Using Predictive Modeling to Reduce Claims Losses in Auto Physical Damage

CAS Loss Reserve Seminar 2003

Session 3 – Private Passenger Automobile Insurance

Frank Cacchione
Carlos Ariza

September 8, 2003

Today’s Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

Powerful new predictive modeling tools are available to capture operational improvement across most elements of the claims process

Technology enabling value drivers for Total Incident Management

Fraud Management	• **Predictive modeling can provide fraud "flags"**
Litigation Management	• **Technology can enable insurers to better determine litigation strategy**

Notification and Assignment
- **Early notification**
- **Predictive triage**
- **Direct to network**

Verification of coverage
- **Customer service**
- **UM pursuit**

Investigation and evaluation
- **Predictive modeling**
- **Image engineering**
- **Adjuster variance**

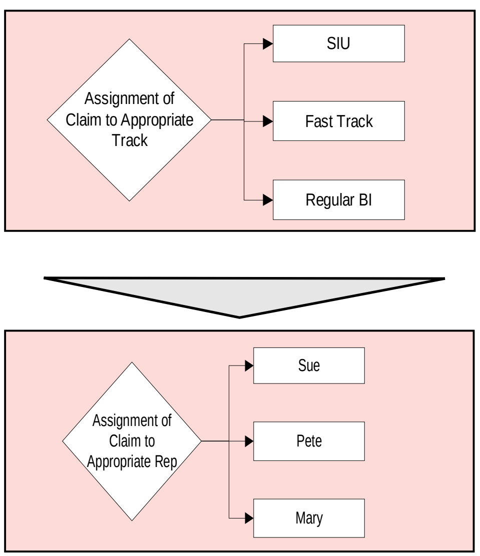
Negotiation and settlement
- **Colossus/ICE**
- **Lawyer profile**
- **Source law firms**
- **Motor repair network**

Salvage/Subrogation
- **Total loss assessment**

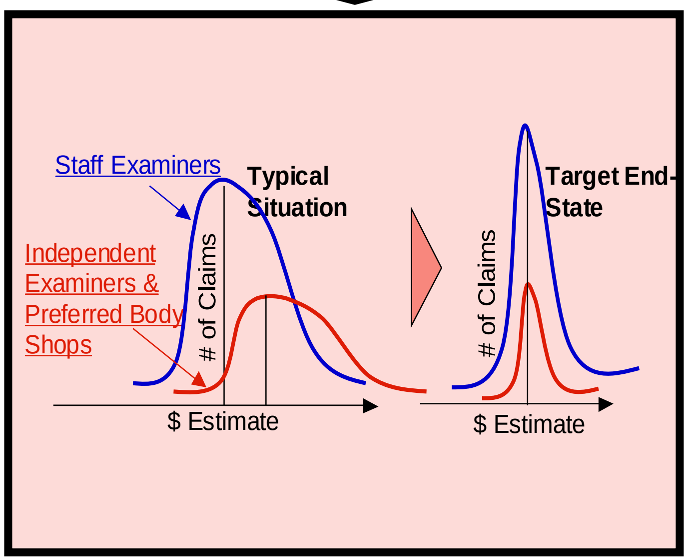
Significant operational improvement requires rapid determination of how a claim should be handled and intensive effort to reduce the variance in adjustors' evaluation of loss

Predictive models can aid the claim representative in determining the likelihood of overpayment and suggest strategy options

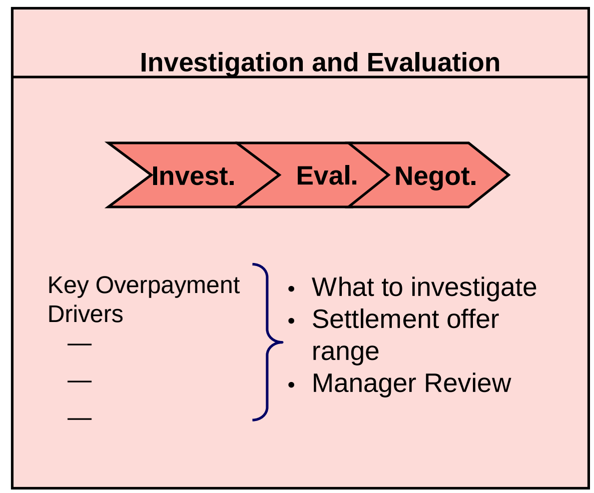
Assignment Triage



Adjuster Calibration



Case Management



Recent studies estimate that Auto Physical Damage represents over 20% of all claims leakage.

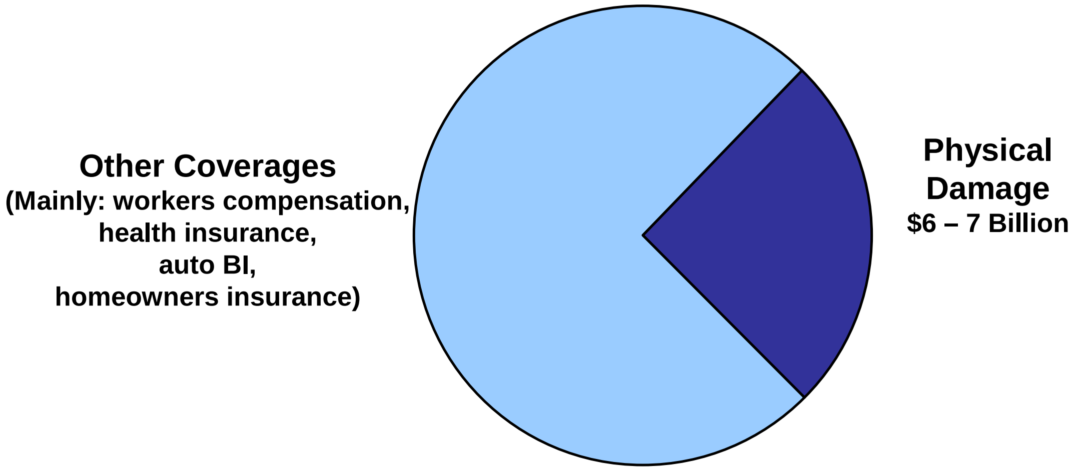
Losses from Claim Leakage

Property and Casualty Insurance

$ Million

100% = $30 Billion

Estimates



- Depending on the location, about **a third** to **a half** of all physical damage claims have substantial leakage
- Losses from '**Hard' fraud** represents a small part of this leakage
- Efforts to date to eliminate this leakage have had only **limited success**

Public Perceptions (All Americans, 2000 Survey):

- **24%** think it is okay to overstate insurance claims to make up for past premiums
- **35%** think it is acceptable to 'pad' claims to make up for deductible

In California*:

- **43%** of all vehicles inspected showed evidence of overwrite by an auto body shop
- Fraudulent repairs equaled **$586 per vehicle**

Source: National Insurance Crime Bureau; California Department of Consumer Affairs; California Bureau of Automotive Repair

* California Bureau of Automotive Repair, ongoing auto-repair re-inspection program. Numbers are as of December 31, 2001

PA's experience indicates that, despite efforts to improve performance, insurers generally continue significantly overpaying Auto Physical Damage claims by up to 9%.

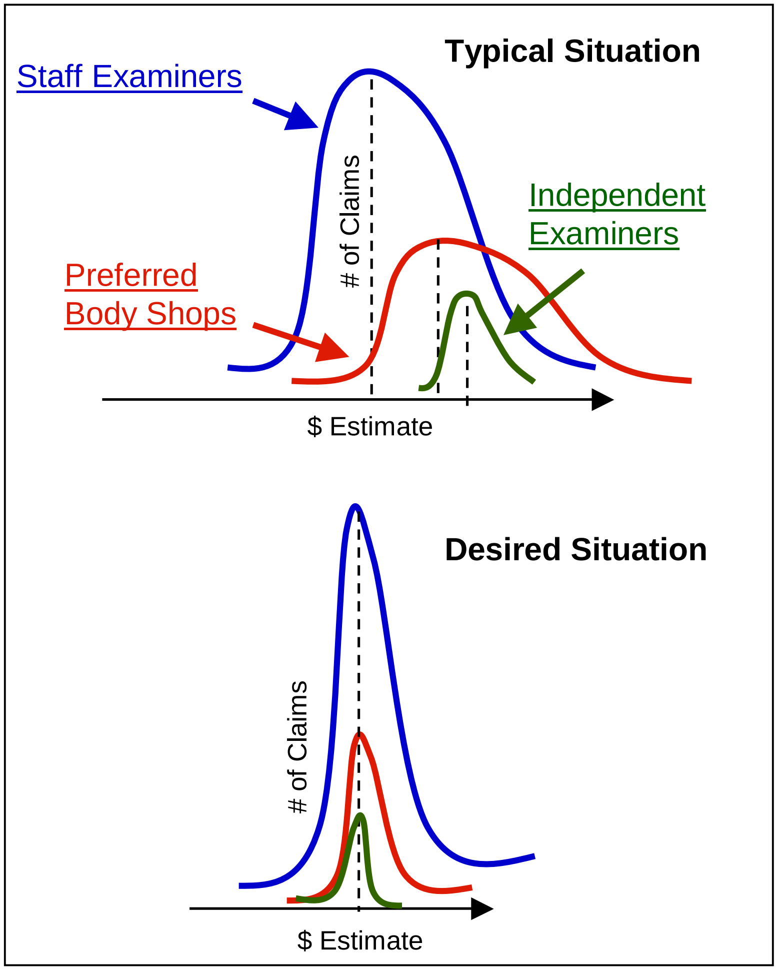
Motor Physical Damage: Distribution of Estimate Amounts

Illustrative

Performance Challenges

- Inadequate Controls (I.e., field reinspections and desk reviews)
- Inconsistent estimates across appraisers and appraiser types, due to broad range of skills
- Wide range of accident types and physical damage profiles
- Emphasis on customer service issues, sometimes at the expense of proper control and safeguards

Significant overpayment **(up to 9%*)** continues in Auto Physical Damage for most insurers



* Based on reinspections by client teams of estimates by "top" front-line claims personnel

Today's Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

Despite global use of training and corrective tools such as reinspections, significant differences still exist between adjusters estimating the claim from the same auto damage event.

Estimate Breakdown: Examples of Differences

Disguised

	Estimator 1	Estimator 2	Estimator 3
Bumper	Overhaul (1.1 hours), LKQ assembly for $187.50,2.5 paint hours	Replace (.7 hours) LKQ assembly for $130, 3.5 paint hours	Replace (.7 hours) LKQ assembly for $160, inspect LKQ (.3 hours) 2.5 paint hours
Lamps	Replace (.3 hours), LKQ tail lamp for $62.50	Replace (.5 hours), LKQ tail lamp for $100	Replace (.5 hours), LKQ tail lamp for $60.00
Rear Body Panel	Not written	Repair body panel (4 hours), 1.7 paint hours	Replace 5.2 hours), OEM body panel ($133.27), 1.4 paint hours
Real Floor Pan	Not written	Not written	Repair (2.5 hours) floor pan assembly, 2 paint hours
Quarter Panel	Repair body panel (7 hours) 2.2 paint hours	Repair body panel (6 hours), 2.2 paint hours	Replace (9.4 hours), LKQ quarter panel($160), 2.6 paint hours
Other	Pinstripe ($10), Flex Material ($12), Car Cover ($5), Hazmat ($4)	Set up & Pull (2 hours), stripe tape ($10), blend (.5 hours), seam sealer ($8), undercoating ($15)	Paint fuel door (.3 hours), flex coat ($12), Hazmat ($4), stripe tape ($15), corrosion protection ($15), seam sealer ($8)
Estimate Total (Indexed)	**100**	**140**	**200**

Note: LKQ refers to salvaged parts
Source: Test of sample claims with representative claims organization

When several adjusters were compared in a blind, quantitative process, startling variations were found in the appraisal performances.

