efficient healthcare ecosystem for all involved.

References

- Ge, Z. (2024). Modeling Health Expenditures Using Generalized Linear Models.
<https://conservancy.umn.edu/server/api/core/bitstreams/45321415-5351-4308-9519-ea356ad062ac/content>
- Hospital price transparency: CMS. CMS.gov. (n.d.).
<https://www.cms.gov/priorities/key-initiatives/hospital-price-transparency>
- J, H., K, L., J, V., E, W.-A., R, S., & E, W. (2023, July 28). No surprises act: Perspectives on the status of consumer protections against Balance Billing. RWJF.
<https://www.rwjf.org/en/insights/our-research/2023/04/no-surprises-act--perspectives-on-status-of-consumer-protections-against-balance-billing.html#:~:text=Conclusion.%20Researchers%20indicate%20the%20No%20Surprises%20Act,required%20to%20protect%20consumers%20from%20unexpected%20charges>
- Morid, M. A., Kawamoto, K., Ault, T., Dorius, J., & Abdelrahman, S. (2018). Supervised Learning Methods for Predicting Healthcare Costs: Systematic Literature Review and Empirical Evaluation. AMIA ... Annual Symposium proceedings. AMIA Symposium, 2017, 1312–1321.
- Pollack, H. A. (2022). Necessity for and Limitations of Price Transparency in American Health Care. AMA Journal of Ethics, 24(11). <https://doi.org/10.1001/amajethics.2022.1069>
- Rakshit, S., Rakshit, S., Twitter, M. R., Claxton, G., Amin, K., & Twitter, C. C. (2024, February 12). *The burden of medical debt in the United States*. Peterson-KFF Health System Tracker.
<https://www.healthsystemtracker.org/brief/the-burden-of-medical-debt-in-the-united-states/#Share%20of%20adults%20who%20have%20medical%20debt,%20by%20health%20status%20and%20disability%20status,%202021>
- Verma, S. (2025, October 16). *Cost to build Ai Insurance App in 2025: Full guide*. Biz4Group.
<https://www.biz4group.com/blog/cost-to-build-ai-insurance-app>