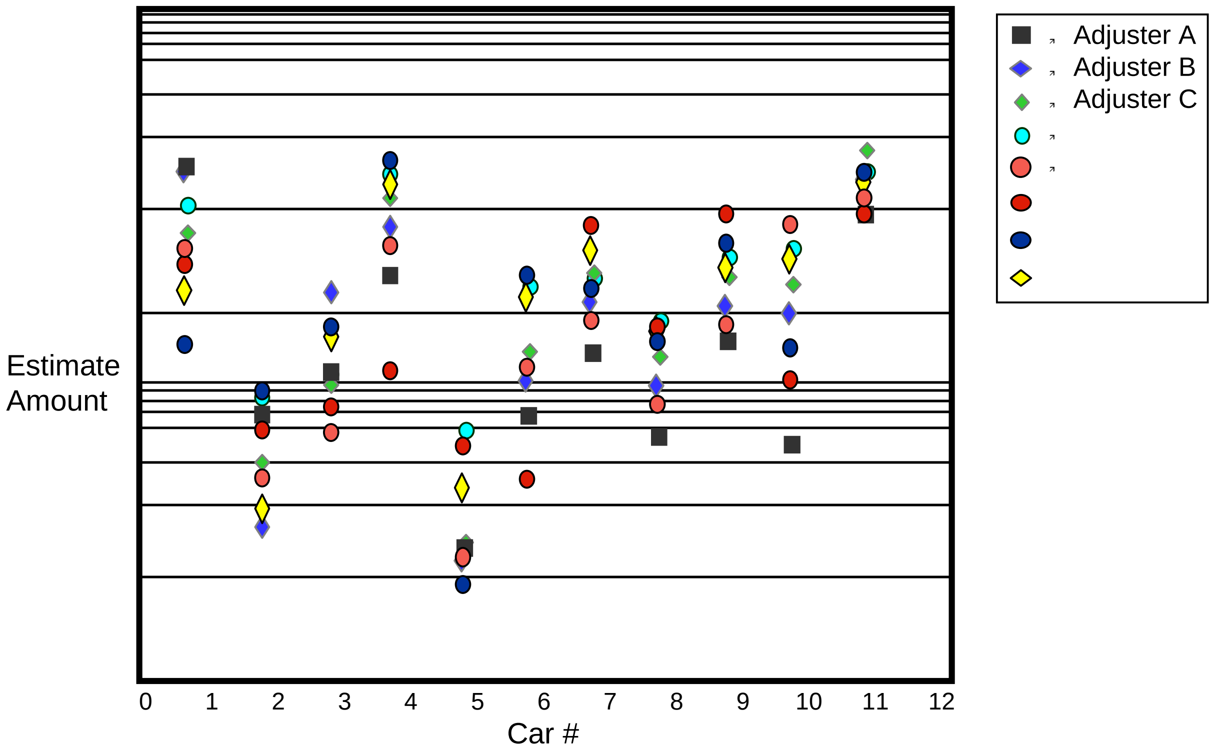
Measuring Estimator Performance

Disguised Example

- Damaged vehicles representing different ages, type and damage locations were collected in a central location
- Estimators were asked to provide repair estimates for every vehicle. Different types of estimators (i.e., insurer, DRP, non-DRP and independent) were chosen
- Estimators were not allowed to share data or observations
- Significant differences <u>between</u> groups (internal and external adjusters), as well as <u>within</u> each group were observed

Repair Estimates

Proportional Scale



Source: Test of sample claims with representative claims organization

This problem of high variance and overpayment in Auto Physical Damage losses can be addressed only through desired **Behavior Modification** on the part of appraisers and shops.

Importance of Behavior Modification

Key Issues in Auto Physical Damage Losses

- Claims losses are primarily determined by a large number of decentralized adjusters, with different performance levels
- The process depends significantly on the expertise of adjusters
- Heterogeneity of claims implies difficulty in designing predictive modeling tools for several risk segments
- Field reinspections and desk reviews tend to be marginally effective because the resources are limited and it is difficult to allocate them effectively

The key objective of any corrective strategy has to be **Behavior Modification** through proper controls and associated training, communications, incentives, etc.

One tool used for control is the reinspections of estimates. Unfortunately, reinspections do not identify most overpayments, nor do they always achieve the desired behavior modification.

Key Issues with Reinspections

- **Unsophisticated claim selection procedure**: reinspection resources tend to focus on the largest $ amount claims, or those that are conveniently accessible on a given day, rather than on those with the maximum potential likelihood of overwriting.
- **Limited reinspection resources**: Reinspection effectiveness is generally limited because of limited resources, and appraisers'/shops' knowledge of these limitations
- **No single reinspection process**: reinspection process may vary between appraisal types (different for staff/internal appraisers from others), making it difficult to evaluate appraisers consistently
- **Inefficient results capture**: Reinspection results are not captured in a single place, generally, and may be saved in different systems for different markets. This makes it difficult to analyze the data for identifying trends or developing corrective strategies

Estimates are typically selected either randomly or through a pre-set and generally known dollar limit, which means that:

- Several of the most problematic claims do not get reinspected, and so **savings are not maximized**
- Reinspections tend to create **limited desired modification in partner/vendor behavior**

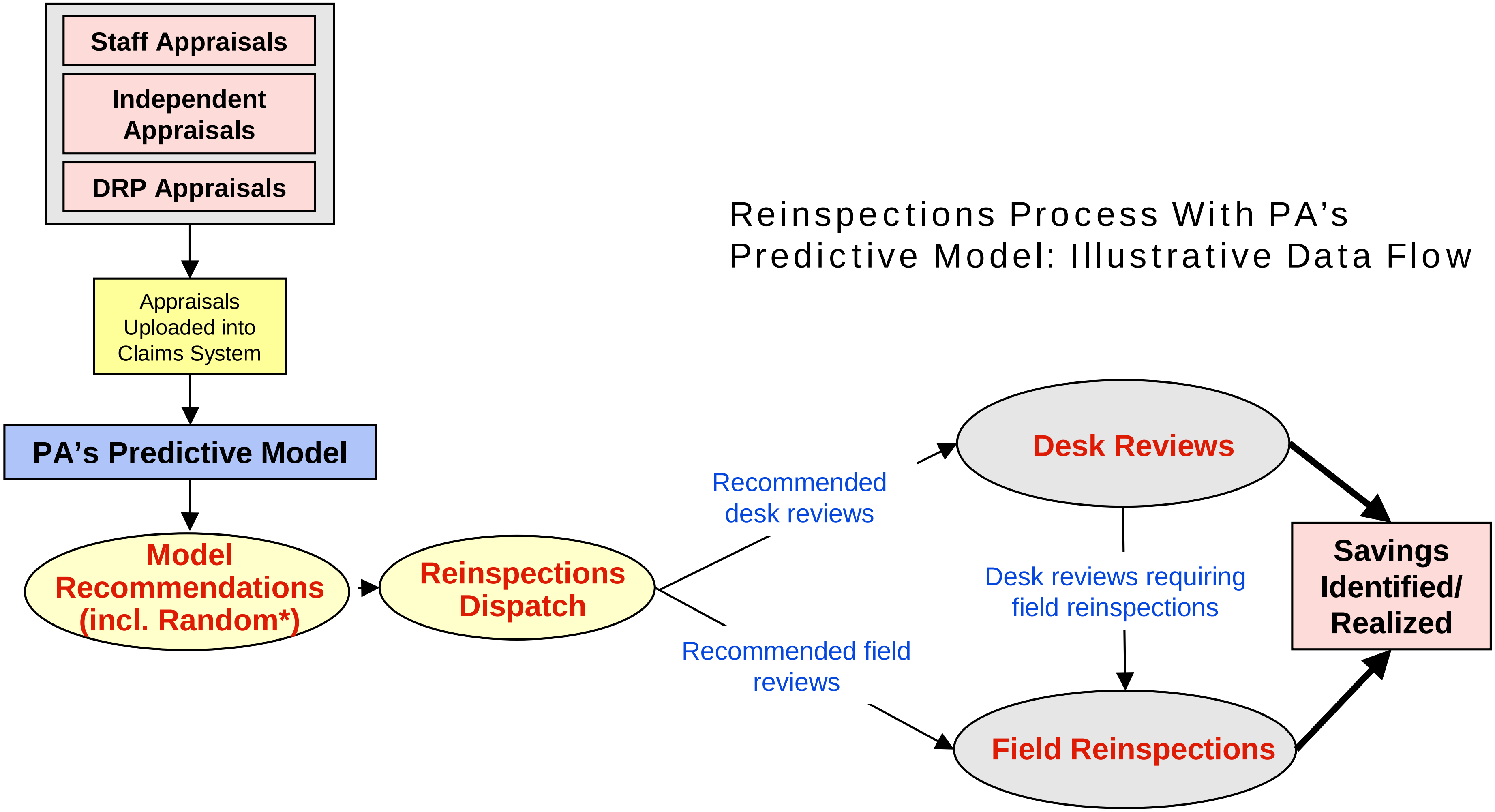
PA tackles reinspections differently, by using a sophisticated model to estimate the overwriting probability of each estimate, and then allocating limited reinspection resources accordingly.

PA's Approach to Targeting Auto Physical Damage Leakage through a Probabilistic Model

- "Expert" decision process information is gathered from experienced adjusters
- Historical claims data is collected for model
- Model is designed, based on a probabilistic Bayesian framework, using historical claims data
- Model is trained on historical reinspection results
- Links are created to allow model to function dynamically with fresh daily claims data, and to 'learn' with data
- Model results are fed into reinspection resource allocation process, resulting in increased 'hit rates' and savings identified
- Over time, process leads to behavior modification and a significantly improved estimation process

Our predictive modeling solution directly integrates with the client claims management system, creating a tool and a process for ongoing claims evaluation and selection for reinspections.

Reinspections Process With PA's Predictive Model: Illustrative Data Flow



*** Typically, a fixed % of random recommendations is added to the mix in order to ensure reinspections of all shops and appraisers over a period of time**

Today's Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

Our methodology is based on **Bayesian Networks**, which provide consistent statistical treatment to both hard data and experts' assumptions about the likelihood of overwriting.

What are Bayesian Networks?

- **BNs are models that represent uncertainty in our knowledge:**
  - **Uncertainty from experts' knowledge**
  - **Uncertainty about the domain being modeled**
  - **Uncertainty from the knowledge engineer (aka model builder)**
  - **Uncertainty about accuracy and availability of the knowledge**

(Source: Russel Greiner, U. of Alberta, Canada)

- **BNs are conditional probability models**
  - **Bayesian inference consists of updating prior beliefs as new information becomes available – BNs are able to "learn" from new data**
  - **Rev. Thomas Bayes (1702-1761) is responsible for the Bayes' Theorem, which is the basis for conditional probability theory**

- **BNs are graphical causal models**
  - **Bayesian methods have been available for a long time, but tend to be computationally difficult**
  - **Only in the last 10 years has software become good enough to develop useful applications that can be solved in a reasonable time**

Bayesian networks use conditional probabilities to represent what we usually call “common sense” – they have found a warm reception in Artificial Intelligence circles

Why Bayesian?

Rev. Thomas Bayes stated in 1761 that

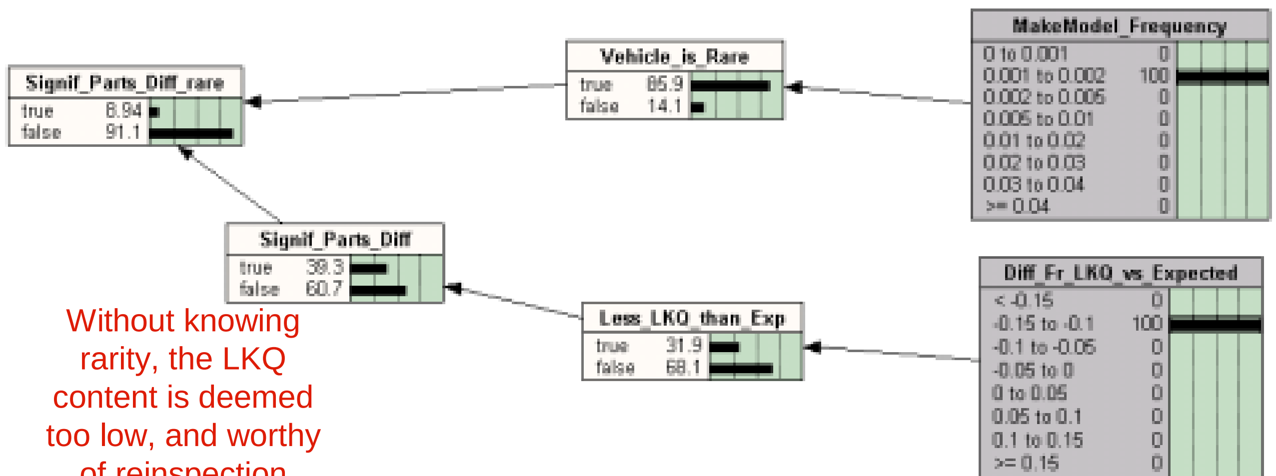
$$P(A \mid B) = P(B \mid A) * P(A) / P(B)$$

The Bayes’ Theorem allows easy manipulation of conditional probabilities, which are critical to develop “reasonable” systems.

An example from the Physical Damage claims world:

The fact that there is a lower than expected amount of used parts in a vehicle repair estimate increases the probability of overwriting. However, if this vehicle is a rare model, then the probability decreases, because we don’t expect to easily find LKQ parts for it.

Conditional on rarity, however, the LKQ content seems appropriate



Rare vehicle

LKQ content lower than expected

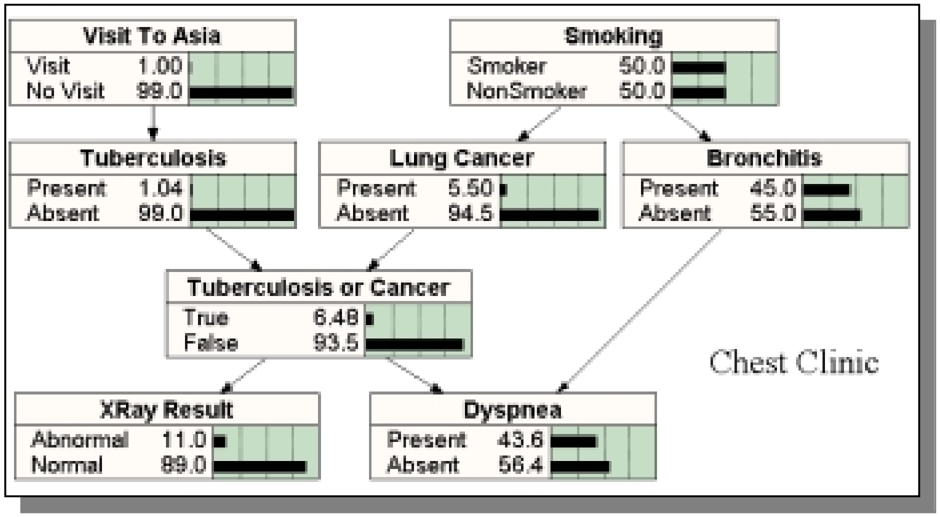
Without knowing rarity, the LKQ content is deemed too low, and worthy of reinspection

Partial view of the predictive model

Bayesian Networks have been applied in diverse areas of engineering, medicine and business

Some examples of applications:

- Medical diagnostics
- Language understanding
- SPAM filtering
- Machine vision
- Fraud detection
- Operational risk management
- Pharmaceutical drug discovery
- Online customer service
- Predictive maintenance
- Fault analysis for complex machinery



The classic "Chest Clinic" Bayesian Network for medical diagnosis*

* (From Norsys's Netica tutorial, based on S. L. Lauritzen and D. J. Spiegelhalter. *Local computations with probabilities on graphical structures and their application to expert systems*. Journal of the Royal Statistical Society, Series B (Methodological), 50(2):157-224, 1988).

Benefits of Bayesian Networks:

- BNs **replicate cause-and-effect relationships** seen in real life, and identified by experts familiar with the industry
- BNs rely on a rigorous probabilistic structure, so **conflicting evidence can be weighed appropriately** instead of biasing the results
- BNs can generate results with **incomplete data**, and account for uncertainty in the inputs appropriately
- BNs stay current because of their **ability to learn** from new data as it is generated, reducing maintenance requirements
- BNs keep **all assumptions visible** and make them easy to share with users, even those without technical training

BNs are ideally suited for Auto Physical Damage Claims analysis

- **Claims data is usually incomplete**, and traditional statistical tools do not handle incomplete information well
- **Experts can quickly identify overwriting** in auto insurance claims, but have a difficult time making their process explicit – modeling with BNs allows companies to capture experts' mental models and replicate them across the company, standardizing best practices
- Historical reinspections results are usually **patchy and unstructured** – a system based on BNs can start with a model created from experts' opinions, and validate them as new results are collected
- Early in the development of the system we explored the possibility of using a more traditional method, such as logistic regression, to estimate the probability of overwriting
  - However, we found logistic regression too rigid for this application, since it has more stringent data requirements
  - Additionally, incorporating the experts' knowledge requires a hierarchical approach to model development, which is difficult to represent using traditional regression tools

Bayesian networks are superior to alternative modeling methodologies.

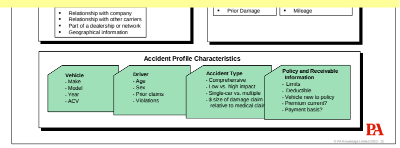
Features	Gut feel and intuition	Regression models	Rules-based expert systems	Neural networks	Bayesian networks
• Can incorporate expert knowledge	✓	✓	✓		✓
• Can handle quantitative data inputs		✓		✓	✓
• Cause and effect explicitly represented		✓			✓
• Non-linear relationships can be easily represented				✓	✓
• Continuous (fuzzy) logic allowed				✓	✓
• Can handle incomplete data sets				✓	✓
• Can learn from data				✓	✓

The predictive model is built by combining expert knowledge with quantitative data.

Model Design and Construction

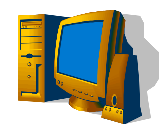
Illustrative

Claims and Policy data…
Appraiser information…
Estimate details..
Repair Shop Characteristics…



Expert Interviews…
PA's experience…

Qualitative model



Quantitative model

The qualitative model represents how different inputs potentially relate to each other

These relationships are represented graphically and validated through interviews

The quantitative model turns the relationships between inputs into probabilities, which are calibrated using claims data and other sources

The key starting point is to tap into the accumulated knowledge of your Material Damage claims experts.

Breaking down the “expert” decision process

“I know a <u>good repair shop</u> when I walk in the door”

“These <u>new</u> staff appraisers always seem to miss the <u>little things</u> that can add up”

“I know a Total Loss <u>when I see it</u>”

“I always jump on those positive supplements that <u>come in late</u>”

“If they all just did what I tell them to do, we wouldn’t need you consultants”

Bayesian networks provide an explicit way of balancing expert knowledge with “hard” data

Illustrative

Prior distribution from expert knowledge is uniform

Variable	
0 to 10	12.5
10 to 20	12.5
20 to 30	12.5
30 to 40	12.5
40 to 50	12.5
50 to 60	12.5
60 to 70	12.5
70 to 80	12.5
40 ± 23	

Model reacts quickly to new data when there is little confidence on prior knowledge…

Confidence on prior expert knowledge is equivalent to 1 data point

Variable	1 data point	10 data points	100 data points	1000 data points
0 to 10	6.25	1.14	0.12	.012
10 to 20	6.25	1.14	0.12	.012
20 to 30	6.25	1.14	3.09	2.81
30 to 40	6.25	46.6	46.7	45.3
40 to 50	56.3	46.6	48.6	49.2
50 to 60	6.25	1.14	1.11	2.71
60 to 70	6.25	1.14	0.12	.012
70 to 80	6.25	1.14	0.12	.012
	42.5 ± 17	40 ± 8.9	39.8 ± 6.8	40.2 ± 6.7

Confidence on prior expert knowledge is equivalent to 100 data points

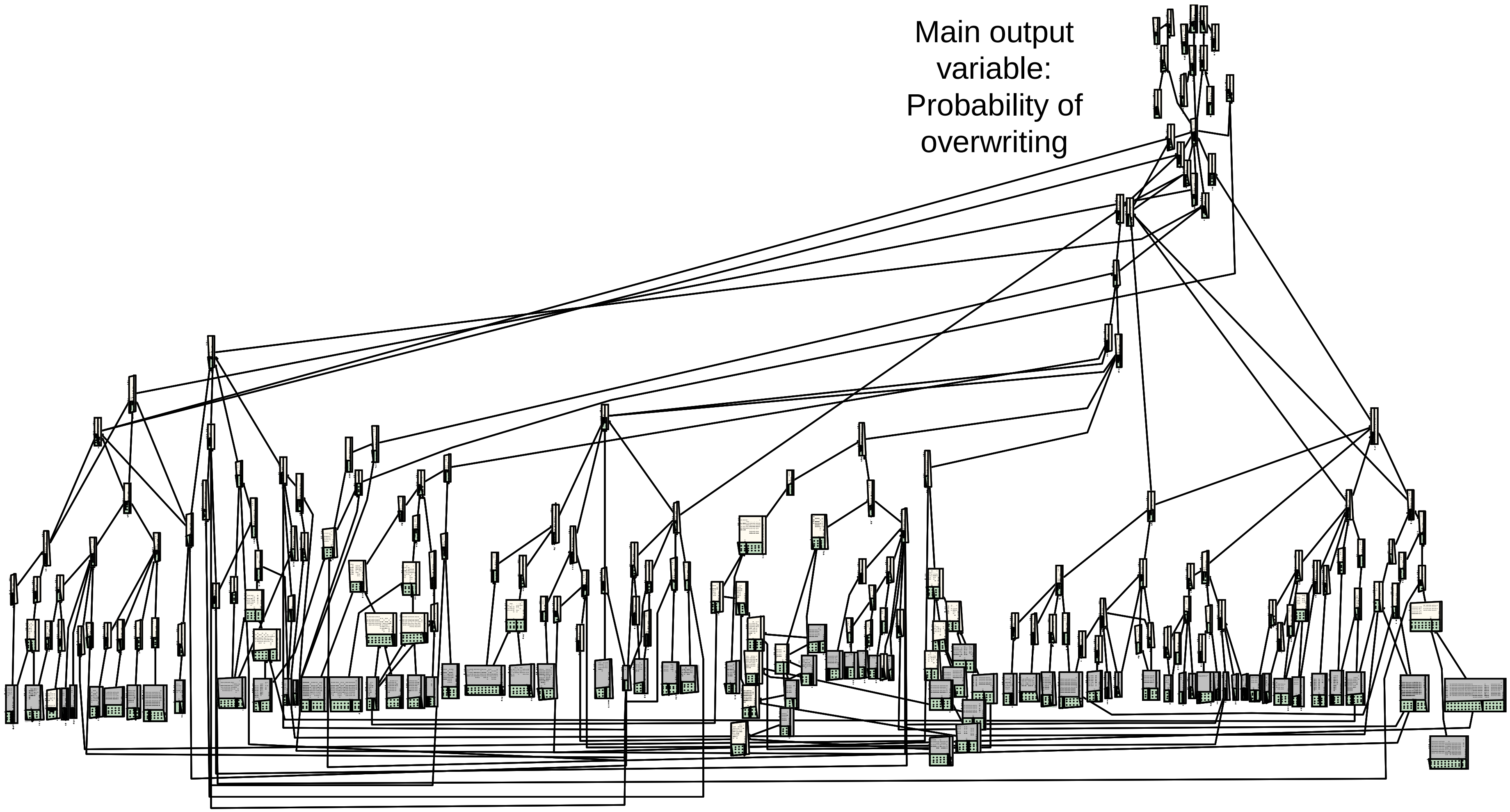
Variable	1 data point	10 data points	100 data points	1000 data points
0 to 10	12.4	11.4	6.25	1.14
10 to 20	12.4	11.4	6.25	1.14
20 to 30	12.4	11.4	7.75	3.68
30 to 40	12.4	15.9	29.8	42.3
40 to 50	13.4	15.9	30.8	45.9
50 to 60	12.4	11.4	6.75	3.59
60 to 70	12.4	11.4	6.25	1.14
70 to 80	12.4	11.4	6.25	1.14
	40 ± 23	40 ± 22	39.9 ± 17	40.2 ± 9.4

If there is more confidence on prior knowledge, the model reacts more slowly to new data

Note: Data points in this example come from a random Normal(40,5) sample

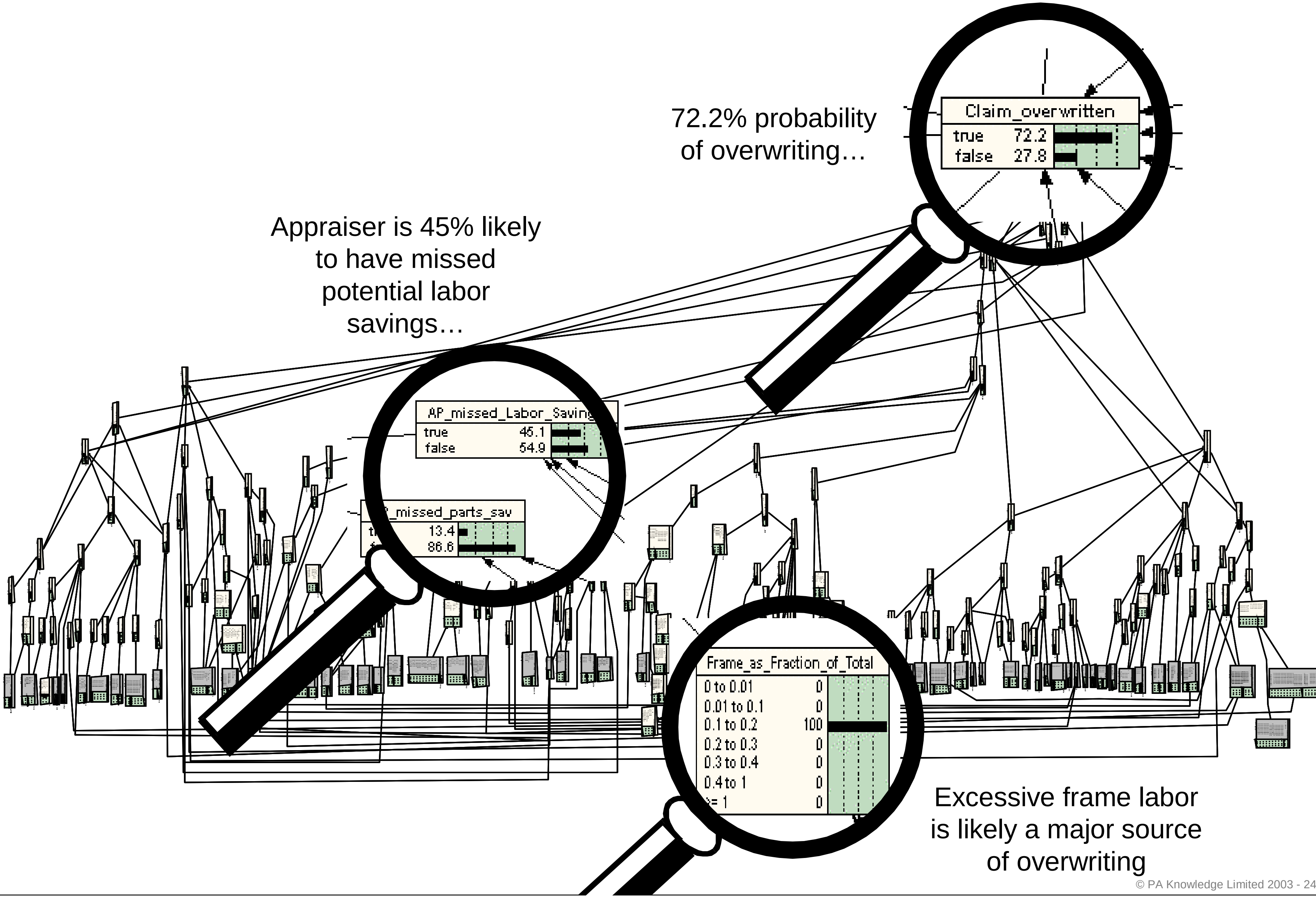
Expert interviews provide the base for creating the Bayesian Network, which captures a rich set of cause-effect relationships between different data inputs

High-Level auto repair model diagram



80+ data inputs, combining hard data in non-linear cause-effect relationships elicited by experts

For example, this claim’s probability of being overwritten is 72.2%, but this value is the result of a complex combination of inputs



72.2% probability of overwriting…

Appraiser is 45% likely to have missed potential labor savings…

Excessive frame labor is likely a major source of overwriting

Policy, claim and other data is gathered to calibrate the predictive model.

Data sources for possible inclusion in model

Examiners and Repair Facilities

Examiner Group

- Independents
- PRS Shops
- Staff Examiners

Characteristics

- Date of joining
- Prior incidences
- Experience level
- Expertise focus
- Training level
- Historical ratios

Repair Facilities

- Relationship with company
- Relationship with other carriers
- Part of a dealership or network
- Geographical information

Appraisal information

Labor
- Labor Rates
- Types of Labor
- Labor Hours

Parts
- Amount
- Type and frequency
- Repair/replace rules

Total Loss
- Threshold Amount
- Vehicle Valuation
- Prior Damage

Miscellaneous
- Repair guidellines
- Betterment
- Mileage

Accident Profile Characteristics

Vehicle
- Make
- Model
- Year
- ACV

Driver
- Age
- Sex
- Prior claims
- Violations

Accident Type
- Comprehensive
- Low vs. high impact
- Single-car vs. multiple
- $ size of damage claim relative to medical clair

Policy and Receivable Information
- Limits
- Deductible
- Vehicle new to policy
- Premium current?
- Payment basis?

Model calibration can be done with a small number of reinspection results, from which the model will continue to learn over time

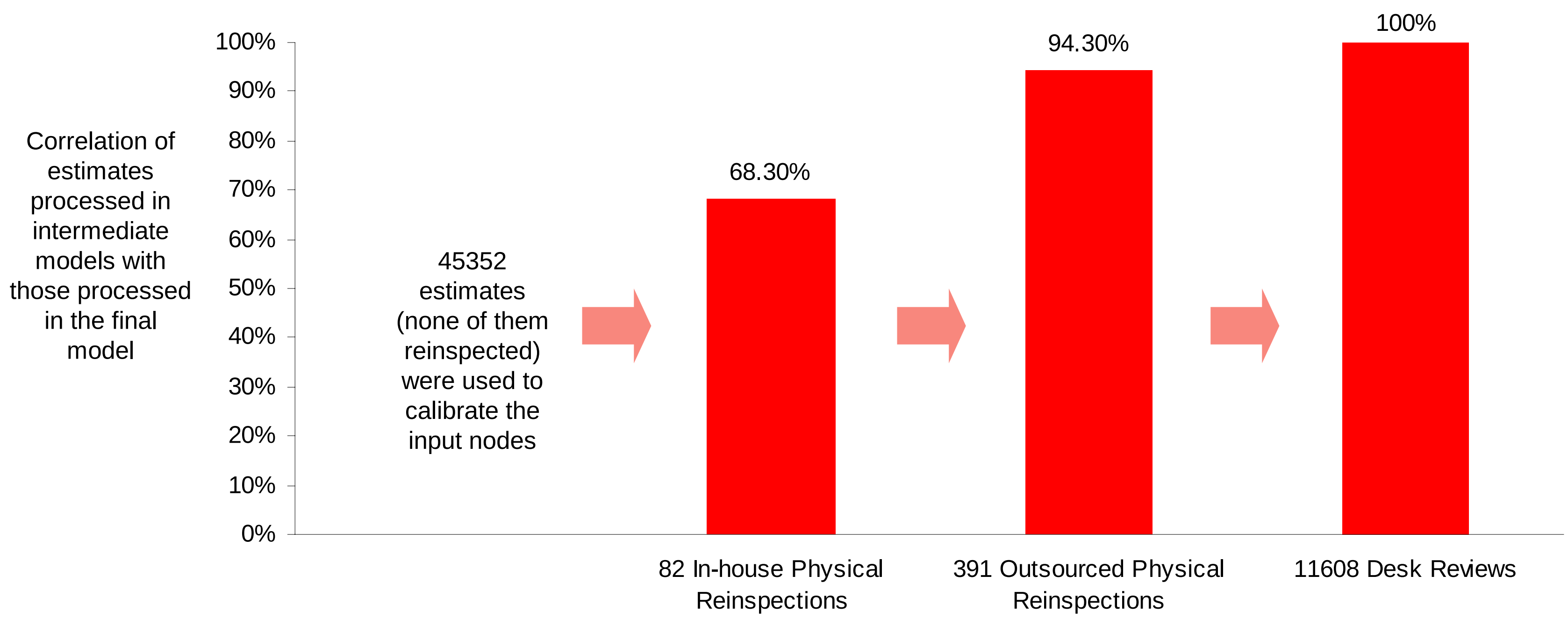
Model Calibration Process

Step 1:

Population Learning
This calibrates model input nodes, using estimates without re-inspection data

Step 2:

Performance Learning*
This step uses re-inspection data to calibrate relationships between variables



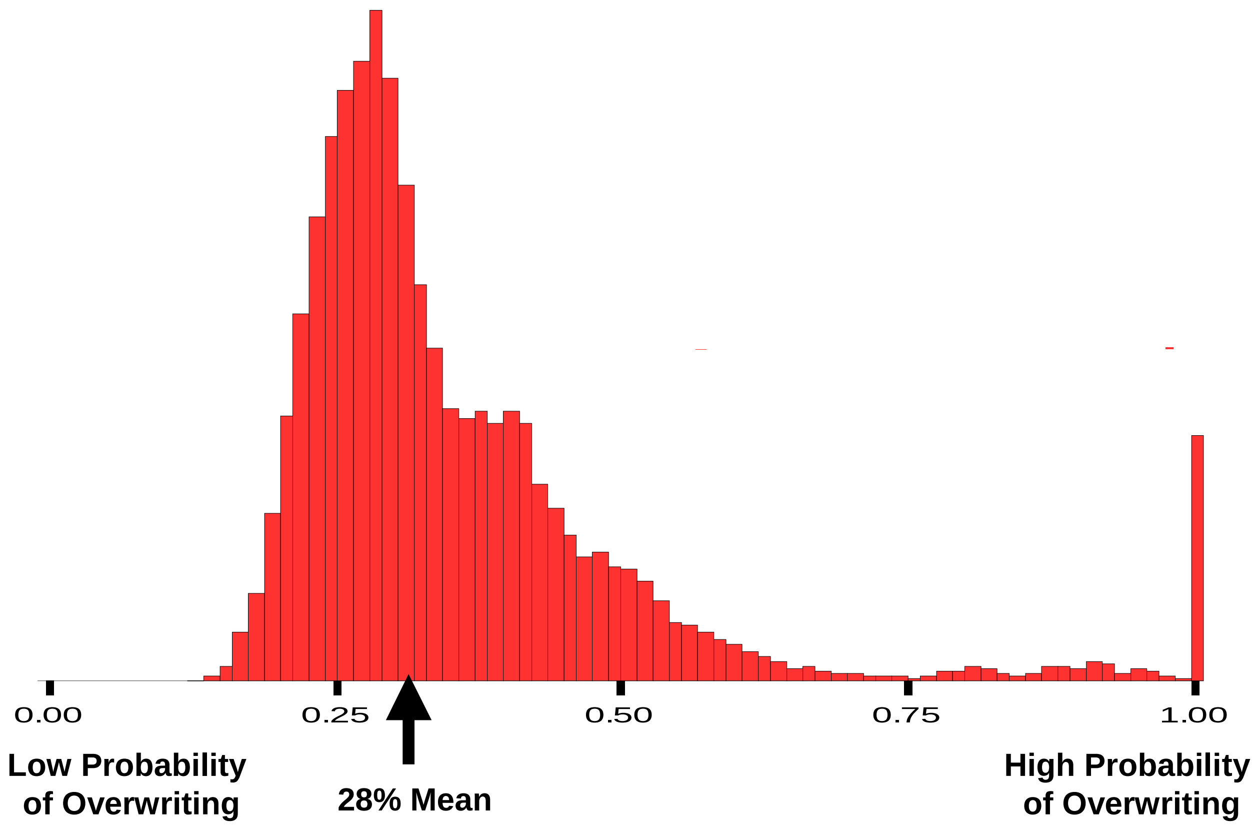
* Performance Learning relies on an iterative optimization algorithm called "Expectation Maximization", or EM. This algorithm maximizes the likelihood that the data used was generated by the model in question, each iteration modifies the conditional probability tables in the model to increase the likelihood that the model would generate the dataset.

Source: PA's Experience with model calibration

Once the model is calibrated, it is used to calculate a probability of overwriting for all the claims processed by the system.

Distribution of probabilities for historical estimates generated by the predictive model

Disguised

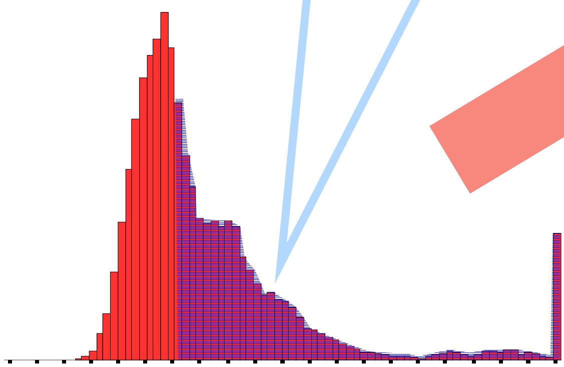


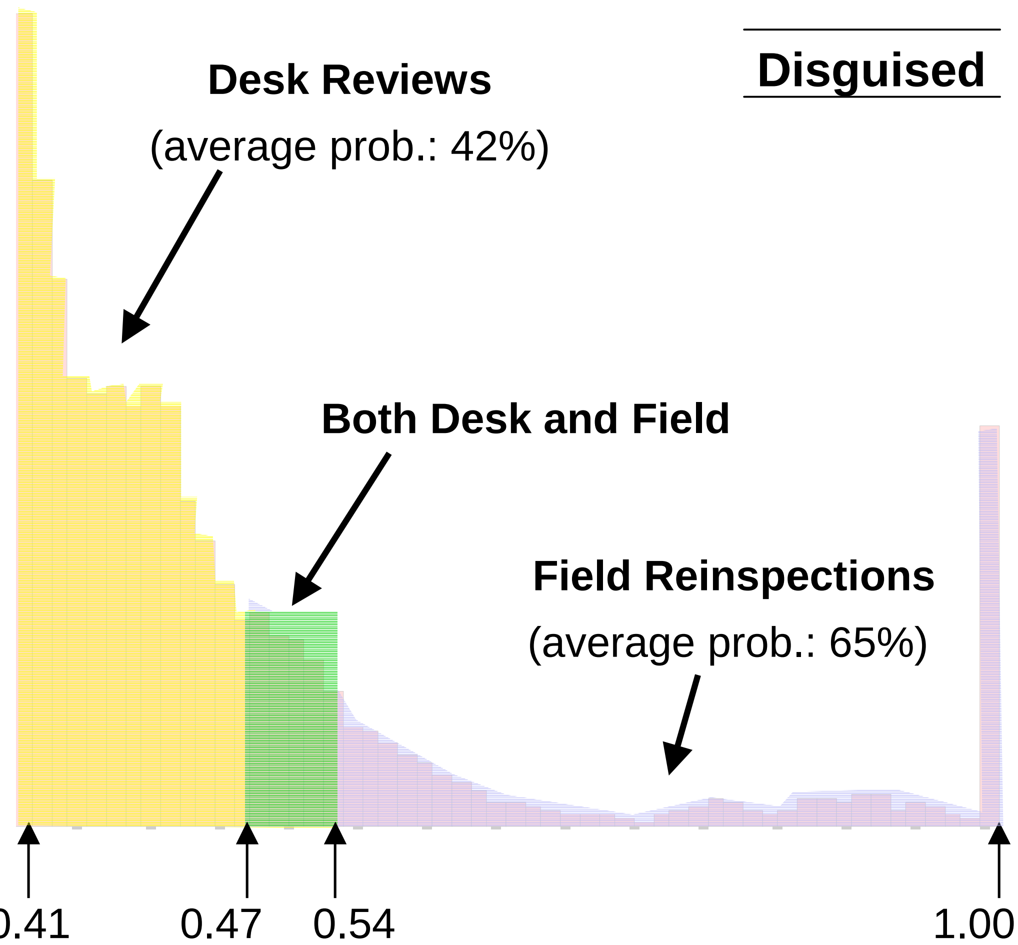
Using the model probabilities, reinspection hit rates on assigned claims – both field and desk – is substantially greater.

Disguised

Allocation of Capacity to Model Recommendations

The shaded area includes the top % of all estimates that can be reinspected by projected capacity



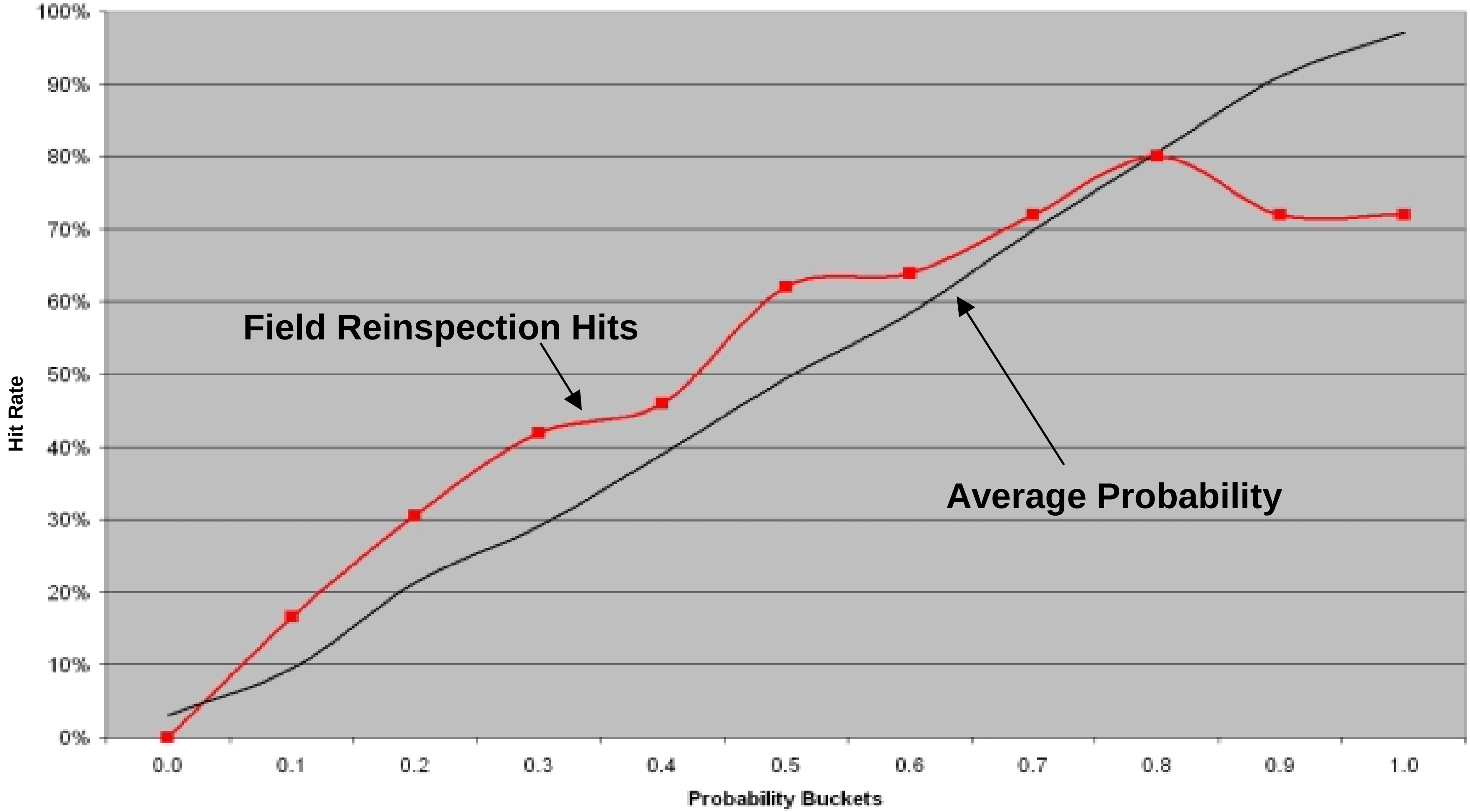


Claims allocation is based on company's reinspection capacity

The actual field hit rate closely matches the model generated probabilities.

CPMS results

Disguised client example



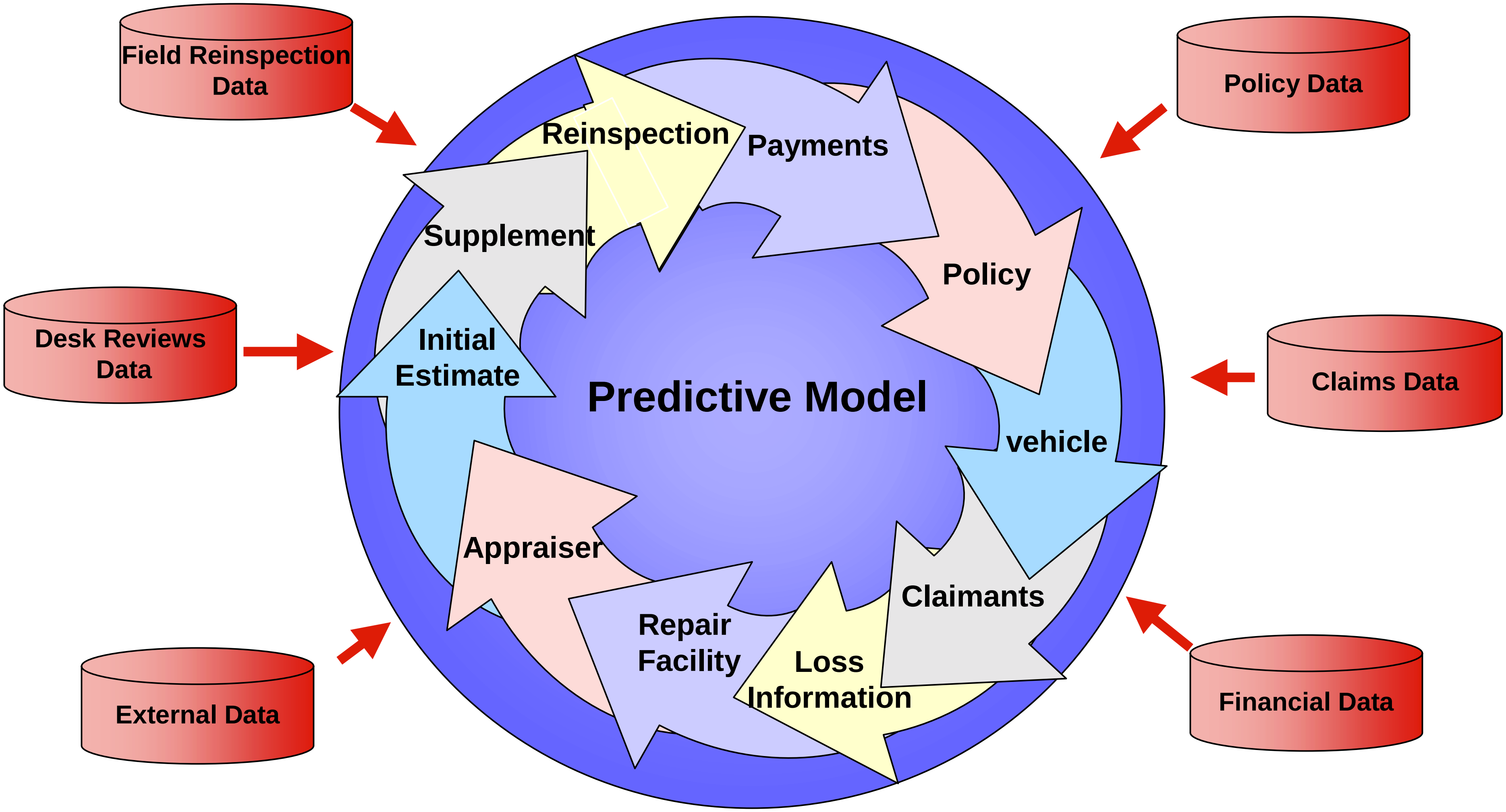
Today's Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

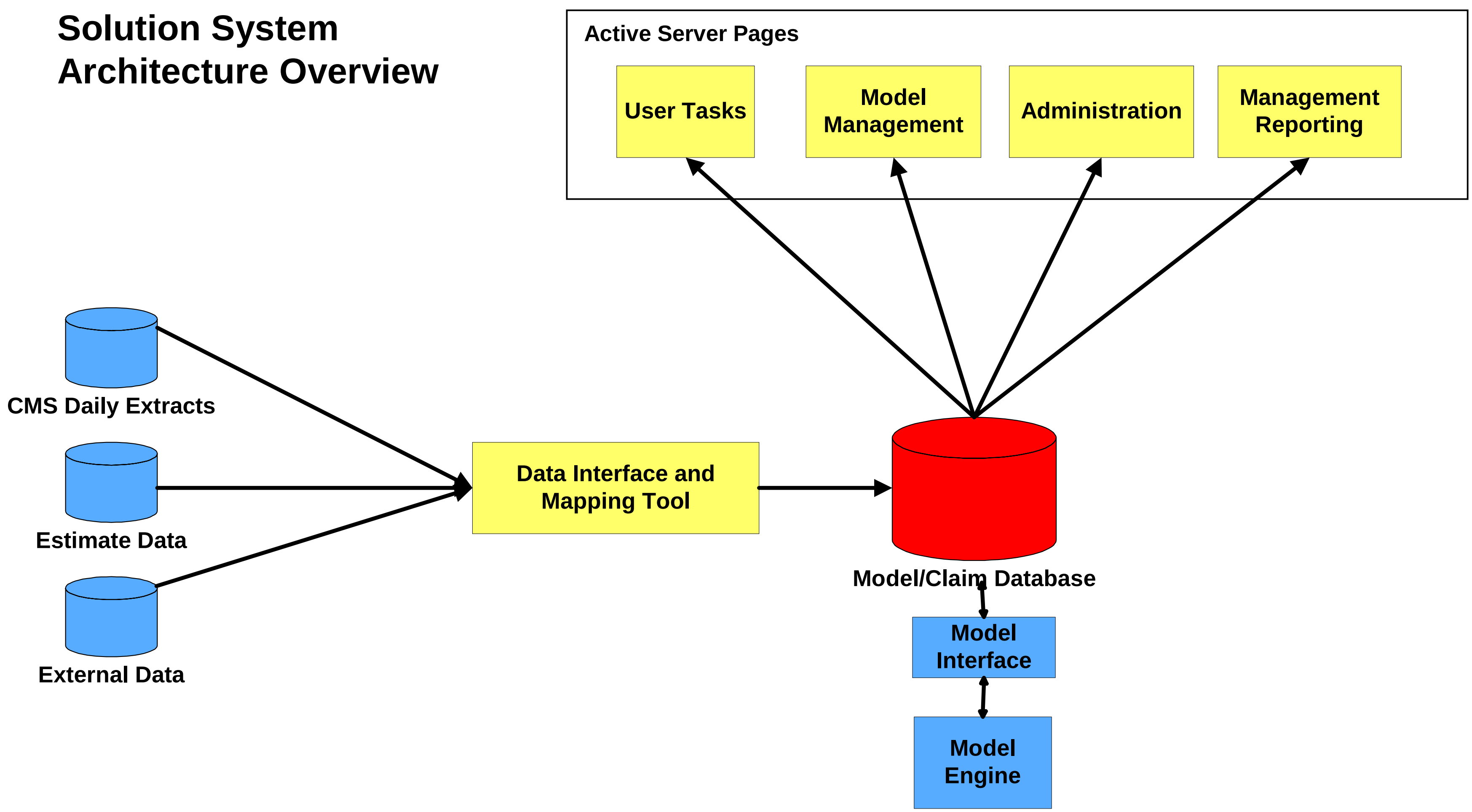
A single system architecture which is necessary to enable the multiple data inputs for the models.

Data Inputs for Predictive Models

Illustrative

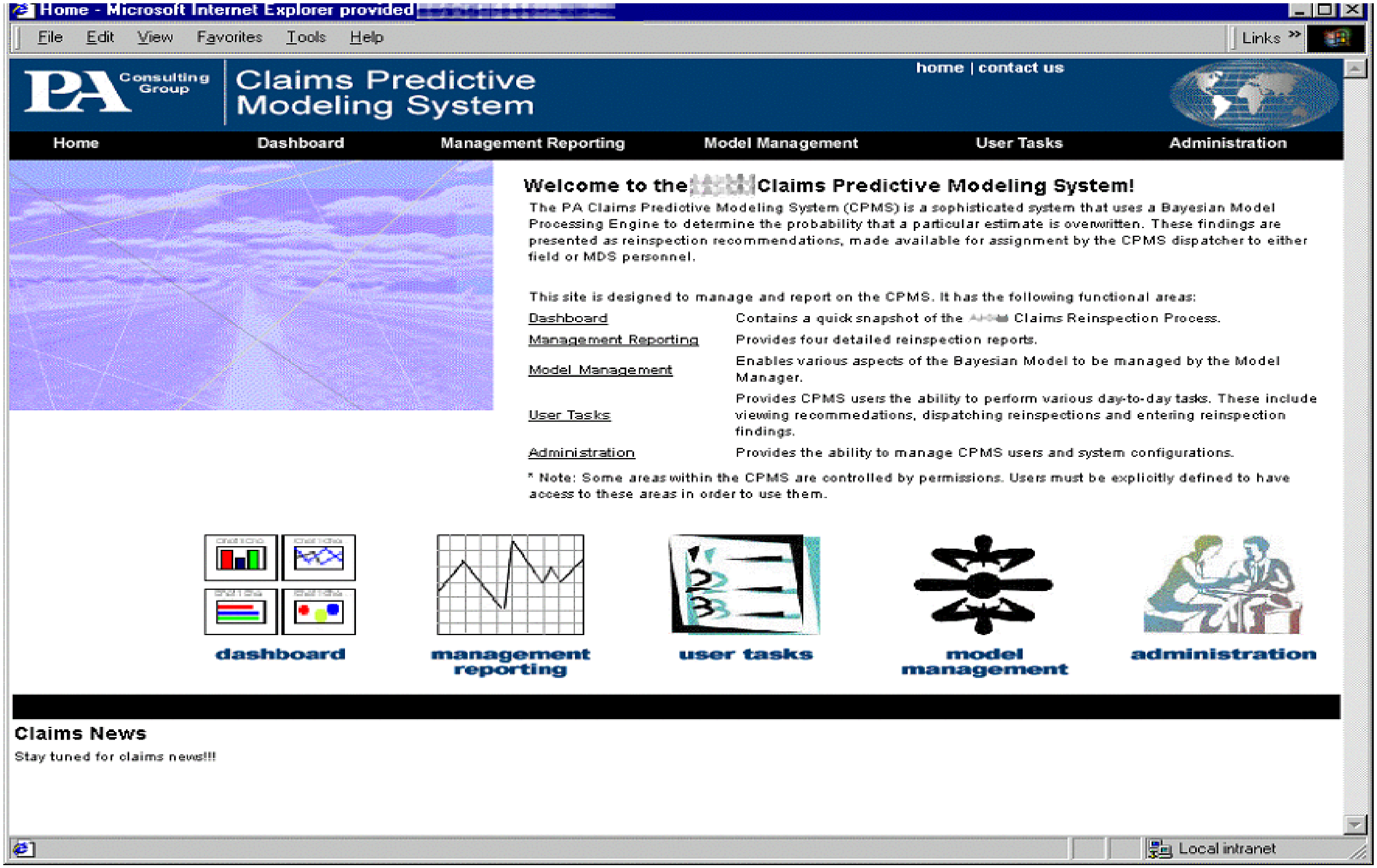


This single system architecture allows for a robust system that, while fully integrated, has a focused purpose.

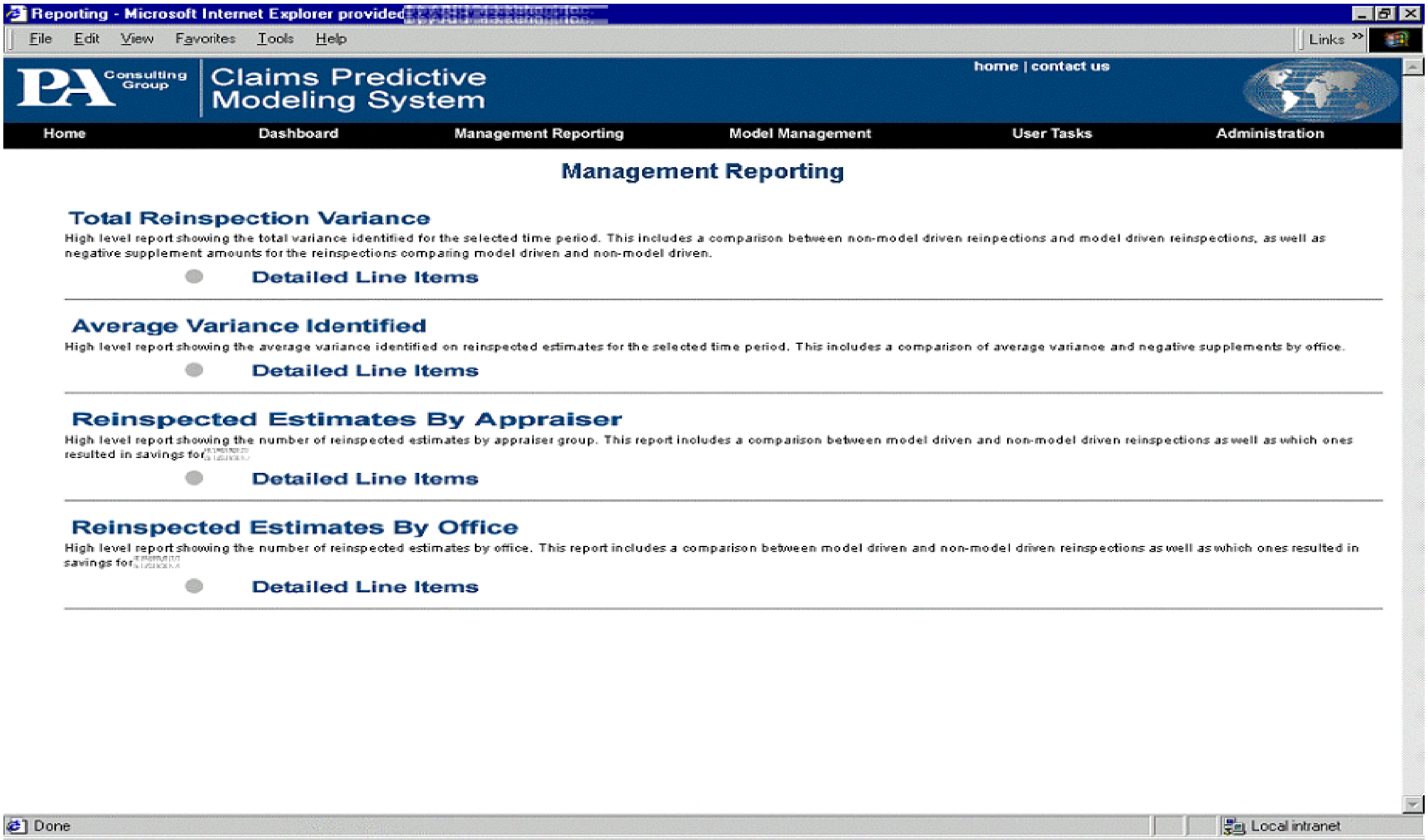


The model is part of an integrated IT solution required for daily analysis of claims and dispatch of reinspectors.

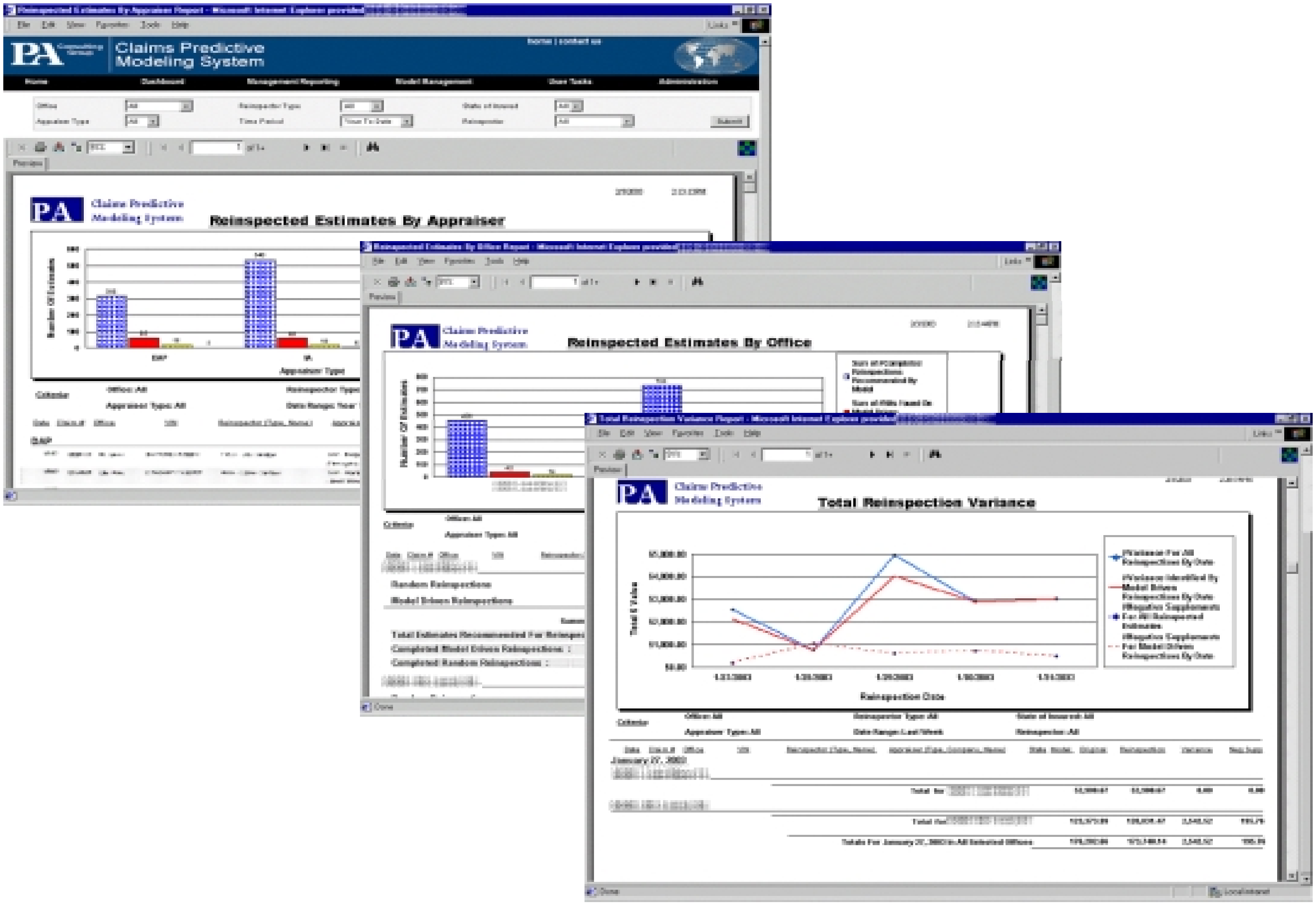
Included with our solution is a user interface that allows management of both the model and the information needed to achieve the benefits expected from your reinspection process….



…including standard and customized management reporting…



…at the Appraiser, Office and Region level, including individual reinspection details.



Today’s Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

A preliminary workplan for the Optimizing Reinspections effort would be divided into 3 phases: planning, detailed design, and pilot launch.

Optimizing Reinspection Process: Auto Physical Damage Claims

Illustrative

	Phase 1 **Planning**	Phase 2 **Solution Design**	Phase 3 **Pilot Implementation**
Timing	2 – 3 Weeks	9 – 11 Weeks	TBD
Key Activities	▪ Identify pilot market/unit ▪ Understand availability of data ▪ Review sources of data, for historical and ongoing data ▪ Obtain historical claim data ▪ Begin analyzing claim data ▪ Lay out detailed claims process ▪ Understand relevant technology infrastructure, and identify links to be created ▪ Develop high-level predictive modeling plan ▪ Develop preliminary reporting requirements ▪ Develop resource requirements - People, costs, systems ▪ Report to management	▪ Design database for pilot market/unit ▪ Design, secure and 'clean' data (e.g., identify and correct for missing variables and outliers) ▪ Develop process and technology for collecting ongoing data into model ▪ Create process for using model results in decision-making (e.g., for determining when to re-inspect) ▪ Develop training program ▪ Design reinspection process around model ▪ Create technology plan ▪ Define pilot implementation team ▪ Design first generation model ▪ Design key management reports ▪ Identify key success metrics ▪ Perform user testing ▪ Report to management	▪ Implement revised reinspection process ▪ Ensure complete model inclusion in decision-making process for claims ▪ Identify and articulate key learnings from model results ▪ Chart progress against objectives ▪ Develop rollout plan for rest of the organization

Rollout

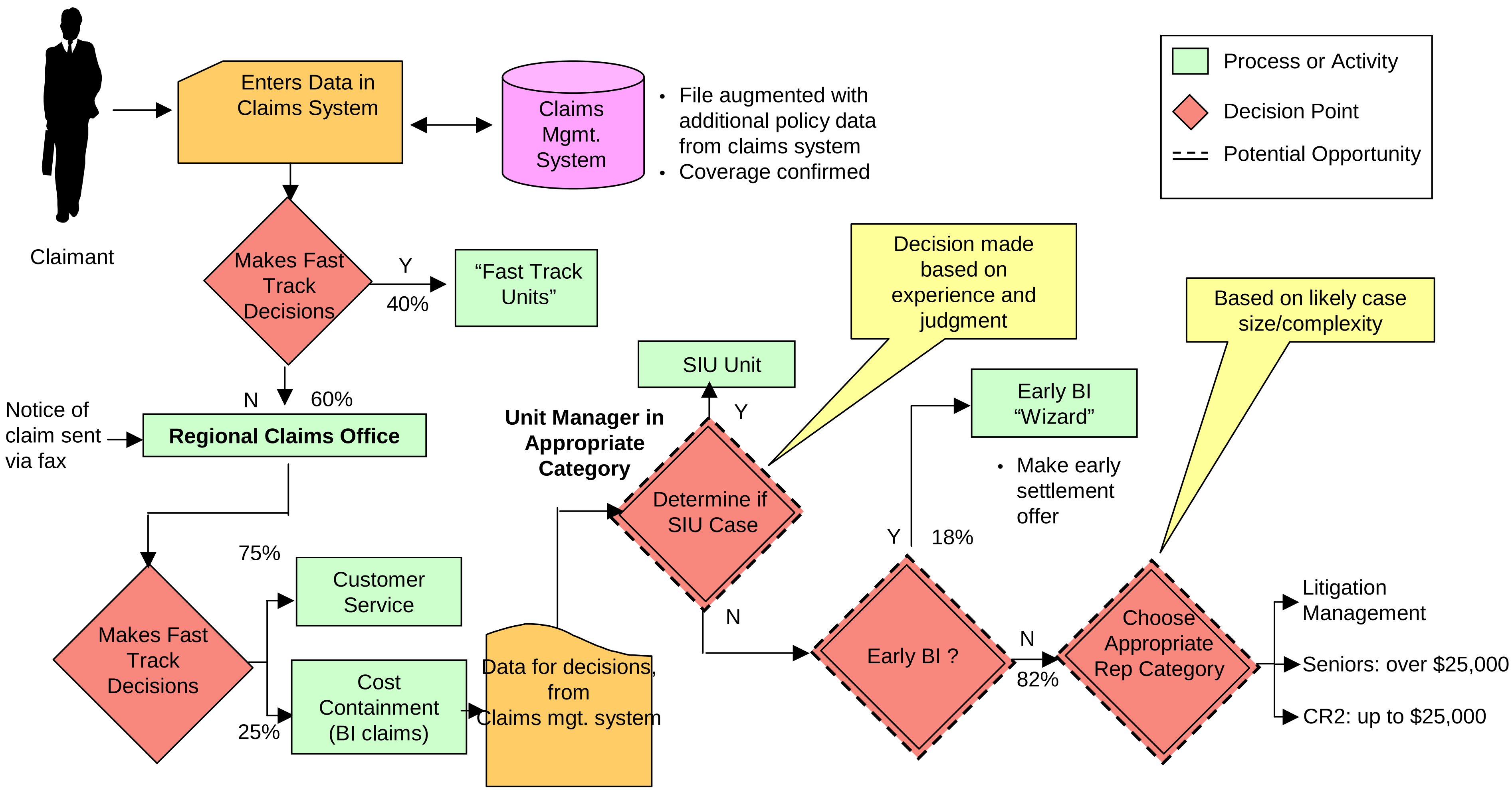
Today's Agenda

- Defining the opportunity
- Approach to Auto Physical Damage claims leakage reduction
- Modeling methodology
- Technical Design
- Typical Process
- Appendix: Application to Bodily Injury Claims

Integrating the models into the claims process begins with a detailed understanding of the claims process in the organization.

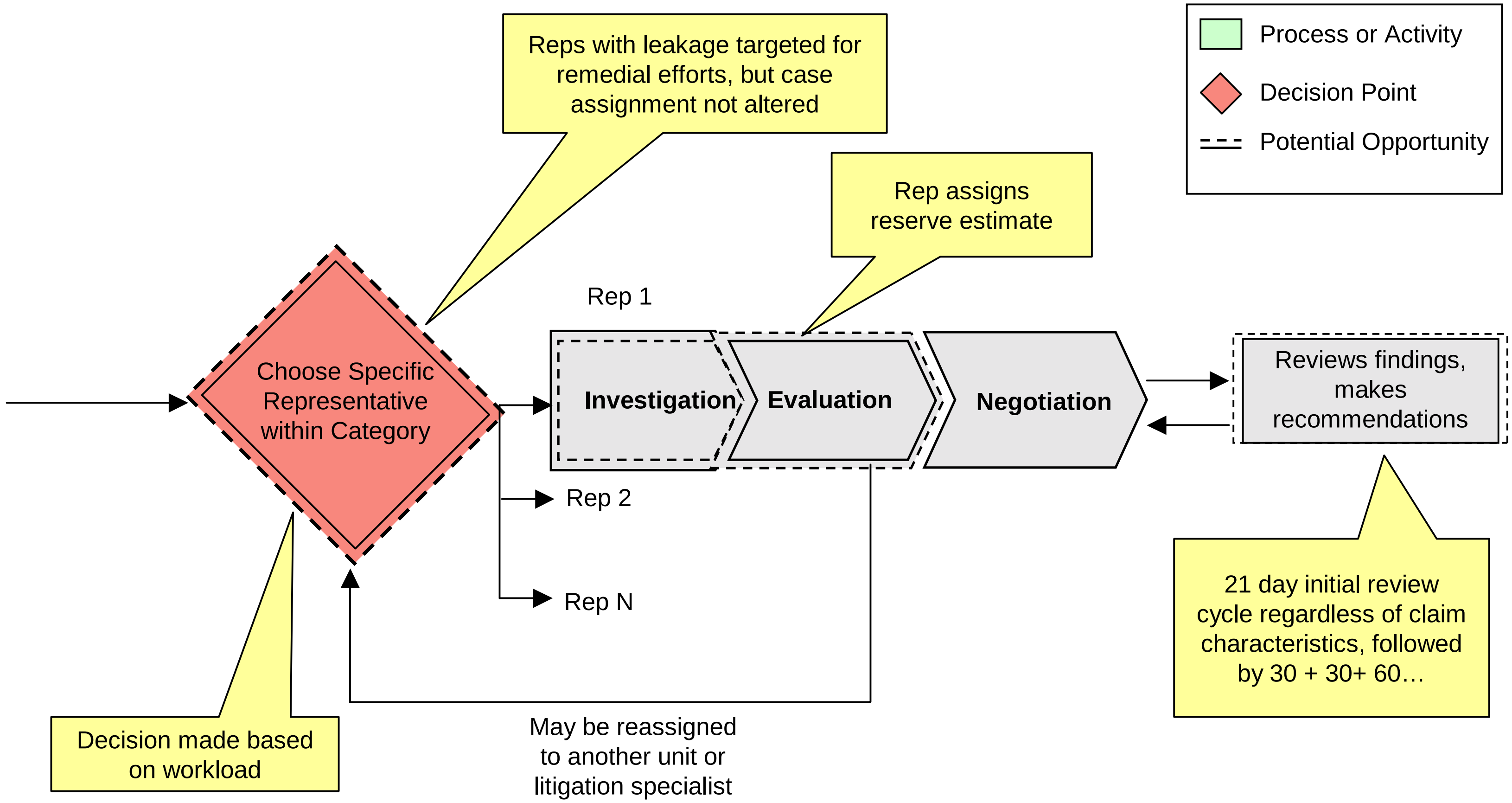
Overview of Generic Claims Process
(and Potential Areas of Improvement)
Automobile BI

Illustrative Only



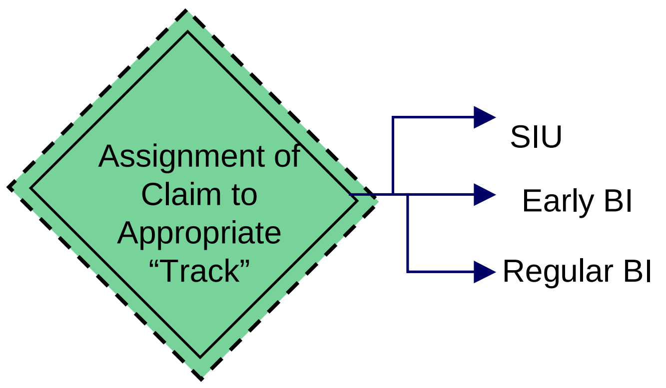
Overview of Generic Claims Process (continued)
(and Potential Areas of Improvement)
Automobile BI

Illustrative

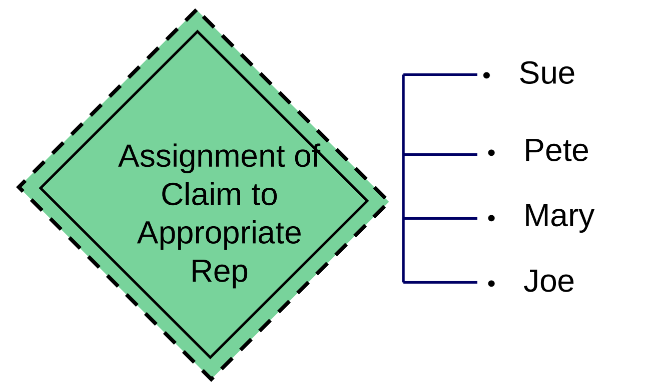


There are at least four areas of potential model application and integration into the BI claims process.

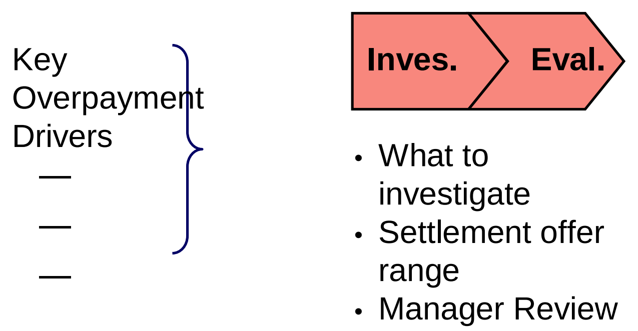
A. Early Triage



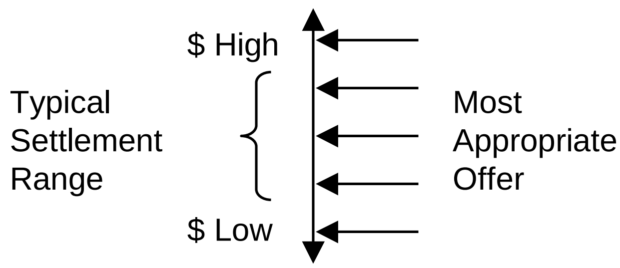
B. Rep Assignment



C. Investigation and Evaluation

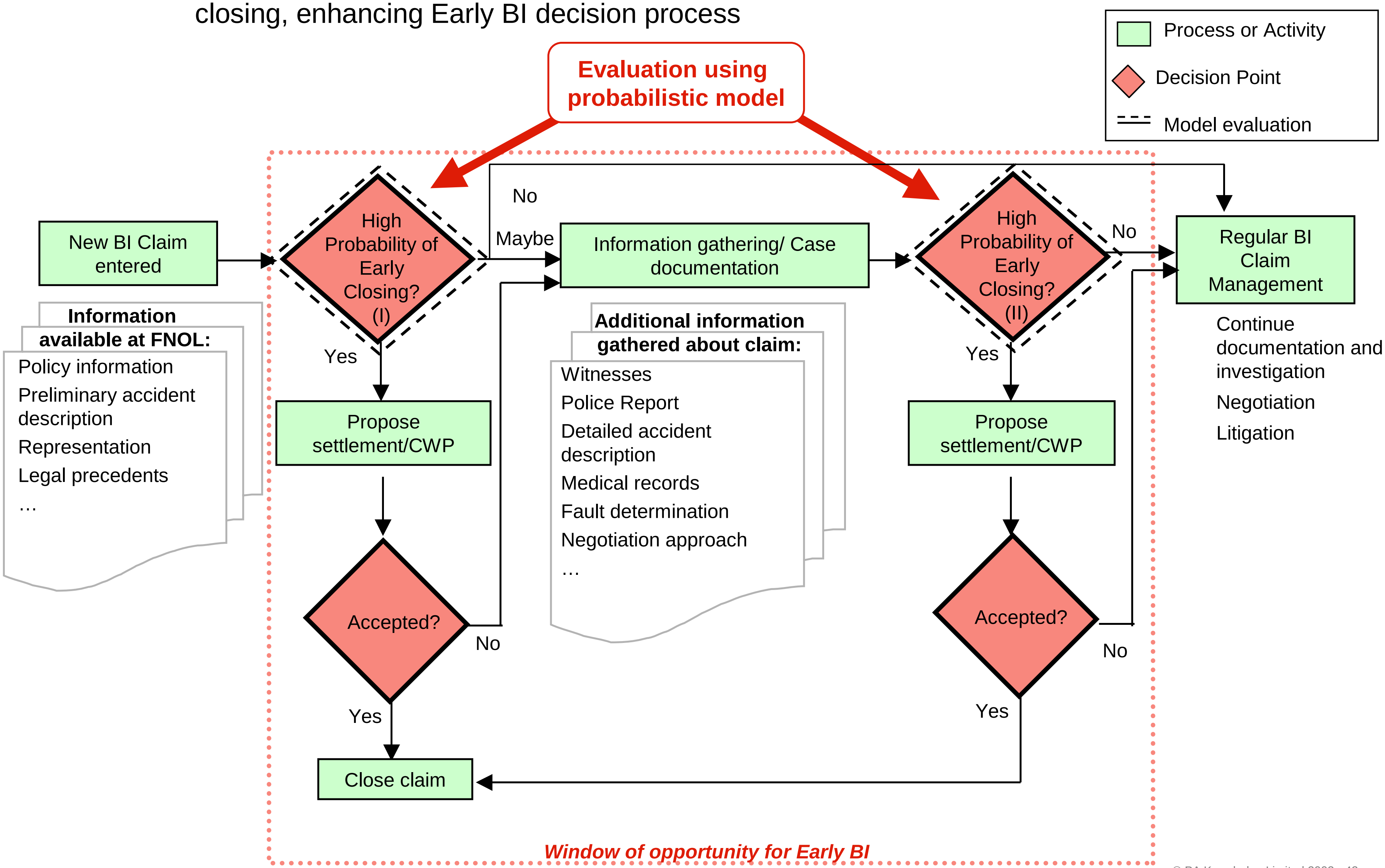


D. Determination of Offer



A. Early Triage

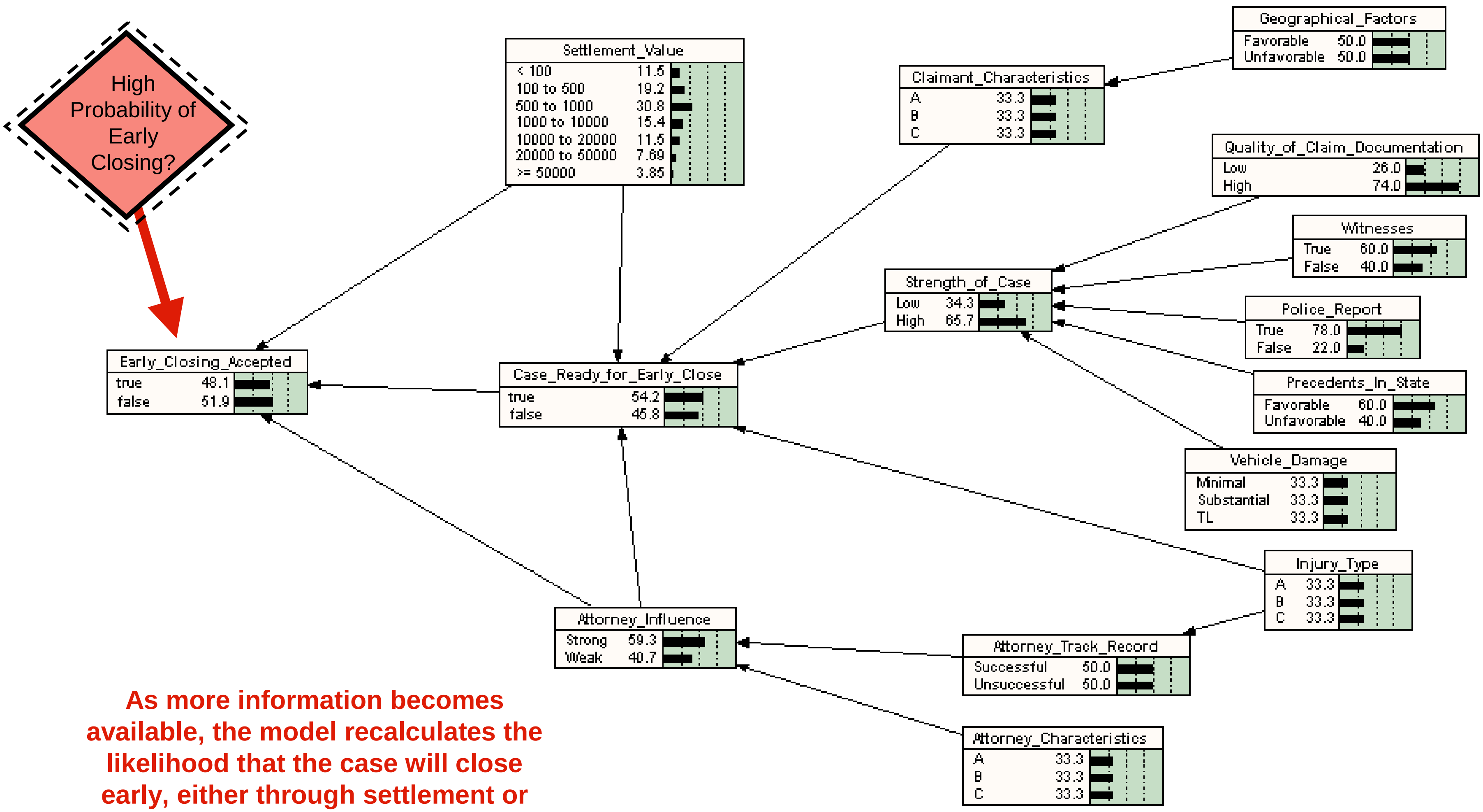
A probabilistic model is used at the early stages of the claim to determine likelihood of early closing, enhancing Early BI decision process



A. Early Triage – Cont.

Data inputs are combined to determine the likelihood of closing as more information is collected

Illustrative



As more information becomes available, the model recalculates the likelihood that the case will close early, either through settlement or through CWP